MATH C285: LINEAR ALGEBRA AND DIFFRENTIAL EQUATIONS (TRAN)



Chau D. Tran
Coastline College

LibreTexts™

Math C285: Linear Algebra and Diffrential Equations (Tran)

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (https://LibreTexts.org) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide and accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by NICE CXone Expert and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptions contact info@LibreTexts.org or visit our main website at https://LibreTexts.org.

This text was compiled on 06/10/2025

TABLE OF CONTENTS

Licensing

1: Systems of Equations

- 1.1: Systems of Equations, Geometry
- 1.2: Systems of Equations, Algebraic Procedures
- 1.3: Gaussian Elimination
- 1.4: Uniqueness of the Reduced Row-Echelon Form
- 1.5: Rank and Homogeneous Systems
- 1.6: Balancing Chemical Reactions
- 1.7: Dimensionless Variables
- 1.8: An Application to Resistor Networks
- 1.9: Exercises

2: Systems of Equations

- 2.1: Systems of Equations, Geometry
- 2.2: Systems of Equations, Algebraic Procedures
- 2.3: Gaussian Elimination
- 2.4: Uniqueness of the Reduced Row-Echelon Form
- 2.5: Rank and Homogeneous Systems
- 2.6: Balancing Chemical Reactions
- 2.7: Dimensionless Variables
- 2.8: An Application to Resistor Networks
- 2.9: Exercises

3: Determinants

- 3.1: Basic Techniques
- 3.2: Properties of Determinants
- 3.3: Finding Determinants using Row Operations
- 3.4: Applications of the Determinant
- 3.E: Exercises

4: R^n

- 4.1: Vectors in R^n
- 4.2: Vector Algebra
- 4.3: Geometric Meaning of Vector Addition
- 4.4: Length of a Vector
- 4.5: Geometric Meaning of Scalar Multiplication
- 4.6: Parametric Lines
- 4.7: The Dot Product
- 4.8: Planes in R^n
- 4.9: The Cross Product
- 4.10: Spanning, Linear Independence and Basis in R^n
- 4.11: Orthogonality
- 4.12: Applications
- 4.E: Exercises

5: Linear Transformations

- 5.1: Linear Transformations
- 5.2: The Matrix of a Linear Transformation I
- 5.3: Properties of Linear Transformations
- 5.4: Special Linear Transformations in R^2
- 5.5: One-to-One and Onto Transformations
- 5.6: Isomorphisms
- 5.7: The Kernel and Image of A Linear Map
- 5.8: The Matrix of a Linear Transformation II
- 5.9: The General Solution of a Linear System
- 5.E: Exercises

6: Complex Numbers

- 6.1: Complex Numbers
- 6.2: Polar Form
- 6.3: Roots of Complex Numbers
- 6.4: The Quadratic Formula
- 6.E: Exercises

7: Spectral Theory

- 7.1: Eigenvalues and Eigenvectors of a Matrix
- 7.2: Diagonalization
- 7.3: Applications of Spectral Theory
- 7.4: Orthogonality
- 7.E: Exercises

8: Some Curvilinear Coordinate Systems

- 8.1: Polar Coordinates and Polar Graphs
- 8.2: Spherical and Cylindrical Coordinates
- 8.E: Exercises

9: Vector Spaces

- 9.1: Algebraic Considerations
- 9.2: Spanning Sets
- 9.3: Linear Independence
- 9.4: Subspaces and Basis
- 9.5: Sums and Intersections
- 9.6: Linear Transformations
- 9.7: Isomorphisms
- 9.8: The Kernel and Image of a Linear Map
- 9.9: The Matrix of a Linear Transformation
- 9.E: Exercises

10: Some Prerequisite Topics

- 10.1: Sets and Set Notation
- 10.2: Well Ordering and Induction

11: Introduction

- 11.1: Notes About These Notes
- 11.2: Introduction to Differential Equations
- 11.3: Classification of Differential Equations
- 11.4: First order ODEs
- 11.E: Introduction (Exercises)

12: First order ODEs

- 12.1: Integrals as solutions
- 12.2: Slope fields
- 12.3: Separable Equations
- 12.4: Linear equations and the integrating factor
- 12.5: Existence and Uniqueness of Solutions of Nonlinear Equations
  - 12.5E: Existence and Uniqueness of Solutions of Nonlinear Equations (Exercises)
- 12.6: Substitution
- 12.7: Autonomous equations
- 12.8: Numerical methods- Euler's method
- 12.9: Exact Equations
- 12.10: First Order Linear PDE
- 12.E: First order ODEs (Exercises)

13: Applications of First Order Equations

- 13.1: Growth and Decay
  - 13.1E: Growth and Decay (Exercises)
- 13.2: Cooling and Mixing
  - 13.2E: Cooling and Mixing (Exercises)
- 13.3: Elementary Mechanics
  - 13.3E: Elementary Mechanics (Exercises)

14: Higher order linear ODEs

- 14.1: Second order linear ODEs
- 14.2: The Method of Undetermined Coefficients I
  - 14.2E: The Method of Undetermined Coefficients I (Exercises)
- 14.3: The Method of Undetermined Coefficients II
  - 14.3E: The Method of Undetermined Coefficients II (Exercises)
- 14.4: Constant coefficient second order linear ODEs
- 14.5: Higher order linear ODEs
- 14.6: Reduction of Order
  - 14.6E: Reduction of Order (Exercises)
- 14.7: Variation of Parameters
  - 14.7E: Variation of Parameters (Exercises)
- 14.8: Mechanical Vibrations
- 14.9: Nonhomogeneous Equations
- 14.10: Forced Oscillations and Resonance
- 14.E: Higher order linear ODEs (Exercises)

15: Fourier series and PDEs

- 15.1: Boundary value problems
- 15.2: The Trigonometric Series
- 15.3: More on the Fourier Series
- 15.4: Sine and Cosine Series
- 15.5: Applications of Fourier Series
- 15.6: PDEs, Separation of Variables, and The Heat Equation
- 15.7: One Dimensional Wave Equation
- 15.8: D'Alembert Solution of The Wave Equation
- 15.9: Steady State Temperature and the Laplacian
- 15.10: Dirichlet Problem in the Circle and the Poisson Kernel
- 15.E: Fourier Series and PDEs (Exercises)

16: Power series methods

- 16.1: Power Series
- 16.2: Series Solutions of Linear Second Order ODEs
- 16.3: Singular Points and the Method of Frobenius
- 16.E: Power series methods (Exercises)

17: Nonlinear Systems

- 17.1: Linearization, Critical Points, and Equilibria
- 17.2: Stability and Classification of Isolated Critical Points
- 17.3: Applications of Nonlinear Systems
- 17.4: Limit cycles
- 17.5: Chaos
- 17.6: Appendix B- Table of Laplace Transforms
- 17.E: Nonlinear Equations (Exercises)

Index

Glossary

Detailed Licensing

Detailed Licensing

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

1: Systems of Equations

1.1: Systems of Equations, Geometry
1.2: Systems of Equations, Algebraic Procedures
1.3: Gaussian Elimination
1.4: Uniqueness of the Reduced Row-Echelon Form
1.5: Rank and Homogeneous Systems
1.6: Balancing Chemical Reactions
1.7: Dimensionless Variables
1.8: An Application to Resistor Networks
1.9: Exercises

This page titled 1: Systems of Equations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

1.1: Systems of Equations, Geometry

Systems of Equations, Geometry

Outcomes

A. Relate the types of solution sets of a system of two (three) variables to the intersections of lines in a plane (the intersection of planes in three space)

As you may remember, linear equations like $2x+3y=6$ can be graphed as straight lines in the coordinate plane. We say that this equation is in two variables, in this case x and y. Suppose you have two such equations, each of which can be graphed as a straight line, and consider the resulting graph of two lines. What would it mean if there exists a point of intersection between the two lines? This point, which lies on *both* graphs, gives x and y values for which both equations are true. In other words, this point gives the ordered pair (x,y) that satisfy both equations. If the point (x,y) is a point of intersection, we say that (x,y) is a **solution** to the two equations. In linear algebra, we often are concerned with finding the solution(s) to a system of equations, if such solutions exist. First, we consider graphical representations of solutions and later we will consider the algebraic methods for finding solutions.

When looking for the intersection of two lines in a graph, several situations may arise. The following picture demonstrates the possible situations when considering two equations (two lines in the graph) involving two variables.

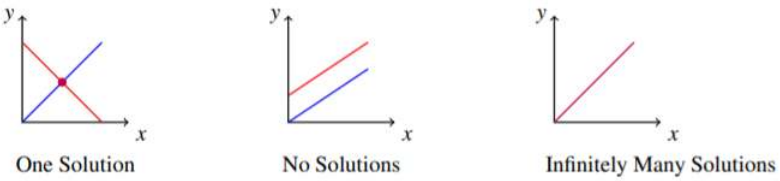


Figure 1.1.1

In the first diagram, there is a unique point of intersection, which means that there is only one (unique) solution to the two equations. In the second, there are no points of intersection and no solution. When no solution exists, this means that the two lines are parallel and they never intersect. The third situation which can occur, as demonstrated in diagram three, is that the two lines are really the same line. For example, $x+y=1$ and $2x+2y=2$ are equations which when graphed yield the same line. In this case there are infinitely many points which are solutions of these two equations, as every ordered pair which is on the graph of the line satisfies both equations. When considering linear systems of equations, there are always three types of solutions possible; exactly one (unique) solution, infinitely many solutions, or no solution.

Example 1.1.1: A Graphical Solution

Use a graph to find the solution to the following system of equations

$$\begin{aligned} x+y&=3\\ y-x&=5 \end{aligned}$$

Solution

Through graphing the above equations and identifying the point of intersection, we can find the solution(s). Remember that we must have either one solution, infinitely many, or no solutions at all. The following graph shows the two equations, as well as the intersection. Remember, the point of intersection represents the solution of the two equations, or the (x,y) which satisfy both equations. In this case, there is one point of intersection at $(-1,4)$ which means we have one unique solution, $x=-1, y=4$.

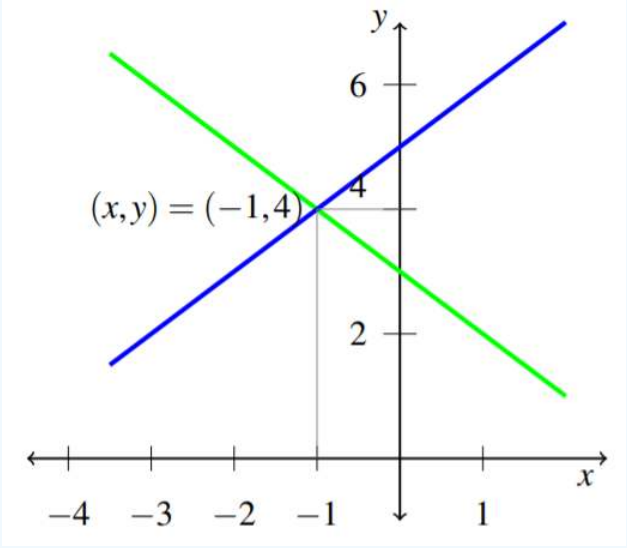


Figure 1.1.2

In the above example, we investigated the intersection point of two equations in two variables, x and y. Now we will consider the graphical solutions of three equations in two variables.

Consider a system of three equations in two variables. Again, these equations can be graphed as straight lines in the plane, so that the resulting graph contains three straight lines. Recall the three possible types of solutions; no solution, one solution, and infinitely many solutions. There are now more complex ways of achieving these situations, due to the presence of the third line. For example, you can imagine the case of three intersecting lines having no common point of intersection. Perhaps you can also imagine three intersecting lines which do intersect at a single point. These two situations are illustrated below.

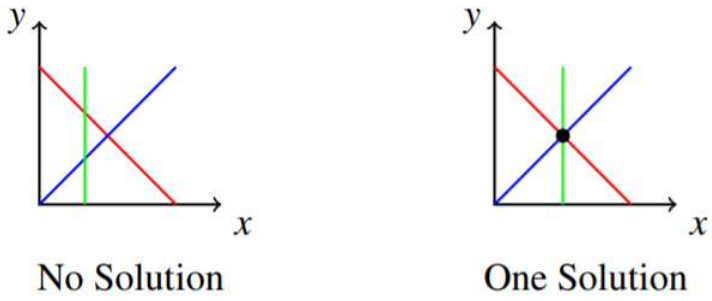


Figure 1.1.3

Consider the first picture above. While all three lines intersect with one another, there is no common point of intersection where all three lines meet at one point. Hence, there is no solution to the three equations. Remember, a solution is a point (x, y) which satisfies **all** three equations. In the case of the second picture, the lines intersect at a common point. This means that there is one solution to the three equations whose graphs are the given lines. You should take a moment now to draw the graph of a system which results in three parallel lines. Next, try the graph of three identical lines. Which type of solution is represented in each of these graphs?

We have now considered the graphical solutions of systems of two equations in two variables, as well as three equations in two variables. However, there is no reason to limit our investigation to equations in two variables. We will now consider equations in three variables.

Below is a video on solving systems of equations graphically



You may recall that equations in three variables, such as $2x+4y-5z=8$, form a plane. Above, we were looking for intersections of lines in order to identify any possible solutions. When graphically solving systems of equations in three variables, we look for intersections of planes. These points of intersection give the (x,y,z) that satisfy all the equations in the system. What types of solutions are possible when working with three variables? Consider the following picture involving two planes, which are given by two equations in three variables.

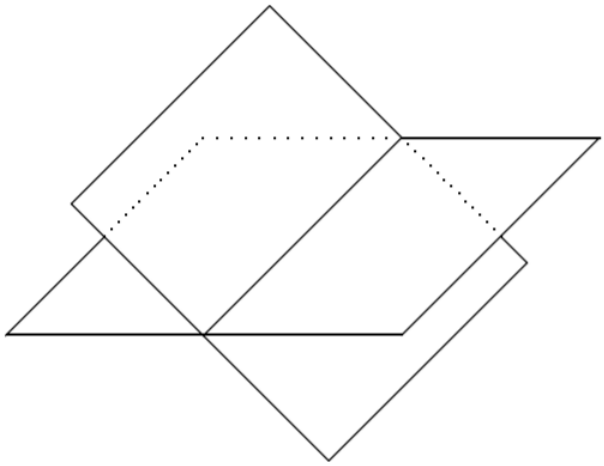


Figure 1.1.4

Notice how these two planes intersect in a line. This means that the points (x,y,z) on this line satisfy both equations in the system. Since the line contains infinitely many points, this system has infinitely many solutions.

It could also happen that the two planes fail to intersect. However, is it possible to have two planes intersect at a single point? Take a moment to attempt drawing this situation, and convince yourself that it is not possible! This means that when we have only two equations in three variables, there is no way to have a unique solution! Hence, the types of solutions possible for two equations in three variables are no solution or infinitely many solutions.

Now imagine adding a third plane. In other words, consider three equations in three variables. What types of solutions are now possible? Consider the following diagram.

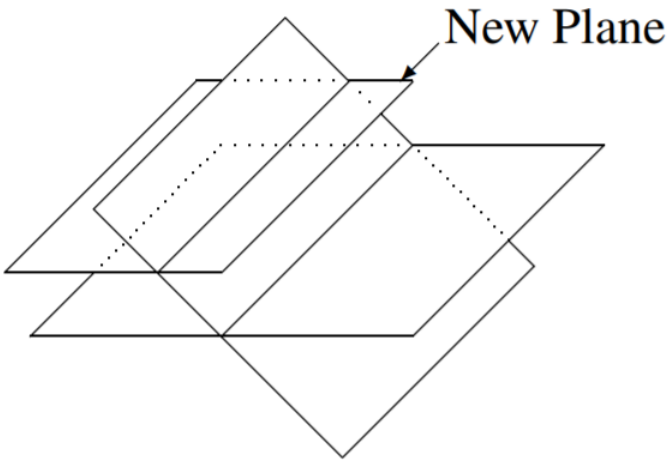


Figure 1.1.5

In this diagram, there is no point which lies in all three planes. There is no intersection between **all** planes so there is no solution. The picture illustrates the situation in which the line of intersection of the new plane with one of the original planes forms a line parallel to the line of intersection of the first two planes. However, in three dimensions, it is possible for two lines to fail to intersect even though they are not parallel. Such lines are called **skew lines**.

Recall that when working with two equations in three variables, it was not possible to have a unique solution. Is it possible when considering three equations in three variables? In fact, it is possible, and we demonstrate this situation in the following picture.

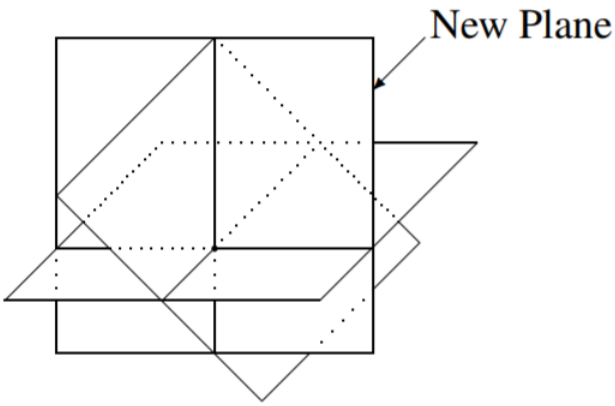


Figure 1.1.6

In this case, the three planes have a single point of intersection. Can you think of other types of solutions possible? Another is that the three planes could intersect in a line, resulting in infinitely many solutions, as in the following diagram.

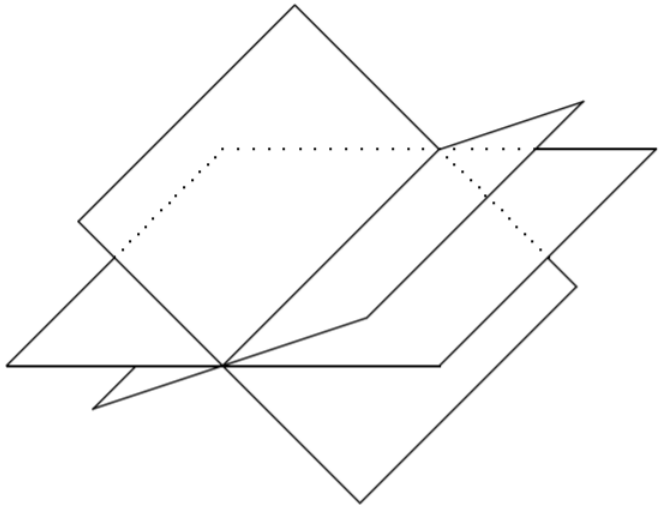


Figure 1.1.7

We have now seen how three equations in three variables can have no solution, a unique solution, or intersect in a line resulting in infinitely many solutions. It is also possible that the three equations graph the same plane, which also leads to infinitely many solutions.

You can see that when working with equations in three variables, there are many more ways to achieve the different types of solutions than when working with two variables. It may prove enlightening to spend time imagining (and drawing) many possible scenarios, and you should take some time to try a few.

You should also take some time to imagine (and draw) graphs of systems in more than three variables. Equations like $x+y-2z+4w=8$ with more than three variables are often called **hyper-planes**. You may soon realize that it is tricky to draw the graphs of hyper-planes! Through the tools of linear algebra, we can algebraically examine these types of systems which are difficult to graph. In the following section, we will consider these algebraic tools.

More Videos:

Solving Systems of Equations Graphically
Solving Systems of Equations by Graphing
Determining the Intersection of Two Graph on the TI83/84
Ex: Identify the Solution to a System of Equation Given a Graph, Then Verify
olve a System of Linear Equations by Graphing (One Solution)
Solve a System of Linear Equations by Graphing (No Solution)
x 1: Solve a System of Equations by Graphing
Ex 2: Solve a System of Equations by Graphing
Ex: Solve a System of Equations by Graphing (Infinite Solutions)
Ex: Solve a System of Equations by Graphing (No Solution)

? Exercise 1.1.1

Graphically, find the point (x_1, y_1) which lies on both lines, $x+3y=1$ and $4x-y=3$. That is, graph each line and see where they intersect.

? Exercise 1.1.2

Graphically, find the point of intersection of the two lines, $3x+y=3$ and $x+2y=1$. That is, graph each line and see where they intersect

Exercise 1.1.3

You have a system of k equations in two variables, $k \geq 2$. Explain the geometric significance of

a. No solution.
b. A unique solution.
c. An infinite number of solutions.

This page titled 1.1: Systems of Equations, Geometry is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.1: Systems of Equations, Geometry** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

1.2: Systems of Equations, Algebraic Procedures

Algebraic Procedures

Outcomes

A. Use elementary operations to find the solution to a linear system of equations.
B. Find the row-echelon form and reduced row-echelon form of a matrix.
C. Determine whether a system of linear equations has no solution, a unique solution or an infinite number of solutions from its .
D. Solve a system of equations using Gaussian Elimination and Gauss-Jordan Elimination.
E. Model a physical system with linear equations and then solve.

We have taken an in depth look at graphical representations of systems of equations, as well as how to find possible solutions graphically. Our attention now turns to working with systems algebraically.

Definition 1.2.1: System of Linear Equations

A **system of linear equations** is a list of equations,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

where a_{ij} and b_j are real numbers. The above is a system of m equations in the n variables, $x_1, x_2 \cdots, x_n$. Written more simply in terms of summation notation, the above can be written in the form

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \ i = 1, 2, 3, \cdots, m$$

The relative size of m and n is not important here. Notice that we have allowed a_{ij} and b_j to be any real number. We can also call these numbers **scalars** . We will use this term throughout the text, so keep in mind that the term **scalar** just means that we are working with real numbers.

Now, suppose we have a system where $b_i = 0$ for all i. In other words every equation equals 0. This is a special type of system.

Definition 1.2.2: Homogeneous System of Equations

A system of equations is called **homogeneous** if each equation in the system is equal to 0. A homogeneous system has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

where a_{ij} are scalars and x_i are variables.

Recall from the previous section that our goal when working with systems of linear equations was to find the point of intersection of the equations when graphed. In other words, we looked for the solutions to the system. We now wish to find these solutions algebraically. We want to find values for $x_1, \cdots, x_n$ which solve all of the equations. If such a set of values exists, we call $(x_1, \cdots, x_n)$ the **solution set**.

Recall the above discussions about the types of solutions possible. We will see that systems of linear equations will have one unique solution, infinitely many solutions, or no solution. Consider the following definition.

Definition 1.2.3: Consistent and Inconsistent Systems

A system of linear equations is called **consistent** if there exists at least one solution. It is called **inconsistent** if there is no solution.

If you think of each equation as a condition which must be satisfied by the variables, consistent would mean there is some choice of variables which can satisfy **all** the conditions. Inconsistent would mean there is no choice of the variables which can satisfy all of the conditions.

The following sections provide methods for determining if a system is consistent or inconsistent, and finding solutions if they exist.

Elementary Operations

We begin this section with an example. Recall from Example 1.1.1 that the solution to the given system was $(x, y) = (-1, 4)$.

Example 1.2.1: Verifying an Ordered Pair is a Solution

Algebraically verify that $(x, y) = (-1, 4)$ is a solution to the following system of equations.

$$\begin{aligned} x + y &= 3 \\ y - x &= 5 \end{aligned}$$

Solution

By graphing these two equations and identifying the point of intersection, we previously found that $(x, y) = (-1, 4)$ is the unique solution.

We can verify algebraically by substituting these values into the original equations, and ensuring that the equations hold. First, we substitute the values into the first equation and check that it equals 3.

$$x + y = (-1) + (4) = 3$$

This equals 3 as needed, so we see that $(-1, 4)$ is a solution to the first equation. Substituting the values into the second equation yields

$$y - x = (4) - (-1) = 4 + 1 = 5$$

which is true. For $(x, y) = (-1, 4)$ each equation is true and therefore, this is a solution to the system.

Now, the interesting question is this: If you were not given these numbers to verify, how could you algebraically determine the solution? Linear algebra gives us the tools needed to answer this question. The following basic operations are important tools that we will utilize.

Definition 1.2.4: Elementary Operations

Elementary operations are those operations consisting of the following.

1. Interchange the order in which the equations are listed.
2. Multiply any equation by a nonzero number.
3. Replace any equation with itself added to a multiple of another equation.

It is important to note that none of these operations will change the set of solutions of the system of equations. In fact, elementary operations are the *key tool* we use in linear algebra to find solutions to systems of equations.

Consider the following example.

Example 1.2.2: Effects of an Elementary Operation

Show that the system

$$\begin{aligned} x+y&=7\\ 2x-y&=8 \end{aligned}$$

has the same solution as the system

$$\begin{aligned} x+y&=7\\ -3y&=-6 \end{aligned}$$

Solution

Notice that the second system has been obtained by taking the second equation of the first system and adding -2 times the first equation, as follows:

$$2x-y+(-2)(x+y)=8+(-2)(7)$$

By simplifying, we obtain

$$-3y=-6$$

which is the second equation in the second system. Now, from here we can solve for y and see that $y=2$. Next, we substitute this value into the first equation as follows

$$x+y=x+2=7$$

Hence $x=5$ and so $(x,y)=(5,2)$ is a solution to the second system. We want to check if $(5,2)$ is also a solution to the first system. We check this by substituting $(x,y)=(5,2)$ into the system and ensuring the equations are true.

$$\begin{aligned} x+y&=(5)+(2)=7\\ 2x-y&=2\,(5)-(2)=8 \end{aligned}$$

Hence, $(5,2)$ is also a solution to the first system.

This example illustrates how an elementary operation applied to a system of two equations in two variables does not affect the solution set. However, a linear system may involve many equations and many variables and there is no reason to limit our study to small systems. For any size of system in any number of variables, the solution set is still the collection of solutions to the equations. In every case, the above operations of Definition 1.2.4 do not change the set of solutions to the system of linear equations.

In the following theorem, we use the notation E_i to represent an equation, while b_i denotes a constant.

Theorem 1.2.1: Elementary Operations and Solutions

Suppose you have a system of two linear equations

$$\begin{aligned} E_1&=b_1\\ E_2&=b_2 \end{aligned} \tag{1.2.1}$$

Then the following systems have the same solution set as (1.2.1):

1. $$\begin{aligned} E_2&=b_2\\ E_1&=b_1 \end{aligned} \tag{1.2.2}$$

2. $$\begin{aligned} E_1&=b_1\\ kE_2&=kb_2 \end{aligned} \tag{1.2.3}$$

   for any scalar k, provided $k\neq 0$.

3. $$\begin{aligned} E_1&=b_1\\ E_2+kE_1&=b_2+kb_1 \end{aligned} \tag{1.2.4}$$

   for any scalar k (including $k=0$).

Before we proceed with the proof of Theorem 1.2.1, let us consider this theorem in context of Example 1.2.2. Then,

$$\begin{aligned} E_1 &= x + y, \quad & b_1 &= 7 \\ E_2 &= 2x - y, \quad & b_2 &= 8 \end{aligned}$$

Recall the elementary operations that we used to modify the system in the solution to the example. First, we added (-2) times the first equation to the second equation. In terms of Theorem 1.2.1, this action is given by

$$E_2 + (-2)\, E_1 = b_2 + (-2)\, b_1$$

or

$$2x - y + (-2)\,(x + y) = 8 + (-2)\,7$$

This gave us the second system in Example 1.2.2, given by

$$\begin{aligned} E_1 &= b_1 \\ E_2 + (-2)\, E_1 &= b_2 + (-2)\, b_1 \end{aligned}$$

From this point, we were able to find the solution to the system. Theorem 1.2.1 tells us that the solution we found is in fact a solution to the original system.

We will now prove Theorem 1.2.1.

Proof

1. The proof that the systems (1.2.1) and (1.2.2) have the same solution set is as follows. Suppose that $(x_1, \cdots, x_n)$ is a solution to $E_1 = b_1, E_2 = b_2$. We want to show that this is a solution to the system in (1.2.2) above. This is clear, because the system in (1.2.2) is the original system, but listed in a different order. Changing the order does not effect the solution set, so $(x_1, \cdots, x_n)$ is a solution to (1.2.2).
2. Next we want to prove that the systems (1.2.1) and (1.2.3) have the same solution set. That is $E_1 = b_1, E_2 = b_2$ has the same solution set as the system $E_1 = b_1, kE_2 = kb_2$ provided $k \neq 0$. Let $(x_1, \cdots, x_n)$ be a solution of $E_1 = b_1, E_2 = b_2,$. We want to show that it is a solution to $E_1 = b_1, kE_2 = kb_2$. Notice that the only difference between these two systems is that the second involves multiplying the equation, $E_2 = b_2$ by the scalar k. Recall that when you multiply both sides of an equation by the same number, the sides are still equal to each other. Hence if $(x_1, \cdots, x_n)$ is a solution to $E_2 = b_2$, then it will also be a solution to $kE_2 = kb_2$. Hence, $(x_1, \cdots, x_n)$ is also a solution to (1.2.3).
   Similarly, let $(x_1, \cdots, x_n)$ be a solution of $E_1 = b_1, kE_2 = kb_2$. Then we can multiply the equation $kE_2 = kb_2$ by the scalar $1/k$, which is possible only because we have required that $k \neq 0$. Just as above, this action preserves equality and we obtain the equation $E_2 = b_2$. Hence $(x_1, \cdots, x_n)$ is also a solution to $E_1 = b_1, E_2 = b_2$.
3. Finally, we will prove that the systems (1.2.1) and (1.2.4) have the same solution set. We will show that any solution of $E_1 = b_1, E_2 = b_2$ is also a solution of (1.2.4). Then, we will show that any solution of (1.2.4) is also a solution of $E_1 = b_1, E_2 = b_2$. Let $(x_1, \cdots, x_n)$ be a solution to $E_1 = b_1, E_2 = b_2$. Then in particular it solves $E_1 = b_1$. Hence, it solves the first equation in (1.2.4). Similarly, it also solves $E_2 = b_2$. By our proof of (1.2.3), it also solves $kE_1 = kb_1$. Notice that if we add E_2 and kE_1, this is equal to $b_2 + kb_1$. Therefore, if $(x_1, \cdots, x_n)$ solves $E_1 = b_1, E_2 = b_2$ it must also solve $E_2 + kE_1 = b_2 + kb_1$.
   Now suppose $(x_1, \cdots, x_n)$ solves the system $E_1 = b_1, E_2 + kE_1 = b_2 + kb_1$. Then in particular it is a solution of $E_1 = b_1$. Again by our proof of (1.2.3), it is also a solution to $kE_1 = kb_1$. Now if we subtract these equal quantities from both sides of $E_2 + kE_1 = b_2 + kb_1$ we obtain $E_2 = b_2$, which shows that the solution also satisfies $E_1 = b_1, E_2 = b_2$.

Stated simply, the above theorem shows that the elementary operations do not change the solution set of a system of equations.

We will now look at an example of a system of three equations and three variables. Similarly to the previous examples, the goal is to find values for x, y, z such that each of the given equations are satisfied when these values are substituted in.

Example 1.2.3: Solving a System of Equations with Elementary Operations

Find the solutions to the system,

$$\begin{aligned} x+3y+6z &= 25 \\ 2x+7y+14z &= 58 \\ 2y+5z &= 19 \end{aligned} \tag{1.2.5}$$

Solution

We can relate this system to Theorem 1.2.1 above. In this case, we have

$$\begin{aligned} E_1 &= x+3y+6z, & b_1 &= 25 \\ E_2 &= 2x+7y+14z, & b_2 &= 58 \\ E_3 &= 2y+5z, & b_3 &= 19 \end{aligned}$$

Theorem 1.2.1 claims that if we do elementary operations on this system, we will not change the solution set. Therefore, we can solve this system using the elementary operations given in Definition 1.2.4. First, replace the second equation by (-2) times the first equation added to the second. This yields the system

$$\begin{aligned} x+3y+6z &= 25 \\ y+2z &= 8 \\ 2y+5z &= 19 \end{aligned} \tag{1.2.6}$$

Now, replace the third equation with (-2) times the second added to the third. This yields the system

$$\begin{aligned} x+3y+6z &= 25 \\ y+2z &= 8 \\ z &= 3 \end{aligned} \tag{1.2.7}$$

At this point, we can easily find the solution. Simply take $z=3$ and substitute this back into the previous equation to solve for y, and similarly to solve for x.

$$\begin{aligned} x+3y+6\left(3\right) = x+3y+18 &= 25 \\ y+2\left(3\right) = y+6 &= 8 \\ z &= 3 \end{aligned}$$

The second equation is now

$$y+6=8$$

You can see from this equation that $y=2$. Therefore, we can substitute this value into the first equation as follows:

$$x+3\left(2\right)+18=25$$

By simplifying this equation, we find that $x=1$. Hence, the solution to this system is $(x,y,z)=(1,2,3)$. This process is called **back substitution**.

Alternatively, in (1.2.7) you could have continued as follows. Add (-2) times the third equation to the second and then add (-6) times the second to the first. This yields

$$\begin{aligned} x+3y &= 7 \\ y &= 2 \\ z &= 3 \end{aligned}$$

Now add (-3) times the second to the first. This yields

$$\begin{aligned} x &= 1 \\ y &= 2 \\ z &= 3 \end{aligned}$$

a system which has the same solution set as the original system. This avoided back substitution and led to the same solution set. It is your decision which you prefer to use, as both methods lead to the correct solution, $(x, y, z) = (1, 2, 3)$.

This page titled 1.2: Systems of Equations, Algebraic Procedures is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.2: Systems of Equations, Algebraic Procedures** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

1.3: Gaussian Elimination

The work we did in the previous section will always find the solution to the system. In this section, we will explore a less cumbersome way to find the solutions. First, we will represent a linear system with an **augmented matrix.** A **matrix** is simply a rectangular array of numbers. The size or dimension of a matrix is defined as $m \times n$ where m is the number of rows and n is the number of columns. In order to construct an augmented matrix from a linear system, we create a **coefficient matrix** from the coefficients of the variables in the system, as well as a **constant matrix** from the constants. The coefficients from one equation of the system create one row of the augmented matrix.

For example, consider the linear system in Example 1.2.3

$$\begin{aligned} x+3y+6z &= 25 \\ 2x+7y+14z &= 58 \\ 2y+5z &= 19 \end{aligned}$$

This system can be written as an augmented matrix, as follows

$$\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \\ 2 & 7 & 14 & 58 \\ 0 & 2 & 5 & 19 \end{array}\right]$$

Notice that it has exactly the same information as the original system. Here it is understood that the first column contains the coefficients from x in each equation, in order, $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$. Similarly, we create a column from the coefficients on y in each equation, $\begin{bmatrix} 3 \\ 7 \\ 2 \end{bmatrix}$ and a column from the coefficients on z in each equation, $\begin{bmatrix} 6 \\ 14 \\ 5 \end{bmatrix}$. For a system of more than three variables, we would continue in this way constructing a column for each variable. Similarly, for a system of less than three variables, we simply construct a column for each variable.

Finally, we construct a column from the constants of the equations, $\begin{bmatrix} 25 \\ 58 \\ 19 \end{bmatrix}$.

The rows of the augmented matrix correspond to the equations in the system. For example, the top row in the augmented matrix, $\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \end{array}\right]$ corresponds to the equation

$$x+3y+6z=25.$$

Consider the following definition.

Definition 1.3.1: Augmented Matrix of a Linear System

For a linear system of the form

$$\begin{aligned} a_{11}x_1+\cdots+a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n &= b_m \end{aligned}$$

where the x_i are variables and the a_{ij} and b_i are constants, the augmented matrix of this system is given by

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

Now, consider elementary operations in the context of the augmented matrix. The elementary operations in Definition 1.2.4 can be used on the rows just as we used them on equations previously. Changes to a system of equations in as a result of an elementary operation are equivalent to changes in the augmented matrix resulting from the corresponding row operation. Note that Theorem 1.2.1 implies that any elementary row operations used on an augmented matrix will not change the solution to the corresponding system of equations. We now formally define elementary row operations. These are the *key tool* we will use to find solutions to systems of equations.

Definition 1.3.2: Elementary Row Operations

The **elementary row operations** (also known as **row operations**) consist of the following

1. Switch two rows.
2. Multiply a row by a nonzero number.
3. Replace a row by any multiple of another row added to it.

Recall how we solved Example 1.2.3. We can do the exact same steps as above, except now in the context of an augmented matrix and using row operations. The augmented matrix of this system is

$$\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \\ 2 & 7 & 14 & 58 \\ 0 & 2 & 5 & 19 \end{array}\right]$$

Thus the first step in solving the system given by (1.2.5) would be to take (-2) times the first row of the augmented matrix and add it to the second row,

$$\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \\ 0 & 1 & 2 & 8 \\ 0 & 2 & 5 & 19 \end{array}\right]$$

Note how this corresponds to (1.2.6). Next take (-2) times the second row and add to the third,

$$\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

This augmented matrix corresponds to the system

$$\begin{aligned} x+3y+6z &= 25 \\ y+2z &= 8 \\ z &= 3 \end{aligned}$$

which is the same as (1.2.7). By back substitution you obtain the solution $x=1, y=2$, and $z=3$.

Through a systematic procedure of row operations, we can simplify an augmented matrix and carry it to **row-echelon form** or **reduced row-echelon form**, which we define next. These forms are used to find the solutions of the system of equations corresponding to the augmented matrix.

In the following definitions, the term **leading entry** refers to the first nonzero entry of a row when scanning the row from left to right.

Definition 1.3.3: Row-Echelon Form

An augmented matrix is in **row-echelon form** if

1. All nonzero rows are above any rows of zeros.
2. Each leading entry of a row is in a column to the right of the leading entries of any row above it.
3. Each leading entry of a row is equal to 1.

We also consider another reduced form of the augmented matrix which has one further condition.

Definition 1.3.4: Reduced Row-Echelon Form

An augmented matrix is in **reduced row-echelon form** if

1. All nonzero rows are above any rows of zeros.
2. Each leading entry of a row is in a column to the right of the leading entries of any rows above it.
3. Each leading entry of a row is equal to 1.
4. All entries in a column above and below a leading entry are zero.

Notice that the first three conditions on a reduced row-echelon form matrix are the same as those for row-echelon form.

Hence, every reduced row-echelon form matrix is also in row-echelon form. The converse is not necessarily true; we cannot assume that every matrix in row-echelon form is also in reduced row-echelon form. However, it often happens that the row-echelon form is sufficient to provide information about the solution of a system.

The following examples describe matrices in these various forms. As an exercise, take the time to carefully verify that they are in the specified form.

Example 1.3.1: Not in Row-Echelon Form

The following augmented matrices are not in row-echelon form (and therefore also not in reduced row-echelon form).

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 4 & -6 \\ 4 & 0 & 7 \end{array}\right], \left[\begin{array}{ccc|c} 0 & 2 & 3 & 3 \\ 1 & 5 & 0 & 2 \\ 7 & 5 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

Example 1.3.2: Matrices in Row-Echelon Form

The following augmented matrices are in row-echelon form, but not in reduced row-echelon form.

$$\left[\begin{array}{ccccc|c} 1 & 0 & 6 & 5 & 8 & 2 \\ 0 & 0 & 1 & 2 & 7 & 3 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{cccc|c} 1 & 3 & 5 & 4 & 4 \\ 0 & 1 & 0 & 7 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \\ 0 & 0 & 0 & 0 & \end{array}\right], \left[\begin{array}{ccc|c} 1 & 0 & 6 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Notice that we could apply further row operations to these matrices to carry them to reduced row-echelon form. Take the time to try that on your own. Consider the following matrices, which are in reduced row-echelon form.

Example 1.3.3: Matrices in Reduced Row-Echelon Form

The following augmented matrices are in reduced row-echelon form.

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

One way in which the row-echelon form of a matrix is useful is in identifying the pivot positions and pivot columns of the matrix.

Definition 1.3.5: Pivot Position and Pivot Column

A **pivot position** in a matrix is the location of a leading entry in the row-echelon form of a matrix.

A **pivot column** is a column that contains a pivot position.

For example consider the following.

Example 1.3.4: Pivot Position

Let

$$A = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 6 \\ 4 & 4 & 4 & 10 \end{array}\right]$$

Where are the pivot positions and pivot columns of the augmented matrix A?

Solution

The row-echelon form of this matrix is

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

This is all we need in this example, but note that this matrix is not in reduced row-echelon form.

In order to identify the pivot positions in the original matrix, we look for the leading entries in the row-echelon form of the matrix. Here, the entry in the first row and first column, as well as the entry in the second row and second column are the leading entries. Hence, these locations are the pivot positions. We identify the pivot positions in the original matrix, as in the following:

$$\left[\begin{array}{ccc|c} \boxed{1} & 2 & 3 & 4 \\ 3 & \boxed{2} & 1 & 6 \\ 4 & 4 & 4 & 10 \end{array}\right]$$

Thus the pivot columns in the matrix are the first two columns.

The following is an algorithm for carrying a matrix to row-echelon form and reduced row-echelon form. You may wish to use this algorithm to carry the above matrix to row-echelon form or reduced row-echelon form yourself for practice.

Algorithm 1.3.1: Reduced Row-Echelon Form Algorithm

This algorithm provides a method for using row operations to take a matrix to its reduced row-echelon form. We begin with the matrix in its original form.

1. Starting from the left, find the first nonzero column. This is the first pivot column, and the position at the top of this column is the first pivot position. Switch rows if necessary to place a nonzero number in the first pivot position.
2. Use row operations to make the entries below the first pivot position (in the first pivot column) equal to zero.
3. Ignoring the row containing the first pivot position, repeat steps 1 and 2 with the remaining rows. Repeat the process until there are no more rows to modify.
4. Divide each nonzero row by the value of the leading entry, so that the leading entry becomes 1. The matrix will then be in row-echelon form.

   The following step will carry the matrix from row-echelon form to reduced row-echelon form.

5. Moving from right to left, use row operations to create zeros in the entries of the pivot columns which are above the pivot positions. The result will be a matrix in reduced row-echelon form.

Most often we will apply this algorithm to an augmented matrix in order to find the solution to a system of linear equations. However, we can use this algorithm to compute the reduced row-echelon form of any matrix which could be useful in other applications.

Consider the following example of Algorithm 1.3.1.

Example 1.3.5: Finding Row-Echelon Form and Reduced Row-Echelon Form of a Matrix

Let

$$A = \begin{bmatrix} 0 & -5 & -4 \\ 1 & 4 & 3 \\ 5 & 10 & 7 \end{bmatrix}$$

Find the row-echelon form of A. Then complete the process until A is in reduced row-echelon form.

Solution

In working through this example, we will use the steps outlined in Algorithm 1.3.1.

1. The first pivot column is the first column of the matrix, as this is the first nonzero column from the left. Hence the first pivot position is the one in the first row and first column. Switch the first two rows to obtain a nonzero entry in the first pivot position, outlined in a box below.

$$\begin{bmatrix} \boxed{1} & 4 & 3 \\ 0 & -5 & -4 \\ 5 & 10 & 7 \end{bmatrix}$$

2. Step two involves creating zeros in the entries below the first pivot position. The first entry of the second row is already a zero. All we need to do is subtract 5 times the first row from the third row. The resulting matrix is

$$\begin{bmatrix} 1 & 4 & 3 \\ 0 & -5 & -4 \\ 0 & 10 & 8 \end{bmatrix}$$

3. Now ignore the top row. Apply steps 1 and 2 to the smaller matrix

$$\begin{bmatrix} -5 & -4 \\ 10 & 8 \end{bmatrix}$$

In this matrix, the first column is a pivot column, and -5 is in the first pivot position. Therefore, we need to create a zero below it. To do this, add 2 times the first row (of this matrix) to the second. The resulting matrix is

$$\begin{bmatrix} -5 & -4 \\ 0 & 0 \end{bmatrix}$$

Our original matrix now looks like

$$\begin{bmatrix} 1 & 4 & 3 \\ 0 & -5 & -4 \\ 0 & 0 & 0 \end{bmatrix}$$

We can see that there are no more rows to modify.

4. Now, we need to create leading 1s in each row. The first row already has a leading 1 so no work is needed here. Divide the second row by -5 to create a leading 1. The resulting matrix is

$$\begin{bmatrix} 1 & 4 & 3 \\ 0 & 1 & \frac{4}{5} \\ 0 & 0 & 0 \end{bmatrix}$$

This matrix is now in row-echelon form.

5. Now create zeros in the entries above pivot positions in each column, in order to carry this matrix all the way to reduced row-echelon form. Notice that there is no pivot position in the third column so we do not need to create any zeros in this column! The column in which we need to create zeros is the second. To do so, subtract 4 times the second row from the first row. The resulting matrix is

$$\begin{bmatrix} 1 & 0 & -\frac{1}{5} \\ 0 & 1 & \frac{4}{5} \\ 0 & 0 & 0 \end{bmatrix}$$

This matrix is now in reduced row-echelon form.

The above algorithm gives you a simple way to obtain the row-echelon form and reduced row-echelon form of a matrix. The main idea is to do row operations in such a way as to end up with a matrix in row-echelon form or reduced row-echelon form. This process is important because the resulting matrix will allow you to describe the solutions to the corresponding linear system of equations in a meaningful way.

In the next example, we look at how to solve a system of equations using the corresponding augmented matrix.

Example 1.3.6: Finding the Solution to a System

Give the complete solution to the following system of equations

$$\begin{aligned} 2x + 4y - 3z &= -1 \\ 5x + 10y - 7z &= -2 \\ 3x + 6y + 5z &= 9 \end{aligned}$$

Solution

The augmented matrix for this system is

$$\left[\begin{array}{ccc|c} 2 & 4 & -3 & -1 \\ 5 & 10 & -7 & -2 \\ 3 & 6 & 5 & 9 \end{array}\right]$$

In order to find the solution to this system, we wish to carry the augmented matrix to reduced row-echelon form. We will do so using Algorithm 1.3.1. Notice that the first column is nonzero, so this is our first pivot column. The first entry in the first row, 2, is the first leading entry and it is in the first pivot position. We will use row operations to create zeros in the entries below the 2. First, replace the second row with -5 times the first row plus 2 times the second row. This yields

$$\left[\begin{array}{ccc|c} 2 & 4 & -3 & -1 \\ 0 & 0 & 1 & 1 \\ 3 & 6 & 5 & 9 \end{array}\right]$$

Now, replace the third row with -3 times the first row plus to 2 times the third row. This yields

$$\left[\begin{array}{ccc|c} 2 & 4 & -3 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 21 \end{array}\right]$$

Now the entries in the first column below the pivot position are zeros. We now look for the second pivot column, which in this case is column three. Here, the 1 in the second row and third column is in the pivot position. We need to do just one row operation to create a zero below the 1.

Taking -1 times the second row and adding it to the third row yields

$$\left[\begin{array}{ccc|c} 2 & 4 & -3 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 20 \end{array}\right]$$

We could proceed with the algorithm to carry this matrix to row-echelon form or reduced row-echelon form. However, remember that we are looking for the solutions to the system of equations. Take another look at the third row of the matrix. Notice that it corresponds to the equation

$$0x+0y+0z=20$$

There is no solution to this equation because for all x, y, z, the left side will equal 0 and $0 \neq 20$. This shows there is no solution to the given system of equations. In other words, this system is inconsistent.

The following is another example of how to find the solution to a system of equations by carrying the corresponding augmented matrix to reduced row-echelon form.

Example 1.3.7: An Infinite Set of Solutions

Give the complete solution to the system of equations

$$\begin{aligned} 3x-y-5z&=9 \\ y-10z&=0 \\ -2x+y&=-6 \end{aligned} \tag{1.3.1}$$

Solution

The augmented matrix of this system is

$$\left[\begin{array}{rrr|r} 3 & -1 & -5 & 9 \\ 0 & 1 & -10 & 0 \\ -2 & 1 & 0 & -6 \end{array}\right]$$

In order to find the solution to this system, we will carry the augmented matrix to reduced row-echelon form, using Algorithm 1.3.1. The first column is the first pivot column. We want to use row operations to create zeros beneath the first entry in this column, which is in the first pivot position. Replace the third row with 2 times the first row added to 3 times the third row. This gives

$$\left[\begin{array}{rrr|r} 3 & -1 & -5 & 9 \\ 0 & 1 & -10 & 0 \\ 0 & 1 & -10 & 0 \end{array}\right]$$

Now, we have created zeros beneath the 3 in the first column, so we move on to the second pivot column (which is the second column) and repeat the procedure. Take -1 times the second row and add to the third row.

$$\left[\begin{array}{rrr|r} 3 & -1 & -5 & 9 \\ 0 & 1 & -10 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The entry below the pivot position in the second column is now a zero. Notice that we have no more pivot columns because we have only two leading entries.

At this stage, we also want the leading entries to be equal to one. To do so, divide the first row by 3.

$$\left[\begin{array}{rrr|r} 1 & -\frac{1}{3} & -\frac{5}{3} & 3 \\ 0 & 1 & -10 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

This matrix is now in row-echelon form.

Let's continue with row operations until the matrix is in reduced row-echelon form. This involves creating zeros above the pivot positions in each pivot column. This requires only one step, which is to add $\frac{1}{3}$ times the second row to the first row.

$$\left[\begin{array}{ccc|c} 1 & 0 & -5 & 3 \\ 0 & 1 & -10 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

This is in reduced row-echelon form, which you should verify using Definition 1.3.4. The equations corresponding to this reduced row-echelon form are

$$x - 5z = 3$$
$$y - 10z = 0$$

or

$$x = 3 + 5z$$
$$y = 10z$$

Observe that z is not restrained by any equation. In fact, z can equal any number. For example, we can let $z = t$, where we can choose t to be any number. In this context t is called a **parameter** . Therefore, the solution set of this system is

$$x = 3 + 5t$$
$$y = 10t$$
$$z = t$$

where t is arbitrary. The system has an infinite set of solutions which are given by these equations. For any value of t we select, x, y, and z will be given by the above equations. For example, if we choose $t = 4$ then the corresponding solution would be

$$x = 3 + 5(4) = 23$$
$$y = 10(4) = 40$$
$$z = 4$$

In Example 1.3.7 the solution involved one parameter. It may happen that the solution to a system involves more than one parameter, as shown in the following example.

Example 1.3.8: A Two Parameter Set of Solutions

Find the solution to the system

$$x + 2y - z + w = 3$$
$$x + y - z + w = 1$$
$$x + 3y - z + w = 5$$

Solution

The augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 1 & 1 & -1 & 1 & 1 \\ 1 & 3 & -1 & 1 & 5 \end{array}\right]$$

We wish to carry this matrix to row-echelon form. Here, we will outline the row operations used. However, make sure that you understand the steps in terms of Algorithm 1.3.1.

Take -1 times the first row and add to the second. Then take -1 times the first row and add to the third. This yields

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 0 & -1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 0 & 2 \end{array}\right]$$

Now add the second row to the third row and divide the second row by -1.

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{1.3.2}$$

This matrix is in row-echelon form and we can see that x and y correspond to pivot columns, while z and w do not. Therefore, we will assign parameters to the variables z and w. Assign the parameter s to z and the parameter t to w. Then the first row yields the equation $x+2y-s+t=3$, while the second row yields the equation $y=2$. Since $y=2$, the first equation becomes $x+4-s+t=3$ showing that the solution is given by

$$\begin{aligned} x &= -1+s-t \\ y &= 2 \\ z &= s \\ w &= t \end{aligned}$$

It is customary to write this solution in the form

$$\left[\begin{array}{c} x \\ y \\ z \\ w \end{array}\right] = \left[\begin{array}{c} -1+s-t \\ 2 \\ s \\ t \end{array}\right] \tag{1.3.3}$$

This example shows a system of equations with an infinite solution set which depends on two parameters. It can be less confusing in the case of an infinite solution set to first place the augmented matrix in reduced row-echelon form rather than just row-echelon form before seeking to write down the description of the solution.

In the above steps, this means we don't stop with the row-echelon form in equation (1.3.2). Instead we first place it in reduced row-echelon form as follows.

$$\left[\begin{array}{cccc|c} 1 & 0 & -1 & 1 & -1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Then the solution is $y=2$ from the second row and $x=-1+z-w$ from the first. Thus letting $z=s$ and $w=t$, the solution is given by (1.3.3).

You can see here that there are two paths to the correct answer, which both yield the same answer. Hence, either approach may be used. The process which we first used in the above solution is called **Gaussian Elimination** This process involves carrying the matrix to row-echelon form, converting back to equations, and using back substitution to find the solution. When you do row operations until you obtain reduced row-echelon form, the process is called **Gauss-Jordan Elimination**.

We have now found solutions for systems of equations with no solution and infinitely many solutions, with one parameter as well as two parameters. Recall the three types of solution sets which we discussed in the previous section; no solution, one solution, and infinitely many solutions. Each of these types of solutions could be identified from the graph of the system. It turns out that we can also identify the type of solution from the reduced row-echelon form of the augmented matrix.

- *No Solution:* In the case where the system of equations has no solution, the row-echelon form of the augmented matrix will have a row of the form

  $$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 1 \end{array}\right]$$

  This row indicates that the system is inconsistent and has no solution.
- *One Solution:* In the case where the system of equations has one solution, every column of the coefficient matrix is a pivot column. The following is an example of an augmented matrix in reduced row-echelon form for a system of equations with one solution.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

- *Infinitely Many Solutions:* In the case where the system of equations has infinitely many solutions, the solution contains parameters. There will be columns of the coefficient matrix which are not pivot columns. The following are examples of augmented matrices in reduced row-echelon form for systems of equations with infinitely many solutions.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

or

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -3 \end{array}\right]$$

This page titled 1.3: Gaussian Elimination is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.3: Gaussian Elimination** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

1.4: Uniqueness of the Reduced Row-Echelon Form

As we have seen in earlier sections, we know that every matrix can be brought into reduced row-echelon form by a sequence of elementary row operations. Here we will prove that the resulting matrix is unique; in other words, the resulting matrix in reduced row-echelon does not depend upon the particular sequence of elementary row operations or the order in which they were performed.

Let A be the augmented matrix of a homogeneous system of linear equations in the variables $x_1, x_2, \cdots, x_n$ which is also in reduced row-echelon form. The matrix A divides the set of variables in two different types. We say that x_i is a *basic variable* whenever A has a leading 1 in column number i, in other words, when column i is a pivot column. Otherwise we say that x_i is a *free variable*.

Recall Example 1.3.8.

Example 1.4.1: Basic and Free Variables

Find the basic and free variables in the system

$$\begin{aligned} x+2y-z+w &= 3 \\ x+y-z+w &= 1 \\ x+3y-z+w &= 5 \end{aligned}$$

Solution

Recall from the solution of Example 1.3.8 that the row-echelon form of the augmented matrix of this system is given by

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

You can see that columns 1 and 2 are pivot columns. These columns correspond to variables x and y, making these the basic variables. Columns 3 and 4 are not pivot columns, which means that z and w are free variables.

We can write the solution to this system as

$$\begin{aligned} x &= -1+s-t \\ y &= 2 \\ z &= s \\ w &= t \end{aligned}$$

Here the free variables are written as parameters, and the basic variables are given by linear functions of these parameters.

In general, all solutions can be written in terms of the free variables. In such a description, the free variables can take any values (they become parameters), while the basic variables become simple linear functions of these parameters. Indeed, a basic variable x_i is a linear function of *only* those free variables x_j with $j > i$. This leads to the following observation.

Proposition 1.4.1: Basic and Free Variables

If x_i is a basic variable of a homogeneous system of linear equations, then any solution of the system with $x_j = 0$ for all those free variables x_j with $j > i$ must also have $x_i = 0$.

Using this proposition, we prove a lemma which will be used in the proof of the main result of this section below.

Lemma 1.4.1: Solutions and the Reduced Row-Echelon Form of a Matrix

Let A and B be two distinct augmented matrices for two homogeneous systems of m equations in n variables, such that A and B are each in reduced row-echelon. Then, the two systems do not have exactly the same solutions.

Proof

With respect to the linear systems associated with the matrices A and B, there are two cases to consider:

- Case 1: the two systems have the same basic variables
- Case 2: the two systems do not have the same basic variables

In case 1, the two matrices will have exactly the same pivot positions. However, since A and B are not identical, there is some row of A which is different from the corresponding row of B and yet the rows each have a pivot in the same column position. Let i be the index of this column position. Since the matrices are in reduced row-echelon form, the two rows must differ at some entry in a column $j > i$. Let these entries be a in A and b in B, where $a \neq b$. Since A is in reduced row-echelon form, if x_j were a basic variable for its linear system, we would have $a = 0$. Similarly, if x_j were a basic variable for the linear system of the matrix B, we would have $b = 0$. Since a and b are unequal, they cannot both be equal to 0, and hence x_j cannot be a basic variable for both linear systems. However, since the systems have the same basic variables, x_j must then be a free variable for each system. We now look at the solutions of the systems in which x_j is set equal to 1 and all other free variables are set equal to 0. For this choice of parameters, the solution of the system for matrix A has $x_j = -a$, while the solution of the system for matrix B has $x_j = -b$, so that the two systems have different solutions.

In case 2, there is a variable x_i which is a basic variable for one matrix, let's say A, and a free variable for the other matrix B. The system for matrix B has a solution in which $x_i = 1$ and $x_j = 0$ for all other free variables x_j. However, by Proposition 1.4.1 this cannot be a solution of the system for the matrix A. This completes the proof of case 2.

Now, we say that the matrix B is **equivalent** to the matrix A provided that B can be obtained from A by performing a sequence of elementary row operations beginning with A. The importance of this concept lies in the following result.

Theorem 1.4.1: Equivalent Matrices

The two linear systems of equations corresponding to two equivalent augmented matrices have exactly the same solutions.

Proof

The proof of this theorem is left as an exercise.

Now, we can use Lemma 1.4.1 and Theorem 1.4.1 to prove the main result of this section.

Theorem 1.4.2: Uniqueness of the Reduced Row-Echelon Form

Every matrix A is equivalent to a unique matrix in reduced row-echelon form.

Proof

Let A be an $m \times n$ matrix and let B and C be matrices in reduced row-echelon form, each equivalent to A. It suffices to show that $B = C$.

Let A^+ be the matrix A augmented with a new rightmost column consisting entirely of zeros. Similarly, augment matrices B and C each with a rightmost column of zeros to obtain B^+ and C^+. Note that B^+ and C^+ are matrices in reduced row-echelon form which are obtained from A^+ by respectively applying the same sequence of elementary row operations which were used to obtain B and C from A.

Now, A^+, B^+, and C^+ can all be considered as augmented matrices of homogeneous linear systems in the variables $x_1, x_2, \cdots, x_n$. Because B^+ and C^+ are each equivalent to A^+, Theorem 1.4.1 ensures that all three homogeneous linear systems have exactly the same solutions. By Lemma 1.4.1 we conclude that $B^+ = C^+$. By construction, we must also have $B = C$.

According to this theorem we can say that each matrix A has a unique reduced row-echelon form.

This page titled 1.4: Uniqueness of the Reduced Row-Echelon Form is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.4: Uniqueness of the Reduced Row-Echelon Form** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

1.5: Rank and Homogeneous Systems

Rank and Homogeneous Systems

There is a special type of system which requires additional study. This type of system is called a homogeneous system of equations, which we defined above in Definition 1.2.3. Our focus in this section is to consider what types of solutions are possible for a homogeneous system of equations.

Consider the following definition.

Consider the homogeneous system of equations given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Then, $x_1 = 0, x_2 = 0, \cdots, x_n = 0$ is always a solution to this system. We call this the **trivial solution** .

If the system has a solution in which not all of the $x_1, \cdots, x_n$ are equal to zero, then we call this solution **nontrivial** . The trivial solution does not tell us much about the system, as it says that $0 = 0$! Therefore, when working with homogeneous systems of equations, we want to know when the system has a nontrivial solution.

Suppose we have a homogeneous system of m equations, using n variables, and suppose that $n > m$. In other words, there are more variables than equations. Then, it turns out that this system always has a nontrivial solution. Not only will the system have a nontrivial solution, but it also will have infinitely many solutions. It is also possible, but not required, to have a nontrivial solution if $n = m$ and $n < m$.

Consider the following example.

Example 1.5.1: Solutions to a Homogeneous System of Equations

Find the nontrivial solutions to the following homogeneous system of equations

$$\begin{aligned} 2x + y - z &= 0 \\ x + 2y - 2z &= 0 \end{aligned}$$

Solution

Notice that this system has $m = 2$ equations and $n = 3$ variables, so $n > m$. Therefore by our previous discussion, we expect this system to have infinitely many solutions.

The process we use to find the solutions for a homogeneous system of equations is the same process we used in the previous section. First, we construct the augmented matrix, given by

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 0 \\ 1 & 2 & -2 & 0 \end{array}\right]$$

Then, we carry this matrix to its reduced row-echelon form, given below.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right]$$

The corresponding system of equations is

$$\begin{aligned} x &= 0 \\ y - z &= 0 \end{aligned}$$

Since z is not restrained by any equation, we know that this variable will become our parameter. Let $z = t$ where t is any number. Therefore, our solution has the form

$$x = 0$$
$$y = z = t$$
$$z = t$$

Hence this system has infinitely many solutions, with one parameter t.

Suppose we were to write the solution to the previous example in another form. Specifically,

$$x = 0$$
$$y = 0 + t$$
$$z = 0 + t$$

can be written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Notice that we have constructed a column from the constants in the solution (all equal to 0), as well as a column corresponding to the coefficients on t in each equation. While we will discuss this form of solution more in further chapters, for now consider the column of coefficients of the parameter t. In this case, this is the column $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

There is a special name for this column, which is **basic solution**. The basic solutions of a system are columns constructed from the coefficients on parameters in the solution. We often denote basic solutions by X_1, X_2 etc., depending on how many solutions occur. Therefore, Example 1.5.1 has the basic solution $X_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

We explore this further in the following example.

Example 1.5.2: Basic Solutions of a Homogeneous System

Consider the following homogeneous system of equations.

$$x + 4y + 3z = 0$$
$$3x + 12y + 9z = 0$$

Find the basic solutions to this system.

Solution

The augmented matrix of this system and the resulting reduced row-echelon form are

$$\left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ 3 & 12 & 9 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

When written in equations, this system is given by

$$x + 4y + 3z = 0$$

Notice that only x corresponds to a pivot column. In this case, we will have two parameters, one for y and one for z. Let $y = s$ and $z = t$ for any numbers s and t. Then, our solution becomes

$$x = -4s - 3t$$
$$y = s$$
$$z = t$$

which can be written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

You can see here that we have two columns of coefficients corresponding to parameters, specifically one for s and one for t. Therefore, this system has two basic solutions! These are

$$X_1 = \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix}, X_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

Below is a video on homogeneous systems of equations



Below is a video on solving a system of equations and determining the particular solution.



Below is a video on solving Ax = 0.



Below is another video on solving Ax = 0.



We now present a new definition.

Definition 1.5.2: Linear Combination

Let $X_1, \cdots, X_n, V$ be column matrices. Then V is said to be a **linear combination** of the columns $X_1, \cdots, X_n$ if there exist scalars, $a_1, \cdots, a_n$ such that

$$V = a_1 X_1 + \cdots + a_n X_n$$

A remarkable result of this section is that a linear combination of the basic solutions is again a solution to the system. Even more remarkable is that every solution can be written as a linear combination of these solutions. Therefore, if we take a linear combination of the two solutions to Example 1.5.2, this would also be a solution. For example, we could take the following linear combination

$$3\begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -18 \\ 3 \\ 2 \end{bmatrix}$$

You should take a moment to verify that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -18 \\ 3 \\ 2 \end{bmatrix}$$

is in fact a solution to the system in Example 1.5.2.

Another way in which we can find out more information about the solutions of a homogeneous system is to consider the **rank** of the associated coefficient matrix. We now define what is meant by the rank of a matrix.

Definition 1.5.3: Rank of a Matrix

Let A be a matrix and consider any row-echelon form of A. Then, the number r of leading entries of A does not depend on the row-echelon form you choose, and is called the **rank** of A. We denote it by Rank(A).

Similarly, we could count the number of pivot positions (or pivot columns) to determine the rank of A.

Example 1.5.3: Finding the Rank of a Matrix

Consider the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 9 \\ 2 & 4 & 6 \end{bmatrix}$$

What is its rank?

Solution

First, we need to find the reduced row-echelon form of A. Through the usual algorithm, we find that this is

$$\begin{bmatrix} \boxed{1} & 0 & -1 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Here we have two leading entries, or two pivot positions, shown above in boxes.The rank of A is $r = 2$.

Notice that we would have achieved the same answer if we had found the row-echelon form of A instead of the reduced row-echelon form.

Suppose we have a homogeneous system of m equations in n variables, and suppose that $n > m$. From our above discussion, we know that this system will have infinitely many solutions. If we consider the rank of the coefficient matrix of this system, we can find out even more about the solution. Note that we are looking at just the coefficient matrix, not the entire augmented matrix.

Theorem 1.5.1: Rank and Solutions to a Homogeneous System

Let A be the $m \times n$ coefficient matrix corresponding to a homogeneous system of equations, and suppose A has rank r. Then, the solution to the corresponding system has $n - r$ parameters.

Consider our above Example 1.5.2 in the context of this theorem. The system in this example has $m = 2$ equations in $n = 3$ variables. First, because $n > m$, we know that the system has a nontrivial solution, and therefore infinitely many solutions. This tells us that the solution will contain at least one parameter. The rank of the coefficient matrix can tell us even more about the solution! The rank of the coefficient matrix of the system is 1, as it has one leading entry in row-echelon form. Theorem 1.5.1 tells us that the solution will have $n - r = 3 - 1 = 2$ parameters. You can check that this is true in the solution to Example 1.5.2.

Notice that if $n = m$ or $n < m$, it is possible to have either a unique solution (which will be the trivial solution) or infinitely many solutions.

We are not limited to homogeneous systems of equations here. The rank of a matrix can be used to learn about the solutions of any system of linear equations. In the previous section, we discussed that a system of equations can have no solution, a unique solution, or infinitely many solutions. Suppose the system is consistent, whether it is homogeneous or not. The following theorem tells us how we can use the rank to learn about the type of solution we have.

Below is a video on Rank, Basic Variables, and Free Variables of a coefficient matrix.



Theorem 1.5.2: Rank and Solutions to a Consistent System of Equations

Let A be the $m \times (n+1)$ augmented matrix corresponding to a consistent system of equations in n variables, and suppose A has rank r. Then

1. the system has a unique solution if $r = n$
2. the system has infinitely many solutions if $r < n$

We will not present a formal proof of this, but consider the following discussions.

1. *No Solution* The above theorem assumes that the system is consistent, that is, that it has a solution. It turns out that it is possible for the augmented matrix of a system with no solution to have any rank r as long as $r > 1$. Therefore, we must know that the system is consistent in order to use this theorem!
2. *Unique Solution* Suppose $r = n$. Then, there is a pivot position in every column of the coefficient matrix of A. Hence, there is a unique solution.
3. *Infinitely Many Solutions* Suppose $r < n$. Then there are infinitely many solutions. There are less pivot positions (and hence less leading entries) than columns, meaning that not every column is a pivot column. The columns which are *not* pivot columns correspond to parameters. In fact, in this case we have $n - r$ parameters.

Below is a video on the geometry of Ax = 0 and Ax = b.



Below is another video on the geometry of Ax = 0 and Ax = b.



Below is a video on the geometry of Ax = 0 and Ax = b, this one with two free variables.



This page titled 1.5: Rank and Homogeneous Systems is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.5: Rank and Homogeneous Systems** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

1.6: Balancing Chemical Reactions

The tools of linear algebra can also be used in the subject area of Chemistry, specifically for balancing chemical reactions.

Consider the chemical reaction

$$SnO_2 + H_2 \rightarrow Sn + H_2O$$

Here the elements involved are tin (Sn), oxygen (O), and hydrogen (H). A chemical reaction occurs and the result is a combination of tin (Sn) and water (H_2O). When considering chemical reactions, we want to investigate how much of each element we began with and how much of each element is involved in the result.

An important theory we will use here is the mass balance theory. It tells us that we cannot create or delete elements within a chemical reaction. For example, in the above expression, we must have the same number of oxygen, tin, and hydrogen on both sides of the reaction. Notice that this is not currently the case. For example, there are two oxygen atoms on the left and only one on the right. In order to fix this, we want to find numbers x, y, z, w such that

$$xSnO_2 + yH_2 \rightarrow zSn + wH_2O$$

where both sides of the reaction have the same number of atoms of the various elements.

This is a familiar problem. We can solve it by setting up a system of equations in the variables x, y, z, w. Thus you need

$$\begin{aligned} Sn: &\quad x = z \\ O: &\quad 2x = w \\ H: &\quad 2y = 2w \end{aligned}$$

We can rewrite these equations as

$$\begin{aligned} Sn: &\quad x - z = 0 \\ O: &\quad 2x - w = 0 \\ H: &\quad 2y - 2w = 0 \end{aligned}$$

The augmented matrix for this system of equations is given by

$$\left[\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 2 & 0 & 0 & -1 & 0 \\ 0 & 2 & 0 & -2 & 0 \end{array}\right]$$

The reduced row-echelon form of this matrix is

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & 0 \end{array}\right]$$

The solution is given by

$$\begin{aligned} x - \tfrac{1}{2}w &= 0 \\ y - w &= 0 \\ z - \tfrac{1}{2}w &= 0 \end{aligned}$$

which we can write as

$$\begin{aligned} x &= \tfrac{1}{2}t \\ y &= t \\ z &= \tfrac{1}{2}t \\ w &= t \end{aligned}$$

For example, let $w = 2$ and this would yield $x = 1, y = 2$, and $z = 1$. We can put these values back into the expression for the reaction which yields

$$SnO_2 + 2H_2 \rightarrow Sn + 2H_2O$$

Observe that each side of the expression contains the same number of atoms of each element. This means that it preserves the total number of atoms, as required, and so the chemical reaction is balanced.

Consider another example.

Example 1.6.1: Balancing a Chemical Reaction

Potassium is denoted by K, oxygen by O, phosphorus by P and hydrogen by H. Consider the reaction given by

$$KOH + H_3PO_4 \rightarrow K_3PO_4 + H_2O$$

Balance this chemical reaction.

Solution

We will use the same procedure as above to solve this problem. We need to find values for x, y, z, w such that

$$xKOH + yH_3PO_4 \rightarrow zK_3PO_4 + wH_2O$$

preserves the total number of atoms of each element.

Finding these values can be done by finding the solution to the following system of equations.

$$\begin{aligned} K: &\quad x = 3z \\ O: &\quad x + 4y = 4z + w \\ H: &\quad x + 3y = 2w \\ P: &\quad y = z \end{aligned}$$

The augmented matrix for this system is

$$\left[\begin{array}{cccc|c} 1 & 0 & -3 & 0 & 0 \\ 1 & 4 & -4 & -1 & 0 \\ 1 & 3 & 0 & -2 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right]$$

and the reduced row-echelon form is

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The solution is given by

$$\begin{aligned} x - w &= 0 \\ y - \tfrac{1}{3}w &= 0 \\ z - \tfrac{1}{3}w &= 0 \end{aligned}$$

which can be written as

$$\begin{aligned} x &= t \\ y &= \tfrac{1}{3}t \\ z &= \tfrac{1}{3}t \\ w &= t \end{aligned}$$

Choose a value for t, say 3. Then $w = 3$ and this yields $x = 3, y = 1, z = 1$. It follows that the balanced reaction is given by

$$3KOH + 1H_3PO_4 \rightarrow 1K_3PO_4 + 3H_2O$$

Note that this results in the same number of atoms on both sides.

Of course these numbers you are finding would typically be the number of moles of the molecules on each side. Thus three moles of KOH added to one mole of H_3PO_4 yields one mole of K_3PO_4 and three moles of H_2O.

This page titled 1.6: Balancing Chemical Reactions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.6: Balancing Chemical Reactions** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

1.7: Dimensionless Variables

This section shows how solving systems of equations can be used to determine appropriate dimensionless variables. It is only an introduction to this topic and considers a specific example of a simple airplane wing shown below. We assume for simplicity that it is a flat plane at an angle to the wind which is blowing against it with speed V as shown.

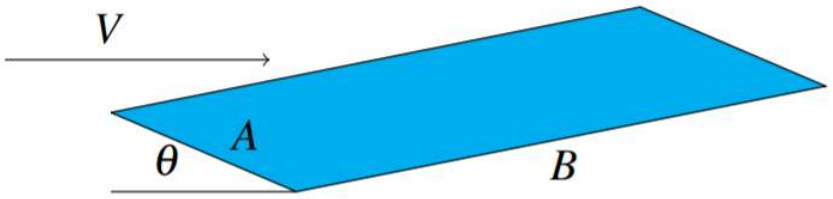


Figure 1.7.1

The angle θ is called the angle of incidence, B is the span of the wing and A is called the chord. Denote by l the lift. Then this should depend on various quantities like θ, V, B, A and so forth. Here is a table which indicates various quantities on which it is reasonable to expect l to depend.

Table 1.7.1

Variable	Symbol	Units
chord	A	m
span	B	m
angle incidence	θ	$m^0 kg^0 sec^0$
speed of wind	V	$msec^{-1}$
speed of sound	V_0	$msec^{-1}$
density of air	ρ	kgm^{-3}
viscosity	μ	$kgsec^{-1}m^{-1}$
lift	l	$kgsec^{-2}m$

Here m denotes meters, sec refers to seconds and kg refers to kilograms. All of these are likely familiar except for μ, which we will discuss in further detail now.

Viscosity is a measure of how much internal friction is experienced when the fluid moves. It is roughly a measure of how “sticky" the fluid is. Consider a piece of area parallel to the direction of motion of the fluid. To say that the viscosity is large is to say that the tangential force applied to this area must be large in order to achieve a given change in speed of the fluid in a direction normal to the tangential force. Thus

$$\mu\,(\text{area})\,(\text{velocity gradient}) = \text{ tangential force}$$

Hence

$$(\text{units on } \mu)\, m^2 \left(\frac{m}{\sec m}\right) = kg \sec^{-2} m$$

Thus the units on μ are

$$kg \sec^{-1} m^{-1}$$

as claimed above.

Returning to our original discussion, you may think that we would want

$$l = f\,(A, B, \theta, V, V_0, \rho, \mu)$$

This is very cumbersome because it depends on seven variables. Also, it is likely that without much care, a change in the units such as going from meters to feet would result in an incorrect value for l. The way to get around this problem is to look for l as a function of dimensionless variables multiplied by something which has units of force. It is helpful because first of all, you will likely have fewer independent variables and secondly, you could expect the formula to hold independent of the way of specifying length, mass and so forth. One looks for

$$l = f\left(g_1, \cdots, g_k\right) \rho V^2 AB$$

where the units on $\rho V^2 AB$ are

$$\frac{kg}{m^3}\left(\frac{m}{\sec}\right)^2 m^2 = \frac{kg \times m}{\sec^2}$$

which are the units of force. Each of these g_i is of the form

$$A^{x_1} B^{x_2} \theta^{x_3} V^{x_4} V_0^{x_5} \rho^{x_6} \mu^{x_7} \tag{1.7.1}$$

and each g_i is independent of the dimensions. That is, this expression must not depend on meters, kilograms, seconds, etc. Thus, placing in the units for each of these quantities, one needs

$$m^{x_1} m^{x_2} \left(m^{x_4} \sec^{-x_4}\right) \left(m^{x_5} \sec^{-x_5}\right) \left(kgm^{-3}\right)^{x_6} \left(kg \sec^{-1} m^{-1}\right)^{x_7} = m^0 kg^0 \sec^0$$

Notice that there are no units on θ because it is just the radian measure of an angle. Hence its dimensions consist of length divided by length, thus it is dimensionless. Then this leads to the following equations for the x_i.

$$\begin{aligned} m: &\quad x_1 + x_2 + x_4 + x_5 - 3x_6 - x_7 = 0 \\ \sec: &\quad -x_4 - x_5 - x_7 = 0 \\ kg: &\quad x_6 + x_7 = 0 \end{aligned}$$

The augmented matrix for this system is

$$\left[\begin{array}{ccccccc|c} 1 & 1 & 0 & 1 & 1 & -3 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

The reduced row-echelon form is given by

$$\left[\begin{array}{ccccccc|c} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

and so the solutions are of the form

$$\begin{aligned} x_1 &= -x_2 - x_7 \\ x_3 &= x_3 \\ x_4 &= -x_5 - x_7 \\ x_6 &= -x_7 \end{aligned}$$

Thus, in terms of vectors, the solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} -x_2 - x_7 \\ x_2 \\ x_3 \\ -x_5 - x_7 \\ x_5 \\ -x_7 \\ x_7 \end{bmatrix}$$

Thus the free variables are x_2, x_3, x_5, x_7. By assigning values to these, we can obtain dimensionless variables by placing the values obtained for the x_i in the formula (1.7.1). For example, let $x_2 = 1$ and all the rest of the free variables are 0. This yields

$$x_1 = -1, x_2 = 1, x_3 = 0, x_4 = 0, x_5 = 0, x_6 = 0, x_7 = 0$$

The dimensionless variable is then $A^{-1}B^1$. This is the ratio between the span and the chord. It is called the aspect ratio, denoted as AR. Next let $x_3 = 1$ and all others equal zero. This gives for a dimensionless quantity the angle θ. Next let $x_5 = 1$ and all others equal zero. This gives

$$x_1 = 0, x_2 = 0, x_3 = 0, x_4 = -1, x_5 = 1, x_6 = 0, x_7 = 0$$

Then the dimensionless variable is $V^{-1}V_0^1$. However, it is written as V/V_0. This is called the Mach number $\mathcal{M}$. Finally, let $x_7 = 1$ and all the other free variables equal 0. Then

$$x_1 = -1, x_2 = 0, x_3 = 0, x_4 = -1, x_5 = 0, x_6 = -1, x_7 = 1$$

then the dimensionless variable which results from this is $A^{-1}V^{-1}\rho^{-1}\mu$. It is customary to write it as $Re = (AV\rho)/\mu$. This one is called the Reynold's number. It is the one which involves viscosity. Thus we would look for

$$l = f(Re, AR, \theta, \mathcal{M})\, kg \times m/\sec^2$$

This is quite interesting because it is easy to vary Re by simply adjusting the velocity or A but it is hard to vary things like μ or ρ. Note that all the quantities are easy to adjust. Now this could be used, along with wind tunnel experiments to get a formula for the lift which would be reasonable. You could also consider more variables and more complicated situations in the same way.

This page titled 1.7: Dimensionless Variables is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.7: Dimensionless Variables** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

1.8: An Application to Resistor Networks

The tools of linear algebra can be used to study the application of resistor networks. An example of an electrical circuit is below.

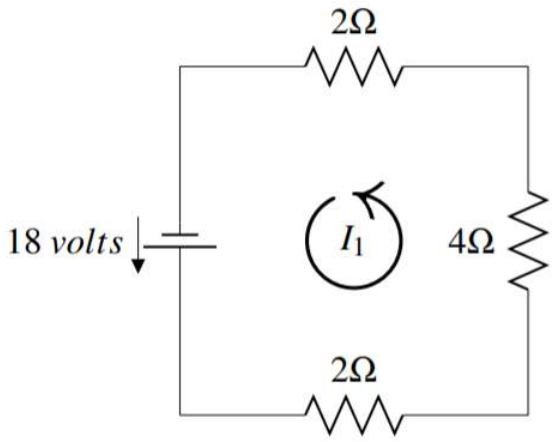


Figure 1.8.1

The jagged lines (—/\/\/\—) denote resistors and the numbers next to them give their resistance in ohms, written as Ω. The voltage source (—|| ||—) causes the current to flow in the direction from the shorter of the two lines toward the longer (as indicated by the arrow). The current for a circuit is labeled I_k.

In the above figure, the current I_1 has been labeled with an arrow in the counter clockwise direction. This is an entirely arbitrary decision and we could have chosen to label the current in the clockwise direction. With our choice of direction here, we define a positive current to flow in the counter clockwise direction and a negative current to flow in the clockwise direction.

The goal of this section is to use the values of resistors and voltage sources in a circuit to determine the current. An essential theorem for this application is Kirchhoff's law.

Theorem 1.8.1: Kirchhoff's Law

The sum of the resistance (R) times the amps (I) in the counter clockwise direction around a loop equals the sum of the voltage sources (V) in the same direction around the loop.

Kirchhoff's law allows us to set up a system of linear equations and solve for any unknown variables. When setting up this system, it is important to trace the circuit in the counter clockwise direction. If a resistor or voltage source is crossed against this direction, the related term must be given a negative sign.

We will explore this in the next example where we determine the value of the current in the initial diagram.

Example 1.8.1: Solving for Current

Applying Kirchhoff's Law to the diagram below, determine the value for I_1.

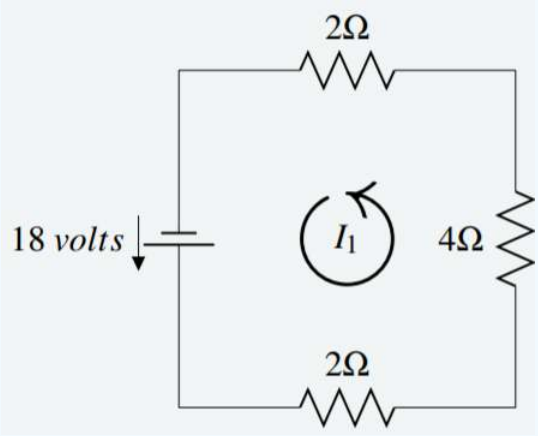


Figure 1.8.2

Solution

Begin in the bottom left corner, and trace the circuit in the counter clockwise direction. At the first resistor, multiplying resistance and current gives $2I_1$. Continuing in this way through all three resistors gives $2I_1+4I_1+2I_1$. This must equal the voltage source in the same direction. Notice that the direction of the voltage source matches the counter clockwise direction specified, so the voltage is positive.

Therefore the equation and solution are given by

$$\begin{aligned} 2I_1+4I_1+2I_1 &= 18 \\ 8I_1 &= 18 \\ I_1 &= \frac{9}{4}A \end{aligned}$$

Since the answer is positive, this confirms that the current flows counter clockwise.

Example 1.8.2 Solving for Current

Applying Kirchhoff's Law to the diagram below, determine the value for I_1.

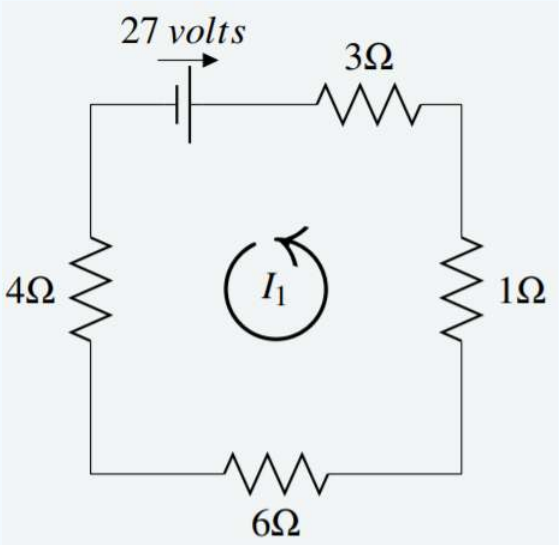


Figure 1.8.3

Solution

Begin in the top left corner this time, and trace the circuit in the counter clockwise direction. At the first resistor, multiplying resistance and current gives $4I_1$. Continuing in this way through the four resistors gives $4I_1+6I_1+1I_1+3I_1$. This must

equal the voltage source in the same direction. Notice that the direction of the voltage source is opposite to the counter clockwise direction, so the voltage is negative.

Therefore the equation and solution are given by

$$\begin{aligned} 4I_1+6I_1+1I_1+3I_1 &= -27 \\ 14I_1 &= -27 \\ I_1 &= -\frac{27}{14}A \end{aligned}$$

Since the answer is negative, this tells us that the current flows clockwise.

A more complicated example follows. Two of the circuits below may be familiar; they were examined in the examples above. However as they are now part of a larger system of circuits, the answers will differ.

Example 1.8.3: Unknown Currents

The diagram below consists of four circuits. The current (I_k) in the four circuits is denoted by I_1, I_2, I_3, I_4. Using Kirchhoff's Law, write an equation for each circuit and solve for each current.

Solution

The circuits are given in the following diagram.

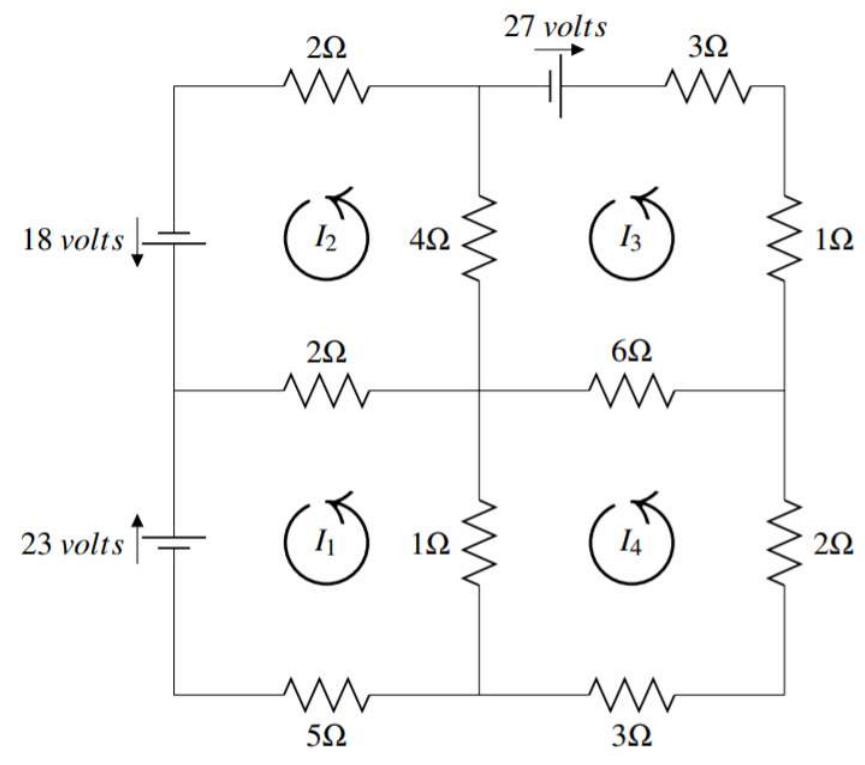


Figure 1.8.4

Starting with the top left circuit, multiply the resistance by the amps and sum the resulting products. Specifically, consider the resistor labeled 2Ω that is part of the circuits of I_1 and I_2. Notice that current I_2 runs through this in a positive (counter clockwise) direction, and I_1 runs through in the opposite (negative) direction. The product of resistance and amps is then $2(I_2-I_1)=2I_2-2I_1$. Continue in this way for each resistor, and set the sum of the products equal to the voltage source to write the equation:

$$2I_2-2I_1+4I_2-4I_3+2I_2=18$$

The above process is used on each of the other three circuits, and the resulting equations are:

Upper right circuit:

$$4I_3 - 4I_2 + 6I_3 - 6I_4 + I_3 + 3I_3 = -27$$

Lower right circuit:

$$3I_4 + 2I_4 + 6I_4 - 6I_3 + I_4 - I_1 = 0$$

Lower left circuit:

$$5I_1 + I_1 - I_4 + 2I_1 - 2I_2 = -23$$

Notice that the voltage for the upper right and lower left circuits are negative due to the clockwise direction they indicate.

The resulting system of four equations in four unknowns is

$$\begin{aligned} 2I_2 - 2I_1 + 4I_2 - 4I_3 + 2I_2 &= 18 \\ 4I_3 - 4I_2 + 6I_3 - 6I_4 + I_3 + I_3 &= -27 \\ 2I_4 + 3I_4 + 6I_4 - 6I_3 + I_4 - I_1 &= 0 \\ 5I_1 + I_1 - I_4 + 2I_1 - 2I_2 &= -23 \end{aligned}$$

Simplifying and rearranging with variables in order, we have:

$$\begin{aligned} -2I_1 + 8I_2 - 4I_3 &= 18 \\ -4I_2 + 14I_3 - 6I_4 &= -27 \\ -I_1 - 6I_3 + 12I_4 &= 0 \\ 8I_1 - 2I_2 - I_4 &= -23 \end{aligned}$$

The augmented matrix is

$$\left[\begin{array}{cccc|c} -2 & 8 & -4 & 0 & 18 \\ 0 & -4 & 14 & -6 & -27 \\ -1 & 0 & -6 & 12 & 0 \\ 8 & -2 & 0 & -1 & -23 \end{array}\right]$$

The solution to this matrix is

$$\begin{aligned} I_1 &= -3A \\ I_2 &= \frac{1}{4}A \\ I_3 &= -\frac{5}{2}A \\ I_4 &= -\frac{3}{2}A \end{aligned}$$

This tells us that currents I_1, I_3, and I_4 travel clockwise while I_2 travels counter clockwise.

This page titled 1.8: An Application to Resistor Networks is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.8: An Application to Resistor Networks** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

1.9: Exercises

Exercises

Exercise 1.9.1

Graphically, find the point (x_1, y_1) which lies on both lines, $x+3y=1$ and $4x-y=3$. That is, graph each line and see where they intersect.

Answer

$\begin{matrix} x+3y=1 \\ 4x-y=3 \end{matrix}$, Solution is: $\left[x=\frac{10}{13}, y=\frac{1}{13}\right]$.

Exercise 1.9.2

Graphically, find the point of intersection of the two lines $3x+y=3$ and $x+2y=1$. That is, graph each line and see where they intersect.

Answer

$\begin{matrix} 3x+y=3 \\ x+2y=1 \end{matrix}$, Solution is: $[x=1, y=0]$

Exercise 1.9.3

Do the three lines, $x+2y=1$, $2x-y=1$, and $4x+3y=3$ have a common point of intersection? If so, find the point and if not, tell why they don't have such a common point of intersection.

Answer

$\begin{matrix} x+2y=1 \\ 2x-y=1 \\ 4x+3y=3 \end{matrix}$, Solution is: $\left[x=\frac{3}{5}, y=\frac{1}{5}\right]$

Exercise 1.9.4

Do the three planes, $x+y-3z=2$, $2x+y+z=1$, and $3x+2y-2z=0$ have a common point of intersection? If so, find one and if not, tell why there is no such point.

Answer

No solution exists. You can see this by writing the augmented matrix and doing row operations. $\begin{bmatrix} 1 & 1 & -3 & 2 \\ 2 & 1 & 1 & 1 \\ 3 & 2 & -2 & 0 \end{bmatrix}$, row echelon form: $\begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & -7 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Thus one of the equation says $0=1$ in an equivalent system of equations.

Exercise 1.9.5

Four times the weight of Gaston is 150 pounds more than the weight of Ichabod. Four times the weight of Ichabod is 660 pounds less than seventeen times the weight of Gaston. Four times the weight of Gaston plus the weight of Siegfried equals 290 pounds. Brunhilde would balance all three of the others. Find the weights of the four people.

Answer

$$\begin{matrix} 4g-I=150 \\ 4I-17g=-660 \\ 4g+s=290 \\ g+I+s-b=0 \end{matrix}$$, Solution is: $\{g=60, I=90, b=200, s=50\}$

? Exercise 1.9.6

Consider the following augmented matrix in which $*$ denotes an arbitrary number and ■ denotes a nonzero number. Determine whether the given augmented matrix is consistent. If consistent, is the solution unique?

$$\left[\begin{array}{ccccc|c} ■ & * & * & * & * & * \\ 0 & ■ & * & * & 0 & * \\ 0 & 0 & ■ & * & * & * \\ 0 & 0 & 0 & 0 & ■ & * \end{array}\right]$$

Answer

The solution exists but is not unique.

? Exercise 1.9.7

Consider the following augmented matrix in which $*$ denotes an arbitrary number and ■ denotes a nonzero number. Determine whether the given augmented matrix is consistent. If consistent, is the solution unique?

$$\left[\begin{array}{ccc|c} ■ & * & * & * \\ 0 & ■ & * & * \\ 0 & 0 & ■ & * \end{array}\right]$$

Answer

A solution exists and is unique.

? Exercise 1.9.8

Consider the following augmented matrix in which $*$ denotes an arbitrary number and ■ denotes a nonzero number. Determine whether the given augmented matrix is consistent. If consistent, is the solution unique?

$$\left[\begin{array}{ccccc|c} ■ & * & * & * & * & * \\ 0 & ■ & 0 & * & 0 & * \\ 0 & 0 & 0 & ■ & * & * \\ 0 & 0 & 0 & 0 & ■ & * \end{array}\right]$$

? Exercise 1.9.9

Consider the following augmented matrix in which $*$ denotes an arbitrary number and ■ denotes a nonzero number. Determine whether the given augmented matrix is consistent. If consistent, is the solution unique?

$$\left[\begin{array}{ccccc|c} ■ & * & * & * & * & * \\ 0 & ■ & * & * & 0 & * \\ 0 & 0 & 0 & 0 & ■ & 0 \\ 0 & 0 & 0 & 0 & * & ■ \end{array}\right]$$

Answer

There might be a solution. If so, there are infinitely many.

Exercise 1.9.10

Suppose a system of equations has fewer equations than variables. Will such a system necessarily be consistent? If so, explain why and if not, give an example which is not consistent.

Answer

No. Consider $x+y+z=2$ and $x+y+z=1$.

Exercise 1.9.11

If a system of equations has more equations than variables, can it have a solution? If so, give an example and if not, tell why not.

Answer

These can have a solution. For example, $x+y=1, 2x+2y=2, 3x+3y=3$ even has an infinite set of solutions.

Exercise 1.9.12

Find h such that

$$\left[\begin{array}{cc|c} 2 & h & 4 \\ 3 & 6 & 7 \end{array}\right]$$

is the augmented matrix of an *inconsistent* system.

Answer

$h=4$

Exercise 1.9.13

Find h such that

$$\left[\begin{array}{cc|c} 1 & h & 3 \\ 2 & 4 & 6 \end{array}\right]$$

is the augmented matrix of a *consistent* system.

Answer

Any h will work.

Exercise 1.9.14

Find h such that

$$\left[\begin{array}{cc|c} 1 & 1 & 4 \\ 3 & h & 12 \end{array}\right]$$

is the augmented matrix of a *consistent* system.

Answer

Any h will work.

Exercise 1.9.15

Choose h and k such that the augmented matrix shown has each of the following:

a. one solution
b. no solution
c. infinitely many solutions

$$\left[\begin{array}{cc|c} 1 & h & 2 \\ 2 & 4 & k \end{array}\right]$$

Answer

If $h \neq 2$ there will be a unique solution for any k. If $h = 2$ and $k \neq 4$, there are no solutions. If $h = 2$ and $k = 4$, then there are infinitely many solutions.

Exercise 1.9.16

Choose h and k such that the augmented matrix shown has each of the following:

a. one solution
b. no solution
c. infinitely many solutions

$$\left[\begin{array}{cc|c} 1 & 2 & 2 \\ 2 & h & k \end{array}\right]$$

Answer

If $h \neq 4$, then there is exactly one solution. If $h = 4$ and $k \neq 4$, then there are no solutions. If $h = 4$ and $k = 4$, then there are infinitely many solutions.

Exercise 1.9.17

Determine if the system is consistent. If so, is the solution unique?

$$\begin{aligned} x + 2y + z - w &= 2 \\ x - y + z + w &= 1 \\ 2x + y - z &= 1 \\ 4x + 2y + z &= 5 \end{aligned}$$

Answer

There is no solution. The system is inconsistent. You can see this from the augmented matrix. $\begin{bmatrix} 1 & 2 & 1 & -1 & 2 \\ 1 & -1 & 1 & 1 & 1 \\ 2 & 1 & -1 & 0 & 1 \\ 4 & 2 & 1 & 0 & 5 \end{bmatrix}$,

reduced row-echelon form: $\begin{bmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{2}{3} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

Exercise 1.9.18

Determine if the system is consistent. If so, is the solution unique?

$$\begin{aligned} x+2y+z-w &= 2 \\ x-y+z+w &= 0 \\ 2x+y-z &= 1 \\ 4x+2y+z &= 3 \end{aligned}$$

Answer

Solution is: $\left[w=\frac{3}{2}y-1, x=\frac{2}{3}-\frac{1}{2}y, z=\frac{1}{3}\right]$

Exercise 1.9.19

Determine which matrices are in reduced row-echelon form.

a. $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 7 \end{bmatrix}$

b. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 5 \\ 0 & 0 & 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{bmatrix}$

Answer

a. This one is not.
b. This one is.
c. This one is.

Exercise 1.9.20

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 2 & -1 & 3 & -1 \\ 1 & 0 & 2 & 1 \\ 1 & -1 & 1 & -2 \end{bmatrix}$$

Exercise 1.9.21

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 0 & 0 & -1 & -1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & -1 \end{bmatrix}$$

Exercise 1.9.22

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 3 & -6 & -7 & -8 \\ 1 & -2 & -2 & -2 \\ 1 & -2 & -3 & -4 \end{bmatrix}$$

Exercise 1.9.23

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 2 & 4 & 5 & 15 \\ 1 & 2 & 3 & 9 \\ 1 & 2 & 2 & 6 \end{bmatrix}$$

Exercise 1.9.24

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 4 & -1 & 7 & 10 \\ 1 & 0 & 3 & 3 \\ 1 & -1 & -2 & 1 \end{bmatrix}$$

Exercise 1.9.25

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 3 & 5 & -4 & 2 \\ 1 & 2 & -1 & 1 \\ 1 & 1 & -2 & 0 \end{bmatrix}$$

Exercise 1.9.26

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} -2 & 3 & -8 & 7 \\ 1 & -2 & 5 & -5 \\ 1 & -3 & 7 & -8 \end{bmatrix}$$

Exercise 1.9.27

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 1 & 3 & 4 & 2 \\ 1 & 0 & 2 & 1 \end{array}\right]$$

Exercise 1.9.28

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 2 & 0 & 1 & 1 \\ 3 & 2 & 1 & 3 \end{array}\right]$$

Answer

The reduced row-echelon form is $\left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & -\frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & 0 \end{array}\right]$. Therefore, the solution is of the form $z = t, y = \frac{3}{4} + t\left(\frac{1}{4}\right), x = \frac{1}{2} - \frac{1}{2}t$ where $t \in \mathbb{R}$.

Exercise 1.9.29

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 1 & 0 & 4 & 2 \end{array}\right]$$

Answer

The reduced row-echelon form is $\left[\begin{array}{ccc|c} 1 & 0 & 4 & 2 \\ 0 & 1 & -4 & -1 \end{array}\right]$ and so the solution is $z = t, y = 4t, x = 2 - 4t$.

Exercise 1.9.30

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccccc|c} 1 & 0 & 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 1 & 2 & 0 & 0 & 1 & 3 \\ 1 & 0 & 1 & 0 & 2 & 2 \end{array}\right]$$

Answer

The reduced row-echelon form is $\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 9 & 3 \\ 0 & 1 & 0 & 0 & -4 & 0 \\ 0 & 0 & 1 & 0 & -7 & -1 \\ 0 & 0 & 0 & 1 & 6 & 1 \end{array}\right]$ and so $x_5 = t, x_4 = 1 - 6t, x_3 = -1 + 7t, x_2 = 4t, x_1 = 3 - 9t$.

Exercise 1.9.31

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccccc|c} 1 & 0 & 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 0 & 2 & 0 & 0 & 1 & 3 \\ 1 & -1 & 2 & 2 & 2 & 0 \end{array}\right]$$

Answer

The reduced row-echelon form is $\left[\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & -\frac{1}{2} & \frac{5}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 0 & 1 & \frac{3}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$. Therefore, let $x_5 = t, x_3 = s$. Then the other variables are given by $x_4 = -\frac{1}{2} - \frac{3}{2}t, x_2 = \frac{3}{2} - t\frac{1}{2}, , x_1 = \frac{5}{2} + \frac{1}{2}t - 2s$.

Exercise 1.9.32

Find the solution to the system of equations, $7x + 14y + 15z = 22$, $2x + 4y + 3z = 5$, and $3x + 6y + 10z = 13$.

Answer

Solution is: $[x = 1 - 2t, z = 1, y = t]$

Exercise 1.9.33

Find the solution to the system of equations, $3x - y + 4z = 6,\ y + 8z = 0,$ and $-2x + y = -4.$

Answer

Solution is: $[x = 2 - 4t, y = -8t, z = t]$

Exercise 1.9.34

Find the solution to the system of equations, $9x - 2y + 4z = -17,\ 13x - 3y + 6z = -25,$ and $-2x - z = 3.$

Answer

Solution is: $[x = -1, y = 2, z = -1]$

Exercise 1.9.35

Find the solution to the system of equations, $65x + 84y + 16z = 546,\ 81x + 105y + 20z = 682,$ and $84x + 110y + 21z = 713.$

Answer

Solution is: $[x = 2, y = 4, z = 5]$

Exercise 1.9.36

Find the solution to the system of equations, $8x + 2y + 3z = -3, 8x + 3y + 3z = -1,$ and $4x + y + 3z = -9.$

Answer

Solution is: $[x = 1, y = 2, z = -5]$

Exercise 1.9.37

Find the solution to the system of equations, $-8x + 2y + 5z = 18, -8x + 3y + 5z = 13,$ and $-4x + y + 5z = 19.$

Answer

Solution is: $[x = -1, y = -5, z = 4]$

Exercise 1.9.38

Find the solution to the system of equations, $3x - y - 2z = 3,\ y - 4z = 0,$ and $-2x + y = -2.$

Answer

Solution is: $[x = 2t + 1, y = 4t, z = t]$

Exercise 1.9.39

Find the solution to the system of equations, $-9x + 15y = 66, -11x + 18y = 79, -x + y = 4$, and $z = 3$.

Answer

Solution is: $[x=1, y=5, z=3]$

Exercise 1.9.40

Find the solution to the system of equations, $-19x+8y=-108$, $-71x+30y=-404$, $-2x+y=-12$, $4x+z=14$.

Answer

Solution is: $[x=4, y=-4, z=-2]$

Exercise 1.9.41

Suppose a system of equations has fewer equations than variables and you have found a solution to this system of equations. Is it possible that your solution is the only one? Explain.

Answer

No. Consider $x+y+z=2$ and $x+y+z=1$.

Exercise 1.9.42

Suppose a system of linear equations has a 2×4 augmented matrix and the last column is a pivot column. Could the system of linear equations be consistent? Explain.

Answer

No. This would lead to $0=1$.

Exercise 1.9.43

Suppose the coefficient matrix of a system of n equations with n variables has the property that every column is a pivot column. Does it follow that the system of equations must have a solution? If so, must the solution be unique? Explain.

Answer

Yes. It has a unique solution.

Exercise 1.9.44

Suppose there is a unique solution to a system of linear equations. What must be true of the pivot columns in the augmented matrix?

Answer

The last column must not be a pivot column. The remaining columns must each be pivot columns.

Exercise 1.9.45

The steady state temperature, u, of a plate solves Laplace's equation, $\Delta u=0$. One way to approximate the solution is to divide the plate into a square mesh and require the temperature at each node to equal the average of the temperature at the four adjacent nodes. In the following picture, the numbers represent the observed temperature at the indicated nodes. Find the temperature at the interior nodes, indicated by $x, y, z,$ and w. One of the equations is $z=0.05in\frac{1}{4}(10+0+w+x)$.

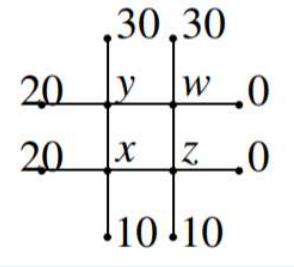


Figure 1.9.1

Answer

You need $\begin{matrix} \frac{1}{4}(20+30+w+x)-y=0 \\ \frac{1}{4}(y+30+0+z)-w=0 \\ \frac{1}{4}(20+y+z+10)-x=0 \\ \frac{1}{4}(x+w+0+10)-z=0 \end{matrix}$, Solution is: $[w=15, x=15, y=20, z=10]$.

? Exercise 1.9.46

Find the rank of the following matrix.

$$\begin{bmatrix} 4 & -16 & -1 & -5 \\ 1 & -4 & 0 & -1 \\ 1 & -4 & -1 & -2 \end{bmatrix}$$

? Exercise 1.9.47

Find the rank of the following matrix.

$$\begin{bmatrix} 3 & 6 & 5 & 12 \\ 1 & 2 & 2 & 5 \\ 1 & 2 & 1 & 2 \end{bmatrix}$$

? Exercise 1.9.48

Find the rank of the following matrix.

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 3 \\ 1 & 4 & 1 & 0 & -8 \\ 1 & 4 & 0 & 1 & 2 \\ -1 & -4 & 0 & -1 & -2 \end{bmatrix}$$

? Exercise 1.9.49

Find the rank of the following matrix.

$$\begin{bmatrix} 4 & -4 & 3 & -9 \\ 1 & -1 & 1 & -2 \\ 1 & -1 & 0 & -3 \end{bmatrix}$$

Exercise 1.9.50

Find the rank of the following matrix.

$$\begin{bmatrix} 2 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 7 \\ 1 & 0 & 0 & 1 & 7 \end{bmatrix}$$

Exercise 1.9.51

Find the rank of the following matrix.

$$\begin{bmatrix} 4 & 15 & 29 \\ 1 & 4 & 8 \\ 1 & 3 & 5 \\ 3 & 9 & 15 \end{bmatrix}$$

Exercise 1.9.52

Find the rank of the following matrix.

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 1 \\ 1 & 2 & 3 & -2 & -18 \\ 1 & 2 & 2 & -1 & -11 \\ -1 & -2 & -2 & 1 & 11 \end{bmatrix}$$

Exercise 1.9.53

Find the rank of the following matrix.

$$\begin{bmatrix} 1 & -2 & 0 & 3 & 11 \\ 1 & -2 & 0 & 4 & 15 \\ 1 & -2 & 0 & 3 & 11 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Exercise 1.9.54

Find the rank of the following matrix.

$$\begin{bmatrix} -2 & -3 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ -3 & 0 & -3 \end{bmatrix}$$

Exercise 1.9.55

Find the rank of the following matrix.

$$\begin{bmatrix} 4 & 4 & 20 & -1 & 17 \\ 1 & 1 & 5 & 0 & 5 \\ 1 & 1 & 5 & -1 & 2 \\ 3 & 3 & 15 & -3 & 6 \end{bmatrix}$$

Exercise 1.9.56

Find the rank of the following matrix.

$$\begin{bmatrix} -1 & 3 & 4 & -3 & 8 \\ 1 & -3 & -4 & 2 & -5 \\ 1 & -3 & -4 & 1 & -2 \\ -2 & 6 & 8 & -2 & 4 \end{bmatrix}$$

Exercise 1.9.57

Suppose A is an $m \times n$ matrix. Explain why the rank of A is always no larger than $\min(m, n)$.

Answer

It is because you cannot have more than $\min(m, n)$ nonzero rows in the reduced row-echelon form. Recall that the number of pivot columns is the same as the number of nonzero rows from the description of this reduced row-echelon form.

Exercise 1.9.58

State whether each of the following sets of data are possible for the matrix equation $AX = B$. If possible, describe the solution set. That is, tell whether there exists a unique solution, no solution or infinitely many solutions. Here, $[A|B]$ denotes the augmented matrix.

a. A is a 5×6 matrix, $rank(A) = 4$ and $rank[A|B] = 4$.
b. A is a 3×4 matrix, $rank(A) = 3$ and $rank[A|B] = 2$.
c. A is a 4×2 matrix, $rank(A) = 4$ and $rank[A|B] = 4$.
d. A is a 5×5 matrix, $rank(A) = 4$ and $rank[A|B] = 5$.
e. A is a 4×2 matrix, $rank(A) = 2$ and $rank[A|B] = 2$.

Answer

a. This says B is in the span of four of the columns. Thus the columns are not independent. Infinite solution set.
b. This surely can't happen. If you add in another column, the rank does not get smaller.
c. This says B is in the span of the columns and the columns must be independent. You can't have the rank equal 4 if you only have two columns.
d. This says B is not in the span of the columns. In this case, there is no solution to the system of equations represented by the augmented matrix.
e. In this case, there is a unique solution since the columns of A are independent.

Exercise 1.9.59

Consider the system $-5x + 2y - z = 0$ and $-5x - 2y - z = 0$. Both equations equal zero and so $-5x + 2y - z = -5x - 2y - z$ which is equivalent to $y = 0$. Does it follow that x and z can equal anything? Notice that when $x = 1$, $z = -4$, and $y = 0$ are plugged in to the equations, the equations do not equal 0. Why?

Answer

These are not legitimate row operations. They do not preserve the solution set of the system.

Exercise 1.9.60

Balance the following chemical reactions.

a. $KNO_3 + H_2CO_3 \rightarrow K_2CO_3 + HNO_3$
b. $AgI + Na_2S \rightarrow Ag_2S + NaI$

c. $Ba_3N_2 + H_2O \rightarrow Ba(OH)_2 + NH_3$
d. $CaCl_2 + Na_3PO_4 \rightarrow Ca_3(PO_4)_2 + NaCl$

? Exercise 1.9.61

In the section on dimensionless variables it was observed that $\rho V^2 AB$ has the units of force. Describe a systematic way to obtain such combinations of the variables which will yield something which has the units of force.

? Exercise 1.9.62

Consider the following diagram of four circuits.

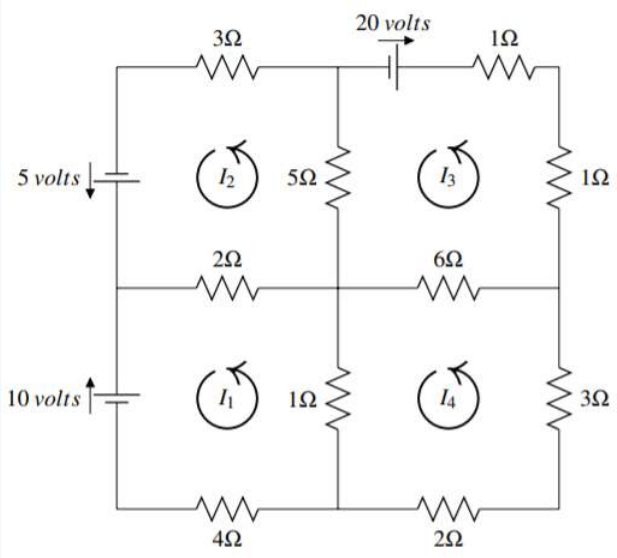


Figure 1.9.2

The current in amps in the four circuits is denoted by I_1, I_2, I_3, I_4 and it is understood that the motion is in the counter clockwise direction. If I_k ends up being negative, then it just means the current flows in the clockwise direction.

In the above diagram, the top left circuit should give the equation

$$2I_2 - 2I_1 + 5I_2 - 5I_3 + 3I_2 = 5$$

For the circuit on the lower left, you should have

$$4I_1 + I_1 - I_4 + 2I_1 - 2I_2 = -10$$

Write equations for each of the other two circuits and then give a solution to the resulting system of equations.

Answer

The other two equations are

$$\begin{aligned} 6I_3 - 6I_4 + I_3 + I_3 + 5I_3 - 5I_2 &= -20 \\ 2I_4 + 3I_4 + 6I_4 - 6I_3 + I_4 - I_1 &= 0 \end{aligned}$$

Then the system is

$$\begin{aligned} 2I_2 - 2I_1 + 5I_2 - 5I_3 + 3I_2 &= 5 \\ 4I_1 + I_1 - I_4 + 2I_1 - 2I_2 &= -10 \\ 6I_3 - 6I_4 + I_3 + I_3 + 5I_3 - 5I_2 &= -20 \\ 2I_4 + 3I_4 + 6I_4 - 6I_3 + I_4 - I_1 &= 0 \end{aligned}$$

The solution is:

$$I_1 = -\frac{750}{373}$$
$$I_2 = -\frac{1421}{1119}$$
$$I_3 = -\frac{3061}{1119}$$
$$I_4 = -\frac{1718}{1119}$$

? Exercise 1.9.63

Consider the following diagram of three circuits.

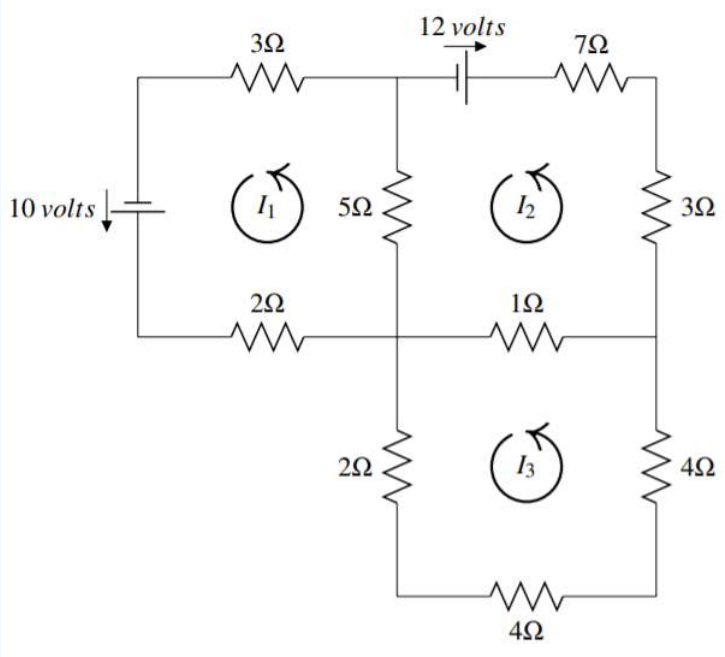


Figure 1.9.3

The current in amps in the four circuits is denoted by I_1, I_2, I_3 and it is understood that the motion is in the counter clockwise direction. If I_k ends up being negative, then it just means the current flows in the clockwise direction.

Find I_1, I_2, I_3.

Answer

You have

$$\begin{aligned} 2I_1 + 5I_1 + 3I_1 - 5I_2 &= 10 \\ I_2 - I_3 + 3I_2 + 7I_2 + 5I_2 - 5I_1 &= -12 \\ 2I_3 + 4I_3 + 4I_3 + I_3 - I_2 &= 0 \end{aligned}$$

Simplifying this yields

$$\begin{aligned} 10I_1 - 5I_2 &= 10 \\ -5I_1 + 16I_2 - I_3 &= -12 \\ -I_2 + 11I_3 &= 0 \end{aligned}$$

The solution is given by

$$I_1 = \frac{218}{295}, I_2 = -\frac{154}{295}, I_3 = -\frac{14}{295}$$

This page titled 1.9: Exercises is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.9: Exercises** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

CHAPTER OVERVIEW

2: Systems of Equations

2.1: Systems of Equations, Geometry
2.2: Systems of Equations, Algebraic Procedures
2.3: Gaussian Elimination
2.4: Uniqueness of the Reduced Row-Echelon Form
2.5: Rank and Homogeneous Systems
2.6: Balancing Chemical Reactions
2.7: Dimensionless Variables
2.8: An Application to Resistor Networks
2.9: Exercises

This page titled 2: Systems of Equations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

2.1: Systems of Equations, Geometry

Systems of Equations, Geometry

Outcomes

A. Relate the types of solution sets of a system of two (three) variables to the intersections of lines in a plane (the intersection of planes in three space)

As you may remember, linear equations like $2x+3y=6$ can be graphed as straight lines in the coordinate plane. We say that this equation is in two variables, in this case x and y. Suppose you have two such equations, each of which can be graphed as a straight line, and consider the resulting graph of two lines. What would it mean if there exists a point of intersection between the two lines? This point, which lies on *both* graphs, gives x and y values for which both equations are true. In other words, this point gives the ordered pair (x,y) that satisfy both equations. If the point (x,y) is a point of intersection, we say that (x,y) is a **solution** to the two equations. In linear algebra, we often are concerned with finding the solution(s) to a system of equations, if such solutions exist. First, we consider graphical representations of solutions and later we will consider the algebraic methods for finding solutions.

When looking for the intersection of two lines in a graph, several situations may arise. The following picture demonstrates the possible situations when considering two equations (two lines in the graph) involving two variables.

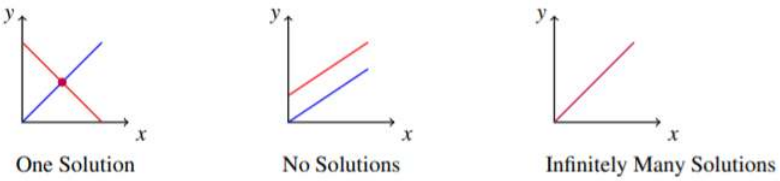


Figure 2.1.1

In the first diagram, there is a unique point of intersection, which means that there is only one (unique) solution to the two equations. In the second, there are no points of intersection and no solution. When no solution exists, this means that the two lines are parallel and they never intersect. The third situation which can occur, as demonstrated in diagram three, is that the two lines are really the same line. For example, $x+y=1$ and $2x+2y=2$ are equations which when graphed yield the same line. In this case there are infinitely many points which are solutions of these two equations, as every ordered pair which is on the graph of the line satisfies both equations. When considering linear systems of equations, there are always three types of solutions possible; exactly one (unique) solution, infinitely many solutions, or no solution.

Example 2.1.1: A Graphical Solution

Use a graph to find the solution to the following system of equations

$$\begin{aligned} x+y&=3\\ y-x&=5 \end{aligned}$$

Solution

Through graphing the above equations and identifying the point of intersection, we can find the solution(s). Remember that we must have either one solution, infinitely many, or no solutions at all. The following graph shows the two equations, as well as the intersection. Remember, the point of intersection represents the solution of the two equations, or the (x,y) which satisfy both equations. In this case, there is one point of intersection at $(-1,4)$ which means we have one unique solution, $x=-1, y=4$.

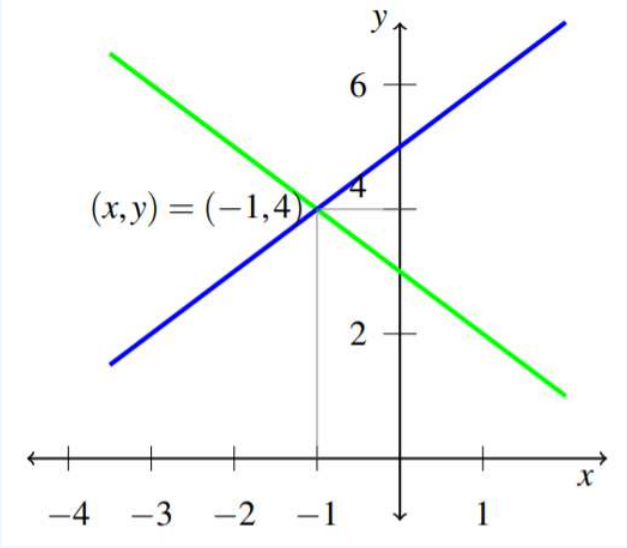


Figure 2.1.2

In the above example, we investigated the intersection point of two equations in two variables, x and y. Now we will consider the graphical solutions of three equations in two variables.

Consider a system of three equations in two variables. Again, these equations can be graphed as straight lines in the plane, so that the resulting graph contains three straight lines. Recall the three possible types of solutions; no solution, one solution, and infinitely many solutions. There are now more complex ways of achieving these situations, due to the presence of the third line. For example, you can imagine the case of three intersecting lines having no common point of intersection. Perhaps you can also imagine three intersecting lines which do intersect at a single point. These two situations are illustrated below.

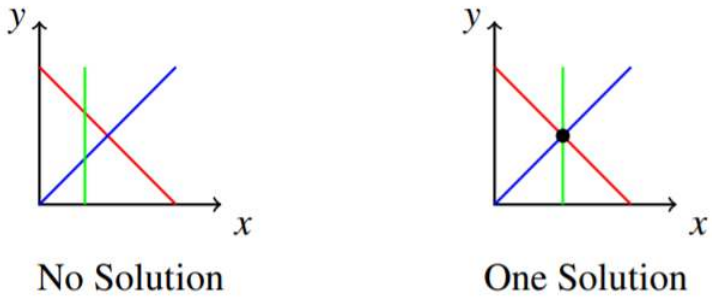


Figure 2.1.3

Consider the first picture above. While all three lines intersect with one another, there is no common point of intersection where all three lines meet at one point. Hence, there is no solution to the three equations. Remember, a solution is a point (x, y) which satisfies **all** three equations. In the case of the second picture, the lines intersect at a common point. This means that there is one solution to the three equations whose graphs are the given lines. You should take a moment now to draw the graph of a system which results in three parallel lines. Next, try the graph of three identical lines. Which type of solution is represented in each of these graphs?

We have now considered the graphical solutions of systems of two equations in two variables, as well as three equations in two variables. However, there is no reason to limit our investigation to equations in two variables. We will now consider equations in three variables.

Below is a video on solving systems of equations graphically



You may recall that equations in three variables, such as $2x+4y-5z=8$, form a plane. Above, we were looking for intersections of lines in order to identify any possible solutions. When graphically solving systems of equations in three variables, we look for intersections of planes. These points of intersection give the (x,y,z) that satisfy all the equations in the system. What types of solutions are possible when working with three variables? Consider the following picture involving two planes, which are given by two equations in three variables.

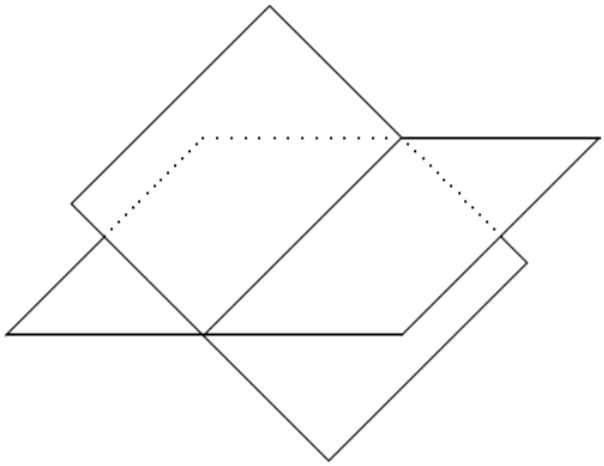


Figure 2.1.4

Notice how these two planes intersect in a line. This means that the points (x,y,z) on this line satisfy both equations in the system. Since the line contains infinitely many points, this system has infinitely many solutions.

It could also happen that the two planes fail to intersect. However, is it possible to have two planes intersect at a single point? Take a moment to attempt drawing this situation, and convince yourself that it is not possible! This means that when we have only two equations in three variables, there is no way to have a unique solution! Hence, the types of solutions possible for two equations in three variables are no solution or infinitely many solutions.

Now imagine adding a third plane. In other words, consider three equations in three variables. What types of solutions are now possible? Consider the following diagram.

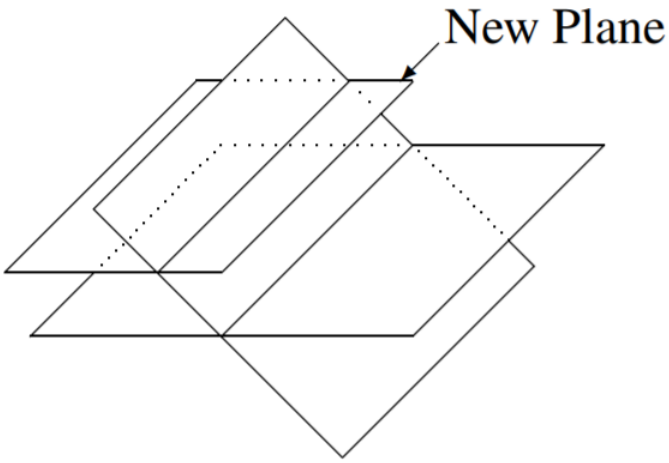


Figure 2.1.5

In this diagram, there is no point which lies in all three planes. There is no intersection between **all** planes so there is no solution. The picture illustrates the situation in which the line of intersection of the new plane with one of the original planes forms a line parallel to the line of intersection of the first two planes. However, in three dimensions, it is possible for two lines to fail to intersect even though they are not parallel. Such lines are called **skew lines**.

Recall that when working with two equations in three variables, it was not possible to have a unique solution. Is it possible when considering three equations in three variables? In fact, it is possible, and we demonstrate this situation in the following picture.

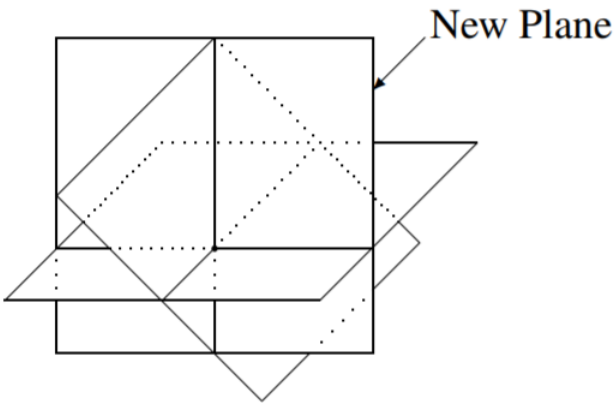


Figure 2.1.6

In this case, the three planes have a single point of intersection. Can you think of other types of solutions possible? Another is that the three planes could intersect in a line, resulting in infinitely many solutions, as in the following diagram.

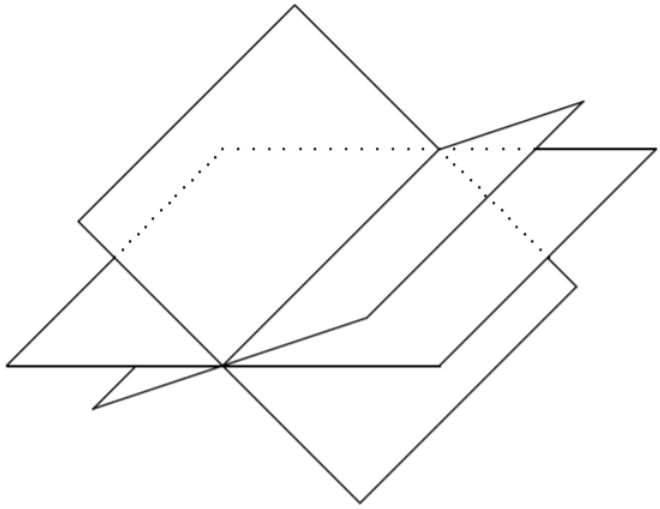


Figure 2.1.7

We have now seen how three equations in three variables can have no solution, a unique solution, or intersect in a line resulting in infinitely many solutions. It is also possible that the three equations graph the same plane, which also leads to infinitely many solutions.

You can see that when working with equations in three variables, there are many more ways to achieve the different types of solutions than when working with two variables. It may prove enlightening to spend time imagining (and drawing) many possible scenarios, and you should take some time to try a few.

You should also take some time to imagine (and draw) graphs of systems in more than three variables. Equations like $x+y-2z+4w=8$ with more than three variables are often called **hyper-planes**. You may soon realize that it is tricky to draw the graphs of hyper-planes! Through the tools of linear algebra, we can algebraically examine these types of systems which are difficult to graph. In the following section, we will consider these algebraic tools.

More Videos:

Solving Systems of Equations Graphically
Solving Systems of Equations by Graphing
Determining the Intersection of Two Graph on the TI83/84
Ex: Identify the Solution to a System of Equation Given a Graph, Then Verify
olve a System of Linear Equations by Graphing (One Solution)
Solve a System of Linear Equations by Graphing (No Solution)
x 1: Solve a System of Equations by Graphing
Ex 2: Solve a System of Equations by Graphing
Ex: Solve a System of Equations by Graphing (Infinite Solutions)
Ex: Solve a System of Equations by Graphing (No Solution)

? Exercise 2.1.1

Graphically, find the point (x_1, y_1) which lies on both lines, $x+3y=1$ and $4x-y=3$. That is, graph each line and see where they intersect.

? Exercise 2.1.2

Graphically, find the point of intersection of the two lines, $3x+y=3$ and $x+2y=1$. That is, graph each line and see where they intersect

Exercise 2.1.3

You have a system of k equations in two variables, $k \geq 2$. Explain the geometric significance of

a. No solution.
b. A unique solution.
c. An infinite number of solutions.

This page titled 2.1: Systems of Equations, Geometry is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.1: Systems of Equations, Geometry** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

2.2: Systems of Equations, Algebraic Procedures

Algebraic Procedures

Outcomes

A. Use elementary operations to find the solution to a linear system of equations.
B. Find the row-echelon form and reduced row-echelon form of a matrix.
C. Determine whether a system of linear equations has no solution, a unique solution or an infinite number of solutions from its .
D. Solve a system of equations using Gaussian Elimination and Gauss-Jordan Elimination.
E. Model a physical system with linear equations and then solve.

We have taken an in depth look at graphical representations of systems of equations, as well as how to find possible solutions graphically. Our attention now turns to working with systems algebraically.

Definition 2.2.1: System of Linear Equations

A **system of linear equations** is a list of equations,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

where a_{ij} and b_j are real numbers. The above is a system of m equations in the n variables, $x_1, x_2 \cdots, x_n$. Written more simply in terms of summation notation, the above can be written in the form

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \ i = 1, 2, 3, \cdots, m$$

The relative size of m and n is not important here. Notice that we have allowed a_{ij} and b_j to be any real number. We can also call these numbers **scalars** . We will use this term throughout the text, so keep in mind that the term **scalar** just means that we are working with real numbers.

Now, suppose we have a system where $b_i = 0$ for all i. In other words every equation equals 0. This is a special type of system.

Definition 2.2.2: Homogeneous System of Equations

A system of equations is called **homogeneous** if each equation in the system is equal to 0. A homogeneous system has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

where a_{ij} are scalars and x_i are variables.

Recall from the previous section that our goal when working with systems of linear equations was to find the point of intersection of the equations when graphed. In other words, we looked for the solutions to the system. We now wish to find these solutions algebraically. We want to find values for $x_1, \cdots, x_n$ which solve all of the equations. If such a set of values exists, we call $(x_1, \cdots, x_n)$ the **solution set**.

Recall the above discussions about the types of solutions possible. We will see that systems of linear equations will have one unique solution, infinitely many solutions, or no solution. Consider the following definition.

Definition 2.2.3: Consistent and Inconsistent Systems

A system of linear equations is called **consistent** if there exists at least one solution. It is called **inconsistent** if there is no solution.

If you think of each equation as a condition which must be satisfied by the variables, consistent would mean there is some choice of variables which can satisfy **all** the conditions. Inconsistent would mean there is no choice of the variables which can satisfy all of the conditions.

The following sections provide methods for determining if a system is consistent or inconsistent, and finding solutions if they exist.

Elementary Operations

We begin this section with an example. Recall from Example 1.1.1 that the solution to the given system was $(x, y) = (-1, 4)$.

Example 2.2.1: Verifying an Ordered Pair is a Solution

Algebraically verify that $(x, y) = (-1, 4)$ is a solution to the following system of equations.

$$\begin{aligned} x + y &= 3 \\ y - x &= 5 \end{aligned}$$

Solution

By graphing these two equations and identifying the point of intersection, we previously found that $(x, y) = (-1, 4)$ is the unique solution.

We can verify algebraically by substituting these values into the original equations, and ensuring that the equations hold. First, we substitute the values into the first equation and check that it equals 3.

$$x + y = (-1) + (4) = 3$$

This equals 3 as needed, so we see that $(-1, 4)$ is a solution to the first equation. Substituting the values into the second equation yields

$$y - x = (4) - (-1) = 4 + 1 = 5$$

which is true. For $(x, y) = (-1, 4)$ each equation is true and therefore, this is a solution to the system.

Now, the interesting question is this: If you were not given these numbers to verify, how could you algebraically determine the solution? Linear algebra gives us the tools needed to answer this question. The following basic operations are important tools that we will utilize.

Definition 2.2.4: Elementary Operations

Elementary operations are those operations consisting of the following.

1. Interchange the order in which the equations are listed.
2. Multiply any equation by a nonzero number.
3. Replace any equation with itself added to a multiple of another equation.

It is important to note that none of these operations will change the set of solutions of the system of equations. In fact, elementary operations are the *key tool* we use in linear algebra to find solutions to systems of equations.

Consider the following example.

Example 2.2.2: Effects of an Elementary Operation

Show that the system

$$\begin{aligned} x+y&=7\\ 2x-y&=8 \end{aligned}$$

has the same solution as the system

$$\begin{aligned} x+y&=7\\ -3y&=-6 \end{aligned}$$

Solution

Notice that the second system has been obtained by taking the second equation of the first system and adding -2 times the first equation, as follows:

$$2x-y+(-2)(x+y)=8+(-2)(7)$$

By simplifying, we obtain

$$-3y=-6$$

which is the second equation in the second system. Now, from here we can solve for y and see that $y=2$. Next, we substitute this value into the first equation as follows

$$x+y=x+2=7$$

Hence $x=5$ and so $(x,y)=(5,2)$ is a solution to the second system. We want to check if $(5,2)$ is also a solution to the first system. We check this by substituting $(x,y)=(5,2)$ into the system and ensuring the equations are true.

$$\begin{aligned} x+y&=(5)+(2)=7\\ 2x-y&=2\,(5)-(2)=8 \end{aligned}$$

Hence, $(5,2)$ is also a solution to the first system.

This example illustrates how an elementary operation applied to a system of two equations in two variables does not affect the solution set. However, a linear system may involve many equations and many variables and there is no reason to limit our study to small systems. For any size of system in any number of variables, the solution set is still the collection of solutions to the equations. In every case, the above operations of Definition 2.2.4 do not change the set of solutions to the system of linear equations.

In the following theorem, we use the notation E_i to represent an equation, while b_i denotes a constant.

Theorem 2.2.1: Elementary Operations and Solutions

Suppose you have a system of two linear equations

$$\begin{aligned} E_1&=b_1\\ E_2&=b_2 \end{aligned} \tag{2.2.1}$$

Then the following systems have the same solution set as (2.2.1):

1. $$\begin{aligned} E_2&=b_2\\ E_1&=b_1 \end{aligned} \tag{2.2.2}$$

2. $$\begin{aligned} E_1&=b_1\\ kE_2&=kb_2 \end{aligned} \tag{2.2.3}$$

   for any scalar k, provided $k\neq 0$.

3. $$\begin{aligned} E_1&=b_1\\ E_2+kE_1&=b_2+kb_1 \end{aligned} \tag{2.2.4}$$

   for any scalar k (including $k=0$).

Before we proceed with the proof of Theorem 2.2.1, let us consider this theorem in context of Example 2.2.2. Then,

$$\begin{aligned} E_1 &= x + y, \quad & b_1 &= 7 \\ E_2 &= 2x - y, \quad & b_2 &= 8 \end{aligned}$$

Recall the elementary operations that we used to modify the system in the solution to the example. First, we added (-2) times the first equation to the second equation. In terms of Theorem 2.2.1, this action is given by

$$E_2 + (-2)\, E_1 = b_2 + (-2)\, b_1$$

or

$$2x - y + (-2)\,(x + y) = 8 + (-2)\,7$$

This gave us the second system in Example 2.2.2, given by

$$\begin{aligned} E_1 &= b_1 \\ E_2 + (-2)\, E_1 &= b_2 + (-2)\, b_1 \end{aligned}$$

From this point, we were able to find the solution to the system. Theorem 2.2.1 tells us that the solution we found is in fact a solution to the original system.

We will now prove Theorem 2.2.1.

Proof

1. The proof that the systems (2.2.1) and (2.2.2) have the same solution set is as follows. Suppose that $(x_1, \cdots, x_n)$ is a solution to $E_1 = b_1, E_2 = b_2$. We want to show that this is a solution to the system in (2.2.2) above. This is clear, because the system in (2.2.2) is the original system, but listed in a different order. Changing the order does not effect the solution set, so $(x_1, \cdots, x_n)$ is a solution to (2.2.2).
2. Next we want to prove that the systems (2.2.1) and (2.2.3) have the same solution set. That is $E_1 = b_1, E_2 = b_2$ has the same solution set as the system $E_1 = b_1, kE_2 = kb_2$ provided $k \neq 0$. Let $(x_1, \cdots, x_n)$ be a solution of $E_1 = b_1, E_2 = b_2,$. We want to show that it is a solution to $E_1 = b_1, kE_2 = kb_2$. Notice that the only difference between these two systems is that the second involves multiplying the equation, $E_2 = b_2$ by the scalar k. Recall that when you multiply both sides of an equation by the same number, the sides are still equal to each other. Hence if $(x_1, \cdots, x_n)$ is a solution to $E_2 = b_2$, then it will also be a solution to $kE_2 = kb_2$. Hence, $(x_1, \cdots, x_n)$ is also a solution to (2.2.3).
   Similarly, let $(x_1, \cdots, x_n)$ be a solution of $E_1 = b_1, kE_2 = kb_2$. Then we can multiply the equation $kE_2 = kb_2$ by the scalar $1/k$, which is possible only because we have required that $k \neq 0$. Just as above, this action preserves equality and we obtain the equation $E_2 = b_2$. Hence $(x_1, \cdots, x_n)$ is also a solution to $E_1 = b_1, E_2 = b_2$.
3. Finally, we will prove that the systems (2.2.1) and (2.2.4) have the same solution set. We will show that any solution of $E_1 = b_1, E_2 = b_2$ is also a solution of (2.2.4). Then, we will show that any solution of (2.2.4) is also a solution of $E_1 = b_1, E_2 = b_2$. Let $(x_1, \cdots, x_n)$ be a solution to $E_1 = b_1, E_2 = b_2$. Then in particular it solves $E_1 = b_1$. Hence, it solves the first equation in (2.2.4). Similarly, it also solves $E_2 = b_2$. By our proof of (2.2.3), it also solves $kE_1 = kb_1$. Notice that if we add E_2 and kE_1, this is equal to $b_2 + kb_1$. Therefore, if $(x_1, \cdots, x_n)$ solves $E_1 = b_1, E_2 = b_2$ it must also solve $E_2 + kE_1 = b_2 + kb_1$.
   Now suppose $(x_1, \cdots, x_n)$ solves the system $E_1 = b_1, E_2 + kE_1 = b_2 + kb_1$. Then in particular it is a solution of $E_1 = b_1$. Again by our proof of (2.2.3), it is also a solution to $kE_1 = kb_1$. Now if we subtract these equal quantities from both sides of $E_2 + kE_1 = b_2 + kb_1$ we obtain $E_2 = b_2$, which shows that the solution also satisfies $E_1 = b_1, E_2 = b_2$.

Stated simply, the above theorem shows that the elementary operations do not change the solution set of a system of equations.

We will now look at an example of a system of three equations and three variables. Similarly to the previous examples, the goal is to find values for x, y, z such that each of the given equations are satisfied when these values are substituted in.

Example 2.2.3: Solving a System of Equations with Elementary Operations

Find the solutions to the system,

$$\begin{aligned} x+3y+6z &= 25 \\ 2x+7y+14z &= 58 \\ 2y+5z &= 19 \end{aligned} \tag{2.2.5}$$

Solution

We can relate this system to Theorem 2.2.1 above. In this case, we have

$$\begin{aligned} E_1 &= x+3y+6z, & b_1 &= 25 \\ E_2 &= 2x+7y+14z, & b_2 &= 58 \\ E_3 &= 2y+5z, & b_3 &= 19 \end{aligned}$$

Theorem 2.2.1 claims that if we do elementary operations on this system, we will not change the solution set. Therefore, we can solve this system using the elementary operations given in Definition 2.2.4. First, replace the second equation by (-2) times the first equation added to the second. This yields the system

$$\begin{aligned} x+3y+6z &= 25 \\ y+2z &= 8 \\ 2y+5z &= 19 \end{aligned} \tag{2.2.6}$$

Now, replace the third equation with (-2) times the second added to the third. This yields the system

$$\begin{aligned} x+3y+6z &= 25 \\ y+2z &= 8 \\ z &= 3 \end{aligned} \tag{2.2.7}$$

At this point, we can easily find the solution. Simply take $z=3$ and substitute this back into the previous equation to solve for y, and similarly to solve for x.

$$\begin{aligned} x+3y+6\left(3\right) = x+3y+18 &= 25 \\ y+2\left(3\right) = y+6 &= 8 \\ z &= 3 \end{aligned}$$

The second equation is now

$$y+6=8$$

You can see from this equation that $y=2$. Therefore, we can substitute this value into the first equation as follows:

$$x+3\left(2\right)+18=25$$

By simplifying this equation, we find that $x=1$. Hence, the solution to this system is $(x,y,z)=(1,2,3)$. This process is called **back substitution**.

Alternatively, in (2.2.7) you could have continued as follows. Add (-2) times the third equation to the second and then add (-6) times the second to the first. This yields

$$\begin{aligned} x+3y &= 7 \\ y &= 2 \\ z &= 3 \end{aligned}$$

Now add (-3) times the second to the first. This yields

$$\begin{aligned} x &= 1 \\ y &= 2 \\ z &= 3 \end{aligned}$$

a system which has the same solution set as the original system. This avoided back substitution and led to the same solution set. It is your decision which you prefer to use, as both methods lead to the correct solution, $(x, y, z) = (1, 2, 3)$.

This page titled 2.2: Systems of Equations, Algebraic Procedures is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.2: Systems of Equations, Algebraic Procedures** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

2.3: Gaussian Elimination

The work we did in the previous section will always find the solution to the system. In this section, we will explore a less cumbersome way to find the solutions. First, we will represent a linear system with an **augmented matrix.** A **matrix** is simply a rectangular array of numbers. The size or dimension of a matrix is defined as $m \times n$ where m is the number of rows and n is the number of columns. In order to construct an augmented matrix from a linear system, we create a **coefficient matrix** from the coefficients of the variables in the system, as well as a **constant matrix** from the constants. The coefficients from one equation of the system create one row of the augmented matrix.

For example, consider the linear system in Example 1.2.3

$$\begin{aligned} x+3y+6z &= 25 \\ 2x+7y+14z &= 58 \\ 2y+5z &= 19 \end{aligned}$$

This system can be written as an augmented matrix, as follows

$$\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \\ 2 & 7 & 14 & 58 \\ 0 & 2 & 5 & 19 \end{array}\right]$$

Notice that it has exactly the same information as the original system. Here it is understood that the first column contains the coefficients from x in each equation, in order, $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$. Similarly, we create a column from the coefficients on y in each equation, $\begin{bmatrix} 3 \\ 7 \\ 2 \end{bmatrix}$ and a column from the coefficients on z in each equation, $\begin{bmatrix} 6 \\ 14 \\ 5 \end{bmatrix}$. For a system of more than three variables, we would continue in this way constructing a column for each variable. Similarly, for a system of less than three variables, we simply construct a column for each variable.

Finally, we construct a column from the constants of the equations, $\begin{bmatrix} 25 \\ 58 \\ 19 \end{bmatrix}$.

The rows of the augmented matrix correspond to the equations in the system. For example, the top row in the augmented matrix, $\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \end{array}\right]$ corresponds to the equation

$$x+3y+6z=25.$$

Consider the following definition.

Definition 2.3.1: Augmented Matrix of a Linear System

For a linear system of the form

$$\begin{aligned} a_{11}x_1+\cdots+a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1+\cdots+a_{mn}x_n &= b_m \end{aligned}$$

where the x_i are variables and the a_{ij} and b_i are constants, the augmented matrix of this system is given by

$$\left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

Now, consider elementary operations in the context of the augmented matrix. The elementary operations in Definition 1.2.4 can be used on the rows just as we used them on equations previously. Changes to a system of equations in as a result of an elementary operation are equivalent to changes in the augmented matrix resulting from the corresponding row operation. Note that Theorem 1.2.1 implies that any elementary row operations used on an augmented matrix will not change the solution to the corresponding system of equations. We now formally define elementary row operations. These are the *key tool* we will use to find solutions to systems of equations.

Definition 2.3.2: Elementary Row Operations

The **elementary row operations** (also known as **row operations**) consist of the following

1. Switch two rows.
2. Multiply a row by a nonzero number.
3. Replace a row by any multiple of another row added to it.

Recall how we solved Example 1.2.3. We can do the exact same steps as above, except now in the context of an augmented matrix and using row operations. The augmented matrix of this system is

$$\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \\ 2 & 7 & 14 & 58 \\ 0 & 2 & 5 & 19 \end{array}\right]$$

Thus the first step in solving the system given by (1.2.5) would be to take (-2) times the first row of the augmented matrix and add it to the second row,

$$\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \\ 0 & 1 & 2 & 8 \\ 0 & 2 & 5 & 19 \end{array}\right]$$

Note how this corresponds to (1.2.6). Next take (-2) times the second row and add to the third,

$$\left[\begin{array}{ccc|c} 1 & 3 & 6 & 25 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

This augmented matrix corresponds to the system

$$\begin{aligned} x+3y+6z &= 25 \\ y+2z &= 8 \\ z &= 3 \end{aligned}$$

which is the same as (1.2.7). By back substitution you obtain the solution $x=1, y=2$, and $z=3$.

Through a systematic procedure of row operations, we can simplify an augmented matrix and carry it to **row-echelon form** or **reduced row-echelon form**, which we define next. These forms are used to find the solutions of the system of equations corresponding to the augmented matrix.

In the following definitions, the term **leading entry** refers to the first nonzero entry of a row when scanning the row from left to right.

Definition 2.3.3: Row-Echelon Form

An augmented matrix is in **row-echelon form** if

1. All nonzero rows are above any rows of zeros.
2. Each leading entry of a row is in a column to the right of the leading entries of any row above it.
3. Each leading entry of a row is equal to 1.

We also consider another reduced form of the augmented matrix which has one further condition.

Definition 2.3.4: Reduced Row-Echelon Form

An augmented matrix is in **reduced row-echelon form** if

1. All nonzero rows are above any rows of zeros.
2. Each leading entry of a row is in a column to the right of the leading entries of any rows above it.
3. Each leading entry of a row is equal to 1.
4. All entries in a column above and below a leading entry are zero.

Notice that the first three conditions on a reduced row-echelon form matrix are the same as those for row-echelon form.

Hence, every reduced row-echelon form matrix is also in row-echelon form. The converse is not necessarily true; we cannot assume that every matrix in row-echelon form is also in reduced row-echelon form. However, it often happens that the row-echelon form is sufficient to provide information about the solution of a system.

The following examples describe matrices in these various forms. As an exercise, take the time to carefully verify that they are in the specified form.

Example 2.3.1: Not in Row-Echelon Form

The following augmented matrices are not in row-echelon form (and therefore also not in reduced row-echelon form).

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 4 & -6 \\ 4 & 0 & 7 \end{array}\right], \left[\begin{array}{ccc|c} 0 & 2 & 3 & 3 \\ 1 & 5 & 0 & 2 \\ 7 & 5 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

Example 2.3.2: Matrices in Row-Echelon Form

The following augmented matrices are in row-echelon form, but not in reduced row-echelon form.

$$\left[\begin{array}{ccccc|c} 1 & 0 & 6 & 5 & 8 & 2 \\ 0 & 0 & 1 & 2 & 7 & 3 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{cccc|c} 1 & 3 & 5 & 4 & 4 \\ 0 & 1 & 0 & 7 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \\ 0 & 0 & 0 & 0 & \end{array}\right], \left[\begin{array}{ccc|c} 1 & 0 & 6 & 0 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Notice that we could apply further row operations to these matrices to carry them to reduced row-echelon form. Take the time to try that on your own. Consider the following matrices, which are in reduced row-echelon form.

Example 2.3.3: Matrices in Reduced Row-Echelon Form

The following augmented matrices are in reduced row-echelon form.

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

One way in which the row-echelon form of a matrix is useful is in identifying the pivot positions and pivot columns of the matrix.

Definition 2.3.5: Pivot Position and Pivot Column

A **pivot position** in a matrix is the location of a leading entry in the row-echelon form of a matrix.

A **pivot column** is a column that contains a pivot position.

For example consider the following.

Example 2.3.4: Pivot Position

Let

$$A = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 6 \\ 4 & 4 & 4 & 10 \end{array}\right]$$

Where are the pivot positions and pivot columns of the augmented matrix A?

Solution

The row-echelon form of this matrix is

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & \frac{3}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

This is all we need in this example, but note that this matrix is not in reduced row-echelon form.

In order to identify the pivot positions in the original matrix, we look for the leading entries in the row-echelon form of the matrix. Here, the entry in the first row and first column, as well as the entry in the second row and second column are the leading entries. Hence, these locations are the pivot positions. We identify the pivot positions in the original matrix, as in the following:

$$\left[\begin{array}{ccc|c} \boxed{1} & 2 & 3 & 4 \\ 3 & \boxed{2} & 1 & 6 \\ 4 & 4 & 4 & 10 \end{array}\right]$$

Thus the pivot columns in the matrix are the first two columns.

The following is an algorithm for carrying a matrix to row-echelon form and reduced row-echelon form. You may wish to use this algorithm to carry the above matrix to row-echelon form or reduced row-echelon form yourself for practice.

Algorithm 2.3.1: Reduced Row-Echelon Form Algorithm

This algorithm provides a method for using row operations to take a matrix to its reduced row-echelon form. We begin with the matrix in its original form.

1. Starting from the left, find the first nonzero column. This is the first pivot column, and the position at the top of this column is the first pivot position. Switch rows if necessary to place a nonzero number in the first pivot position.
2. Use row operations to make the entries below the first pivot position (in the first pivot column) equal to zero.
3. Ignoring the row containing the first pivot position, repeat steps 1 and 2 with the remaining rows. Repeat the process until there are no more rows to modify.
4. Divide each nonzero row by the value of the leading entry, so that the leading entry becomes 1. The matrix will then be in row-echelon form.

   The following step will carry the matrix from row-echelon form to reduced row-echelon form.

5. Moving from right to left, use row operations to create zeros in the entries of the pivot columns which are above the pivot positions. The result will be a matrix in reduced row-echelon form.

Most often we will apply this algorithm to an augmented matrix in order to find the solution to a system of linear equations. However, we can use this algorithm to compute the reduced row-echelon form of any matrix which could be useful in other applications.

Consider the following example of Algorithm 2.3.1.

Example 2.3.5: Finding Row-Echelon Form and Reduced Row-Echelon Form of a Matrix

Let

$$A = \begin{bmatrix} 0 & -5 & -4 \\ 1 & 4 & 3 \\ 5 & 10 & 7 \end{bmatrix}$$

Find the row-echelon form of A. Then complete the process until A is in reduced row-echelon form.

Solution

In working through this example, we will use the steps outlined in Algorithm 2.3.1.

1. The first pivot column is the first column of the matrix, as this is the first nonzero column from the left. Hence the first pivot position is the one in the first row and first column. Switch the first two rows to obtain a nonzero entry in the first pivot position, outlined in a box below.

$$\begin{bmatrix} \boxed{1} & 4 & 3 \\ 0 & -5 & -4 \\ 5 & 10 & 7 \end{bmatrix}$$

2. Step two involves creating zeros in the entries below the first pivot position. The first entry of the second row is already a zero. All we need to do is subtract 5 times the first row from the third row. The resulting matrix is

$$\begin{bmatrix} 1 & 4 & 3 \\ 0 & -5 & -4 \\ 0 & 10 & 8 \end{bmatrix}$$

3. Now ignore the top row. Apply steps 1 and 2 to the smaller matrix

$$\begin{bmatrix} -5 & -4 \\ 10 & 8 \end{bmatrix}$$

In this matrix, the first column is a pivot column, and -5 is in the first pivot position. Therefore, we need to create a zero below it. To do this, add 2 times the first row (of this matrix) to the second. The resulting matrix is

$$\begin{bmatrix} -5 & -4 \\ 0 & 0 \end{bmatrix}$$

Our original matrix now looks like

$$\begin{bmatrix} 1 & 4 & 3 \\ 0 & -5 & -4 \\ 0 & 0 & 0 \end{bmatrix}$$

We can see that there are no more rows to modify.

4. Now, we need to create leading 1s in each row. The first row already has a leading 1 so no work is needed here. Divide the second row by -5 to create a leading 1. The resulting matrix is

$$\begin{bmatrix} 1 & 4 & 3 \\ 0 & 1 & \frac{4}{5} \\ 0 & 0 & 0 \end{bmatrix}$$

This matrix is now in row-echelon form.

5. Now create zeros in the entries above pivot positions in each column, in order to carry this matrix all the way to reduced row-echelon form. Notice that there is no pivot position in the third column so we do not need to create any zeros in this column! The column in which we need to create zeros is the second. To do so, subtract 4 times the second row from the first row. The resulting matrix is

$$\begin{bmatrix} 1 & 0 & -\frac{1}{5} \\ 0 & 1 & \frac{4}{5} \\ 0 & 0 & 0 \end{bmatrix}$$

This matrix is now in reduced row-echelon form.

The above algorithm gives you a simple way to obtain the row-echelon form and reduced row-echelon form of a matrix. The main idea is to do row operations in such a way as to end up with a matrix in row-echelon form or reduced row-echelon form. This process is important because the resulting matrix will allow you to describe the solutions to the corresponding linear system of equations in a meaningful way.

In the next example, we look at how to solve a system of equations using the corresponding augmented matrix.

Example 2.3.6: Finding the Solution to a System

Give the complete solution to the following system of equations

$$\begin{aligned} 2x + 4y - 3z &= -1 \\ 5x + 10y - 7z &= -2 \\ 3x + 6y + 5z &= 9 \end{aligned}$$

Solution

The augmented matrix for this system is

$$\left[\begin{array}{ccc|c} 2 & 4 & -3 & -1 \\ 5 & 10 & -7 & -2 \\ 3 & 6 & 5 & 9 \end{array}\right]$$

In order to find the solution to this system, we wish to carry the augmented matrix to reduced row-echelon form. We will do so using Algorithm 2.3.1. Notice that the first column is nonzero, so this is our first pivot column. The first entry in the first row, 2, is the first leading entry and it is in the first pivot position. We will use row operations to create zeros in the entries below the 2. First, replace the second row with -5 times the first row plus 2 times the second row. This yields

$$\left[\begin{array}{ccc|c} 2 & 4 & -3 & -1 \\ 0 & 0 & 1 & 1 \\ 3 & 6 & 5 & 9 \end{array}\right]$$

Now, replace the third row with -3 times the first row plus to 2 times the third row. This yields

$$\left[\begin{array}{ccc|c} 2 & 4 & -3 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 21 \end{array}\right]$$

Now the entries in the first column below the pivot position are zeros. We now look for the second pivot column, which in this case is column three. Here, the 1 in the second row and third column is in the pivot position. We need to do just one row operation to create a zero below the 1.

Taking -1 times the second row and adding it to the third row yields

$$\left[\begin{array}{ccc|c} 2 & 4 & -3 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 20 \end{array}\right]$$

We could proceed with the algorithm to carry this matrix to row-echelon form or reduced row-echelon form. However, remember that we are looking for the solutions to the system of equations. Take another look at the third row of the matrix. Notice that it corresponds to the equation

$$0x+0y+0z=20$$

There is no solution to this equation because for all x,y,z, the left side will equal 0 and $0\neq 20$. This shows there is no solution to the given system of equations. In other words, this system is inconsistent.

The following is another example of how to find the solution to a system of equations by carrying the corresponding augmented matrix to reduced row-echelon form.

Example 2.3.7: An Infinite Set of Solutions

Give the complete solution to the system of equations

$$\begin{aligned} 3x-y-5z&=9\\ y-10z&=0\\ -2x+y&=-6 \end{aligned} \tag{2.3.1}$$

Solution

The augmented matrix of this system is

$$\left[\begin{array}{ccc|c} 3 & -1 & -5 & 9\\ 0 & 1 & -10 & 0\\ -2 & 1 & 0 & -6 \end{array}\right]$$

In order to find the solution to this system, we will carry the augmented matrix to reduced row-echelon form, using Algorithm 2.3.1. The first column is the first pivot column. We want to use row operations to create zeros beneath the first entry in this column, which is in the first pivot position. Replace the third row with 2 times the first row added to 3 times the third row. This gives

$$\left[\begin{array}{ccc|c} 3 & -1 & -5 & 9\\ 0 & 1 & -10 & 0\\ 0 & 1 & -10 & 0 \end{array}\right]$$

Now, we have created zeros beneath the 3 in the first column, so we move on to the second pivot column (which is the second column) and repeat the procedure. Take -1 times the second row and add to the third row.

$$\left[\begin{array}{ccc|c} 3 & -1 & -5 & 9\\ 0 & 1 & -10 & 0\\ 0 & 0 & 0 & 0 \end{array}\right]$$

The entry below the pivot position in the second column is now a zero. Notice that we have no more pivot columns because we have only two leading entries.

At this stage, we also want the leading entries to be equal to one. To do so, divide the first row by 3.

$$\left[\begin{array}{ccc|c} 1 & -\frac{1}{3} & -\frac{5}{3} & 3\\ 0 & 1 & -10 & 0\\ 0 & 0 & 0 & 0 \end{array}\right]$$

This matrix is now in row-echelon form.

Let's continue with row operations until the matrix is in reduced row-echelon form. This involves creating zeros above the pivot positions in each pivot column. This requires only one step, which is to add $\frac{1}{3}$ times the second row to the first row.

$$\left[\begin{array}{ccc|c} 1 & 0 & -5 & 3 \\ 0 & 1 & -10 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

This is in reduced row-echelon form, which you should verify using Definition 2.3.4. The equations corresponding to this reduced row-echelon form are

$$x - 5z = 3$$
$$y - 10z = 0$$

or

$$x = 3 + 5z$$
$$y = 10z$$

Observe that z is not restrained by any equation. In fact, z can equal any number. For example, we can let $z = t$, where we can choose t to be any number. In this context t is called a **parameter** . Therefore, the solution set of this system is

$$x = 3 + 5t$$
$$y = 10t$$
$$z = t$$

where t is arbitrary. The system has an infinite set of solutions which are given by these equations. For any value of t we select, x, y, and z will be given by the above equations. For example, if we choose $t = 4$ then the corresponding solution would be

$$x = 3 + 5(4) = 23$$
$$y = 10(4) = 40$$
$$z = 4$$

In Example 2.3.7 the solution involved one parameter. It may happen that the solution to a system involves more than one parameter, as shown in the following example.

Example 2.3.8: A Two Parameter Set of Solutions

Find the solution to the system

$$x + 2y - z + w = 3$$
$$x + y - z + w = 1$$
$$x + 3y - z + w = 5$$

Solution

The augmented matrix is

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 1 & 1 & -1 & 1 & 1 \\ 1 & 3 & -1 & 1 & 5 \end{array}\right]$$

We wish to carry this matrix to row-echelon form. Here, we will outline the row operations used. However, make sure that you understand the steps in terms of Algorithm 2.3.1.

Take -1 times the first row and add to the second. Then take -1 times the first row and add to the third. This yields

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 0 & -1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 0 & 2 \end{array}\right]$$

Now add the second row to the third row and divide the second row by -1.

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{2.3.2}$$

This matrix is in row-echelon form and we can see that x and y correspond to pivot columns, while z and w do not. Therefore, we will assign parameters to the variables z and w. Assign the parameter s to z and the parameter t to w. Then the first row yields the equation $x+2y-s+t=3$, while the second row yields the equation $y=2$. Since $y=2$, the first equation becomes $x+4-s+t=3$ showing that the solution is given by

$$\begin{aligned} x &= -1+s-t \\ y &= 2 \\ z &= s \\ w &= t \end{aligned}$$

It is customary to write this solution in the form

$$\left[\begin{array}{c} x \\ y \\ z \\ w \end{array}\right] = \left[\begin{array}{c} -1+s-t \\ 2 \\ s \\ t \end{array}\right] \tag{2.3.3}$$

This example shows a system of equations with an infinite solution set which depends on two parameters. It can be less confusing in the case of an infinite solution set to first place the augmented matrix in reduced row-echelon form rather than just row-echelon form before seeking to write down the description of the solution.

In the above steps, this means we don't stop with the row-echelon form in equation (2.3.2). Instead we first place it in reduced row-echelon form as follows.

$$\left[\begin{array}{cccc|c} 1 & 0 & -1 & 1 & -1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Then the solution is $y=2$ from the second row and $x=-1+z-w$ from the first. Thus letting $z=s$ and $w=t$, the solution is given by (2.3.3).

You can see here that there are two paths to the correct answer, which both yield the same answer. Hence, either approach may be used. The process which we first used in the above solution is called **Gaussian Elimination** This process involves carrying the matrix to row-echelon form, converting back to equations, and using back substitution to find the solution. When you do row operations until you obtain reduced row-echelon form, the process is called **Gauss-Jordan Elimination**.

We have now found solutions for systems of equations with no solution and infinitely many solutions, with one parameter as well as two parameters. Recall the three types of solution sets which we discussed in the previous section; no solution, one solution, and infinitely many solutions. Each of these types of solutions could be identified from the graph of the system. It turns out that we can also identify the type of solution from the reduced row-echelon form of the augmented matrix.

- *No Solution:* In the case where the system of equations has no solution, the row-echelon form of the augmented matrix will have a row of the form

  $$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 1 \end{array}\right]$$

  This row indicates that the system is inconsistent and has no solution.
- *One Solution:* In the case where the system of equations has one solution, every column of the coefficient matrix is a pivot column. The following is an example of an augmented matrix in reduced row-echelon form for a system of equations with one solution.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

- *Infinitely Many Solutions:* In the case where the system of equations has infinitely many solutions, the solution contains parameters. There will be columns of the coefficient matrix which are not pivot columns. The following are examples of augmented matrices in reduced row-echelon form for systems of equations with infinitely many solutions.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

or

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -3 \end{array}\right]$$

This page titled 2.3: Gaussian Elimination is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.3: Gaussian Elimination** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

2.4: Uniqueness of the Reduced Row-Echelon Form

As we have seen in earlier sections, we know that every matrix can be brought into reduced row-echelon form by a sequence of elementary row operations. Here we will prove that the resulting matrix is unique; in other words, the resulting matrix in reduced row-echelon does not depend upon the particular sequence of elementary row operations or the order in which they were performed.

Let A be the augmented matrix of a homogeneous system of linear equations in the variables $x_1, x_2, \cdots, x_n$ which is also in reduced row-echelon form. The matrix A divides the set of variables in two different types. We say that x_i is a *basic variable* whenever A has a leading 1 in column number i, in other words, when column i is a pivot column. Otherwise we say that x_i is a *free variable*.

Recall Example 1.3.8.

Example 2.4.1: Basic and Free Variables

Find the basic and free variables in the system

$$\begin{aligned} x+2y-z+w &= 3 \\ x+y-z+w &= 1 \\ x+3y-z+w &= 5 \end{aligned}$$

Solution

Recall from the solution of Example 1.3.8 that the row-echelon form of the augmented matrix of this system is given by

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

You can see that columns 1 and 2 are pivot columns. These columns correspond to variables x and y, making these the basic variables. Columns 3 and 4 are not pivot columns, which means that z and w are free variables.

We can write the solution to this system as

$$\begin{aligned} x &= -1+s-t \\ y &= 2 \\ z &= s \\ w &= t \end{aligned}$$

Here the free variables are written as parameters, and the basic variables are given by linear functions of these parameters.

In general, all solutions can be written in terms of the free variables. In such a description, the free variables can take any values (they become parameters), while the basic variables become simple linear functions of these parameters. Indeed, a basic variable x_i is a linear function of *only* those free variables x_j with $j > i$. This leads to the following observation.

Proposition 2.4.1: Basic and Free Variables

If x_i is a basic variable of a homogeneous system of linear equations, then any solution of the system with $x_j = 0$ for all those free variables x_j with $j > i$ must also have $x_i = 0$.

Using this proposition, we prove a lemma which will be used in the proof of the main result of this section below.

Lemma 2.4.1: Solutions and the Reduced Row-Echelon Form of a Matrix

Let A and B be two distinct augmented matrices for two homogeneous systems of m equations in n variables, such that A and B are each in reduced row-echelon. Then, the two systems do not have exactly the same solutions.

Proof

With respect to the linear systems associated with the matrices A and B, there are two cases to consider:

- Case 1: the two systems have the same basic variables
- Case 2: the two systems do not have the same basic variables

In case 1, the two matrices will have exactly the same pivot positions. However, since A and B are not identical, there is some row of A which is different from the corresponding row of B and yet the rows each have a pivot in the same column position. Let i be the index of this column position. Since the matrices are in reduced row-echelon form, the two rows must differ at some entry in a column $j > i$. Let these entries be a in A and b in B, where $a \neq b$. Since A is in reduced row-echelon form, if x_j were a basic variable for its linear system, we would have $a = 0$. Similarly, if x_j were a basic variable for the linear system of the matrix B, we would have $b = 0$. Since a and b are unequal, they cannot both be equal to 0, and hence x_j cannot be a basic variable for both linear systems. However, since the systems have the same basic variables, x_j must then be a free variable for each system. We now look at the solutions of the systems in which x_j is set equal to 1 and all other free variables are set equal to 0. For this choice of parameters, the solution of the system for matrix A has $x_j = -a$, while the solution of the system for matrix B has $x_j = -b$, so that the two systems have different solutions.

In case 2, there is a variable x_i which is a basic variable for one matrix, let's say A, and a free variable for the other matrix B. The system for matrix B has a solution in which $x_i = 1$ and $x_j = 0$ for all other free variables x_j. However, by Proposition 2.4.1 this cannot be a solution of the system for the matrix A. This completes the proof of case 2.

Now, we say that the matrix B is **equivalent** to the matrix A provided that B can be obtained from A by performing a sequence of elementary row operations beginning with A. The importance of this concept lies in the following result.

Theorem 2.4.1: Equivalent Matrices

The two linear systems of equations corresponding to two equivalent augmented matrices have exactly the same solutions.

Proof

The proof of this theorem is left as an exercise.

Now, we can use Lemma 2.4.1 and Theorem 2.4.1 to prove the main result of this section.

Theorem 2.4.2: Uniqueness of the Reduced Row-Echelon Form

Every matrix A is equivalent to a unique matrix in reduced row-echelon form.

Proof

Let A be an $m \times n$ matrix and let B and C be matrices in reduced row-echelon form, each equivalent to A. It suffices to show that $B = C$.

Let A^+ be the matrix A augmented with a new rightmost column consisting entirely of zeros. Similarly, augment matrices B and C each with a rightmost column of zeros to obtain B^+ and C^+. Note that B^+ and C^+ are matrices in reduced row-echelon form which are obtained from A^+ by respectively applying the same sequence of elementary row operations which were used to obtain B and C from A.

Now, A^+, B^+, and C^+ can all be considered as augmented matrices of homogeneous linear systems in the variables $x_1, x_2, \cdots, x_n$. Because B^+ and C^+ are each equivalent to A^+, Theorem 2.4.1 ensures that all three homogeneous linear systems have exactly the same solutions. By Lemma 2.4.1 we conclude that $B^+ = C^+$. By construction, we must also have $B = C$.

According to this theorem we can say that each matrix A has a unique reduced row-echelon form.

This page titled 2.4: Uniqueness of the Reduced Row-Echelon Form is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.4: Uniqueness of the Reduced Row-Echelon Form** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

2.5: Rank and Homogeneous Systems

Rank and Homogeneous Systems

There is a special type of system which requires additional study. This type of system is called a homogeneous system of equations, which we defined above in Definition 1.2.3. Our focus in this section is to consider what types of solutions are possible for a homogeneous system of equations.

Consider the following definition.

Consider the homogeneous system of equations given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Then, $x_1 = 0, x_2 = 0, \cdots, x_n = 0$ is always a solution to this system. We call this the **trivial solution** .

If the system has a solution in which not all of the $x_1, \cdots, x_n$ are equal to zero, then we call this solution **nontrivial** . The trivial solution does not tell us much about the system, as it says that $0 = 0$! Therefore, when working with homogeneous systems of equations, we want to know when the system has a nontrivial solution.

Suppose we have a homogeneous system of m equations, using n variables, and suppose that $n > m$. In other words, there are more variables than equations. Then, it turns out that this system always has a nontrivial solution. Not only will the system have a nontrivial solution, but it also will have infinitely many solutions. It is also possible, but not required, to have a nontrivial solution if $n = m$ and $n < m$.

Consider the following example.

Example 2.5.1: Solutions to a Homogeneous System of Equations

Find the nontrivial solutions to the following homogeneous system of equations

$$\begin{aligned} 2x + y - z &= 0 \\ x + 2y - 2z &= 0 \end{aligned}$$

Solution

Notice that this system has $m = 2$ equations and $n = 3$ variables, so $n > m$. Therefore by our previous discussion, we expect this system to have infinitely many solutions.

The process we use to find the solutions for a homogeneous system of equations is the same process we used in the previous section. First, we construct the augmented matrix, given by

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 0 \\ 1 & 2 & -2 & 0 \end{array}\right]$$

Then, we carry this matrix to its reduced row-echelon form, given below.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right]$$

The corresponding system of equations is

$$\begin{aligned} x &= 0 \\ y - z &= 0 \end{aligned}$$

Since z is not restrained by any equation, we know that this variable will become our parameter. Let $z = t$ where t is any number. Therefore, our solution has the form

$$x = 0$$
$$y = z = t$$
$$z = t$$

Hence this system has infinitely many solutions, with one parameter t.

Suppose we were to write the solution to the previous example in another form. Specifically,

$$x = 0$$
$$y = 0 + t$$
$$z = 0 + t$$

can be written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Notice that we have constructed a column from the constants in the solution (all equal to 0), as well as a column corresponding to the coefficients on t in each equation. While we will discuss this form of solution more in further chapters, for now consider the column of coefficients of the parameter t. In this case, this is the column $\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

There is a special name for this column, which is **basic solution**. The basic solutions of a system are columns constructed from the coefficients on parameters in the solution. We often denote basic solutions by X_1, X_2 etc., depending on how many solutions occur. Therefore, Example 2.5.1 has the basic solution $X_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

We explore this further in the following example.

Example 2.5.2: Basic Solutions of a Homogeneous System

Consider the following homogeneous system of equations.

$$x + 4y + 3z = 0$$
$$3x + 12y + 9z = 0$$

Find the basic solutions to this system.

Solution

The augmented matrix of this system and the resulting reduced row-echelon form are

$$\left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ 3 & 12 & 9 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 4 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

When written in equations, this system is given by

$$x + 4y + 3z = 0$$

Notice that only x corresponds to a pivot column. In this case, we will have two parameters, one for y and one for z. Let $y = s$ and $z = t$ for any numbers s and t. Then, our solution becomes

$$x = -4s - 3t$$
$$y = s$$
$$z = t$$

which can be written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + s \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

You can see here that we have two columns of coefficients corresponding to parameters, specifically one for s and one for t. Therefore, this system has two basic solutions! These are

$$X_1 = \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix}, X_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

Below is a video on homogeneous systems of equations



Below is a video on solving a system of equations and determining the particular solution.



Below is a video on solving Ax = 0.



Below is another video on solving Ax = 0.



We now present a new definition.

Definition 2.5.2: Linear Combination

Let $X_1, \cdots, X_n, V$ be column matrices. Then V is said to be a **linear combination** of the columns $X_1, \cdots, X_n$ if there exist scalars, $a_1, \cdots, a_n$ such that

$$V = a_1 X_1 + \cdots + a_n X_n$$

A remarkable result of this section is that a linear combination of the basic solutions is again a solution to the system. Even more remarkable is that every solution can be written as a linear combination of these solutions. Therefore, if we take a linear combination of the two solutions to Example 2.5.2, this would also be a solution. For example, we could take the following linear combination

$$3\begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -18 \\ 3 \\ 2 \end{bmatrix}$$

You should take a moment to verify that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -18 \\ 3 \\ 2 \end{bmatrix}$$

is in fact a solution to the system in Example 2.5.2.

Another way in which we can find out more information about the solutions of a homogeneous system is to consider the **rank** of the associated coefficient matrix. We now define what is meant by the rank of a matrix.

Definition 2.5.3: Rank of a Matrix

Let A be a matrix and consider any row-echelon form of A. Then, the number r of leading entries of A does not depend on the row-echelon form you choose, and is called the **rank** of A. We denote it by Rank(A).

Similarly, we could count the number of pivot positions (or pivot columns) to determine the rank of A.

Example 2.5.3: Finding the Rank of a Matrix

Consider the matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 5 & 9 \\ 2 & 4 & 6 \end{bmatrix}$$

What is its rank?

Solution

First, we need to find the reduced row-echelon form of A. Through the usual algorithm, we find that this is

$$\begin{bmatrix} \boxed{1} & 0 & -1 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Here we have two leading entries, or two pivot positions, shown above in boxes.The rank of A is $r = 2$.

Notice that we would have achieved the same answer if we had found the row-echelon form of A instead of the reduced row-echelon form.

Suppose we have a homogeneous system of m equations in n variables, and suppose that $n > m$. From our above discussion, we know that this system will have infinitely many solutions. If we consider the rank of the coefficient matrix of this system, we can find out even more about the solution. Note that we are looking at just the coefficient matrix, not the entire augmented matrix.

Theorem 2.5.1: Rank and Solutions to a Homogeneous System

Let A be the $m \times n$ coefficient matrix corresponding to a homogeneous system of equations, and suppose A has rank r. Then, the solution to the corresponding system has $n - r$ parameters.

Consider our above Example 2.5.2 in the context of this theorem. The system in this example has $m = 2$ equations in $n = 3$ variables. First, because $n > m$, we know that the system has a nontrivial solution, and therefore infinitely many solutions. This tells us that the solution will contain at least one parameter. The rank of the coefficient matrix can tell us even more about the solution! The rank of the coefficient matrix of the system is 1, as it has one leading entry in row-echelon form. Theorem 2.5.1 tells us that the solution will have $n - r = 3 - 1 = 2$ parameters. You can check that this is true in the solution to Example 2.5.2.

Notice that if $n = m$ or $n < m$, it is possible to have either a unique solution (which will be the trivial solution) or infinitely many solutions.

We are not limited to homogeneous systems of equations here. The rank of a matrix can be used to learn about the solutions of any system of linear equations. In the previous section, we discussed that a system of equations can have no solution, a unique solution, or infinitely many solutions. Suppose the system is consistent, whether it is homogeneous or not. The following theorem tells us how we can use the rank to learn about the type of solution we have.

Below is a video on Rank, Basic Variables, and Free Variables of a coefficient matrix.



Theorem 2.5.2: Rank and Solutions to a Consistent System of Equations

Let A be the $m \times (n+1)$ augmented matrix corresponding to a consistent system of equations in n variables, and suppose A has rank r. Then

1. the system has a unique solution if $r = n$
2. the system has infinitely many solutions if $r < n$

We will not present a formal proof of this, but consider the following discussions.

1. *No Solution* The above theorem assumes that the system is consistent, that is, that it has a solution. It turns out that it is possible for the augmented matrix of a system with no solution to have any rank r as long as $r > 1$. Therefore, we must know that the system is consistent in order to use this theorem!
2. *Unique Solution* Suppose $r = n$. Then, there is a pivot position in every column of the coefficient matrix of A. Hence, there is a unique solution.
3. *Infinitely Many Solutions* Suppose $r < n$. Then there are infinitely many solutions. There are less pivot positions (and hence less leading entries) than columns, meaning that not every column is a pivot column. The columns which are *not* pivot columns correspond to parameters. In fact, in this case we have $n - r$ parameters.

Below is a video on the geometry of Ax = 0 and Ax = b.



Below is another video on the geometry of Ax = 0 and Ax = b.



Below is a video on the geometry of Ax = 0 and Ax = b, this one with two free variables.



This page titled 2.5: Rank and Homogeneous Systems is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.5: Rank and Homogeneous Systems** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

2.6: Balancing Chemical Reactions

The tools of linear algebra can also be used in the subject area of Chemistry, specifically for balancing chemical reactions.

Consider the chemical reaction

$$SnO_2 + H_2 \rightarrow Sn + H_2O$$

Here the elements involved are tin (Sn), oxygen (O), and hydrogen (H). A chemical reaction occurs and the result is a combination of tin (Sn) and water (H_2O). When considering chemical reactions, we want to investigate how much of each element we began with and how much of each element is involved in the result.

An important theory we will use here is the mass balance theory. It tells us that we cannot create or delete elements within a chemical reaction. For example, in the above expression, we must have the same number of oxygen, tin, and hydrogen on both sides of the reaction. Notice that this is not currently the case. For example, there are two oxygen atoms on the left and only one on the right. In order to fix this, we want to find numbers x, y, z, w such that

$$xSnO_2 + yH_2 \rightarrow zSn + wH_2O$$

where both sides of the reaction have the same number of atoms of the various elements.

This is a familiar problem. We can solve it by setting up a system of equations in the variables x, y, z, w. Thus you need

$$\begin{aligned} Sn: &\quad x = z \\ O: &\quad 2x = w \\ H: &\quad 2y = 2w \end{aligned}$$

We can rewrite these equations as

$$\begin{aligned} Sn: &\quad x - z = 0 \\ O: &\quad 2x - w = 0 \\ H: &\quad 2y - 2w = 0 \end{aligned}$$

The augmented matrix for this system of equations is given by

$$\left[\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 2 & 0 & 0 & -1 & 0 \\ 0 & 2 & 0 & -2 & 0 \end{array}\right]$$

The reduced row-echelon form of this matrix is

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & 0 \end{array}\right]$$

The solution is given by

$$\begin{aligned} x - \tfrac{1}{2}w &= 0 \\ y - w &= 0 \\ z - \tfrac{1}{2}w &= 0 \end{aligned}$$

which we can write as

$$\begin{aligned} x &= \tfrac{1}{2}t \\ y &= t \\ z &= \tfrac{1}{2}t \\ w &= t \end{aligned}$$

For example, let $w = 2$ and this would yield $x = 1, y = 2$, and $z = 1$. We can put these values back into the expression for the reaction which yields

$$SnO_2 + 2H_2 \rightarrow Sn + 2H_2O$$

Observe that each side of the expression contains the same number of atoms of each element. This means that it preserves the total number of atoms, as required, and so the chemical reaction is balanced.

Consider another example.

Example 2.6.1: Balancing a Chemical Reaction

Potassium is denoted by K, oxygen by O, phosphorus by P and hydrogen by H. Consider the reaction given by

$$KOH + H_3PO_4 \rightarrow K_3PO_4 + H_2O$$

Balance this chemical reaction.

Solution

We will use the same procedure as above to solve this problem. We need to find values for x, y, z, w such that

$$xKOH + yH_3PO_4 \rightarrow zK_3PO_4 + wH_2O$$

preserves the total number of atoms of each element.

Finding these values can be done by finding the solution to the following system of equations.

$$\begin{aligned} K: &\quad x = 3z \\ O: &\quad x + 4y = 4z + w \\ H: &\quad x + 3y = 2w \\ P: &\quad y = z \end{aligned}$$

The augmented matrix for this system is

$$\left[\begin{array}{cccc|c} 1 & 0 & -3 & 0 & 0 \\ 1 & 4 & -4 & -1 & 0 \\ 1 & 3 & 0 & -2 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right]$$

and the reduced row-echelon form is

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

The solution is given by

$$\begin{aligned} x - w &= 0 \\ y - \tfrac{1}{3}w &= 0 \\ z - \tfrac{1}{3}w &= 0 \end{aligned}$$

which can be written as

$$\begin{aligned} x &= t \\ y &= \tfrac{1}{3}t \\ z &= \tfrac{1}{3}t \\ w &= t \end{aligned}$$

Choose a value for t, say 3. Then $w = 3$ and this yields $x = 3, y = 1, z = 1$. It follows that the balanced reaction is given by

$$3KOH + 1H_3PO_4 \rightarrow 1K_3PO_4 + 3H_2O$$

Note that this results in the same number of atoms on both sides.

Of course these numbers you are finding would typically be the number of moles of the molecules on each side. Thus three moles of KOH added to one mole of H_3PO_4 yields one mole of K_3PO_4 and three moles of H_2O.

This page titled 2.6: Balancing Chemical Reactions is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.6: Balancing Chemical Reactions** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

2.7: Dimensionless Variables

This section shows how solving systems of equations can be used to determine appropriate dimensionless variables. It is only an introduction to this topic and considers a specific example of a simple airplane wing shown below. We assume for simplicity that it is a flat plane at an angle to the wind which is blowing against it with speed V as shown.

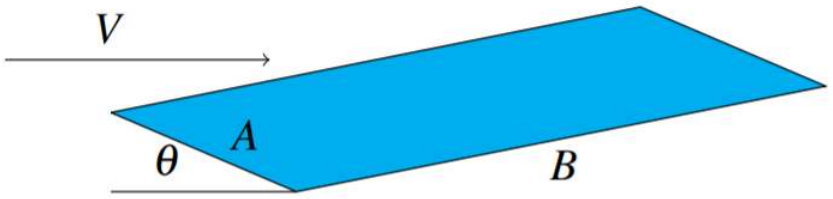


Figure 2.7.1

The angle θ is called the angle of incidence, B is the span of the wing and A is called the chord. Denote by l the lift. Then this should depend on various quantities like θ, V, B, A and so forth. Here is a table which indicates various quantities on which it is reasonable to expect l to depend.

Table 2.7.1

Variable	Symbol	Units
chord	A	m
span	B	m
angle incidence	θ	$m^0 kg^0 sec^0$
speed of wind	V	$msec^{-1}$
speed of sound	V_0	$msec^{-1}$
density of air	ρ	kgm^{-3}
viscosity	μ	$kgsec^{-1}m^{-1}$
lift	l	$kgsec^{-2}m$

Here m denotes meters, sec refers to seconds and kg refers to kilograms. All of these are likely familiar except for μ, which we will discuss in further detail now.

Viscosity is a measure of how much internal friction is experienced when the fluid moves. It is roughly a measure of how “sticky" the fluid is. Consider a piece of area parallel to the direction of motion of the fluid. To say that the viscosity is large is to say that the tangential force applied to this area must be large in order to achieve a given change in speed of the fluid in a direction normal to the tangential force. Thus

$$\mu\,(\text{area})\,(\text{velocity gradient}) = \text{ tangential force}$$

Hence

$$(\text{units on } \mu)\, m^2 \left(\frac{m}{\sec m}\right) = kg \sec^{-2} m$$

Thus the units on μ are

$$kg \sec^{-1} m^{-1}$$

as claimed above.

Returning to our original discussion, you may think that we would want

$$l = f\,(A, B, \theta, V, V_0, \rho, \mu)$$

This is very cumbersome because it depends on seven variables. Also, it is likely that without much care, a change in the units such as going from meters to feet would result in an incorrect value for l. The way to get around this problem is to look for l as a function of dimensionless variables multiplied by something which has units of force. It is helpful because first of all, you will likely have fewer independent variables and secondly, you could expect the formula to hold independent of the way of specifying length, mass and so forth. One looks for

$$l = f\left(g_1, \cdots, g_k\right) \rho V^2 AB$$

where the units on $\rho V^2 AB$ are

$$\frac{kg}{m^3}\left(\frac{m}{\sec}\right)^2 m^2 = \frac{kg \times m}{\sec^2}$$

which are the units of force. Each of these g_i is of the form

$$A^{x_1} B^{x_2} \theta^{x_3} V^{x_4} V_0^{x_5} \rho^{x_6} \mu^{x_7} \tag{2.7.1}$$

and each g_i is independent of the dimensions. That is, this expression must not depend on meters, kilograms, seconds, etc. Thus, placing in the units for each of these quantities, one needs

$$m^{x_1} m^{x_2} \left(m^{x_4} \sec^{-x_4}\right) \left(m^{x_5} \sec^{-x_5}\right) \left(kgm^{-3}\right)^{x_6} \left(kg \sec^{-1} m^{-1}\right)^{x_7} = m^0 kg^0 \sec^0$$

Notice that there are no units on θ because it is just the radian measure of an angle. Hence its dimensions consist of length divided by length, thus it is dimensionless. Then this leads to the following equations for the x_i.

$$\begin{aligned} m: &\quad x_1 + x_2 + x_4 + x_5 - 3x_6 - x_7 = 0 \\ \sec: &\quad -x_4 - x_5 - x_7 = 0 \\ kg: &\quad x_6 + x_7 = 0 \end{aligned}$$

The augmented matrix for this system is

$$\left[\begin{array}{ccccccc|c} 1 & 1 & 0 & 1 & 1 & -3 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

The reduced row-echelon form is given by

$$\left[\begin{array}{ccccccc|c} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right]$$

and so the solutions are of the form

$$\begin{aligned} x_1 &= -x_2 - x_7 \\ x_3 &= x_3 \\ x_4 &= -x_5 - x_7 \\ x_6 &= -x_7 \end{aligned}$$

Thus, in terms of vectors, the solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} -x_2 - x_7 \\ x_2 \\ x_3 \\ -x_5 - x_7 \\ x_5 \\ -x_7 \\ x_7 \end{bmatrix}$$

Thus the free variables are x_2, x_3, x_5, x_7. By assigning values to these, we can obtain dimensionless variables by placing the values obtained for the x_i in the formula (2.7.1). For example, let $x_2 = 1$ and all the rest of the free variables are 0. This yields

$$x_1 = -1, x_2 = 1, x_3 = 0, x_4 = 0, x_5 = 0, x_6 = 0, x_7 = 0$$

The dimensionless variable is then $A^{-1}B^1$. This is the ratio between the span and the chord. It is called the aspect ratio, denoted as AR. Next let $x_3 = 1$ and all others equal zero. This gives for a dimensionless quantity the angle θ. Next let $x_5 = 1$ and all others equal zero. This gives

$$x_1 = 0, x_2 = 0, x_3 = 0, x_4 = -1, x_5 = 1, x_6 = 0, x_7 = 0$$

Then the dimensionless variable is $V^{-1}V_0^1$. However, it is written as V/V_0. This is called the Mach number $\mathcal{M}$. Finally, let $x_7 = 1$ and all the other free variables equal 0. Then

$$x_1 = -1, x_2 = 0, x_3 = 0, x_4 = -1, x_5 = 0, x_6 = -1, x_7 = 1$$

then the dimensionless variable which results from this is $A^{-1}V^{-1}\rho^{-1}\mu$. It is customary to write it as $Re = (AV\rho)/\mu$. This one is called the Reynold's number. It is the one which involves viscosity. Thus we would look for

$$l = f(Re, AR, \theta, \mathcal{M})\, kg \times m/\sec^2$$

This is quite interesting because it is easy to vary Re by simply adjusting the velocity or A but it is hard to vary things like μ or ρ. Note that all the quantities are easy to adjust. Now this could be used, along with wind tunnel experiments to get a formula for the lift which would be reasonable. You could also consider more variables and more complicated situations in the same way.

This page titled 2.7: Dimensionless Variables is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.7: Dimensionless Variables** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

2.8: An Application to Resistor Networks

The tools of linear algebra can be used to study the application of resistor networks. An example of an electrical circuit is below.

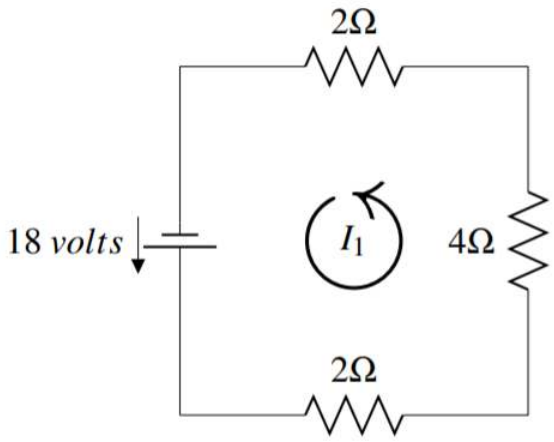


Figure 2.8.1

The jagged lines (—/\/\/\—) denote resistors and the numbers next to them give their resistance in ohms, written as Ω. The voltage source (—|| ||—) causes the current to flow in the direction from the shorter of the two lines toward the longer (as indicated by the arrow). The current for a circuit is labeled I_k.

In the above figure, the current I_1 has been labeled with an arrow in the counter clockwise direction. This is an entirely arbitrary decision and we could have chosen to label the current in the clockwise direction. With our choice of direction here, we define a positive current to flow in the counter clockwise direction and a negative current to flow in the clockwise direction.

The goal of this section is to use the values of resistors and voltage sources in a circuit to determine the current. An essential theorem for this application is Kirchhoff's law.

Theorem 2.8.1: Kirchhoff's Law

The sum of the resistance (R) times the amps (I) in the counter clockwise direction around a loop equals the sum of the voltage sources (V) in the same direction around the loop.

Kirchhoff's law allows us to set up a system of linear equations and solve for any unknown variables. When setting up this system, it is important to trace the circuit in the counter clockwise direction. If a resistor or voltage source is crossed against this direction, the related term must be given a negative sign.

We will explore this in the next example where we determine the value of the current in the initial diagram.

Example 2.8.1: Solving for Current

Applying Kirchhoff's Law to the diagram below, determine the value for I_1.

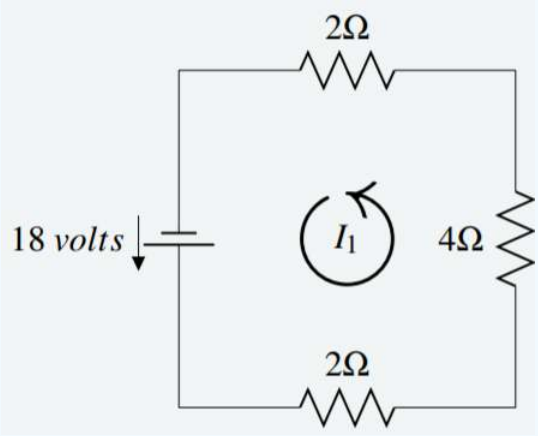


Figure 2.8.2

Solution

Begin in the bottom left corner, and trace the circuit in the counter clockwise direction. At the first resistor, multiplying resistance and current gives $2I_1$. Continuing in this way through all three resistors gives $2I_1+4I_1+2I_1$. This must equal the voltage source in the same direction. Notice that the direction of the voltage source matches the counter clockwise direction specified, so the voltage is positive.

Therefore the equation and solution are given by

$$\begin{aligned} 2I_1+4I_1+2I_1 &= 18 \\ 8I_1 &= 18 \\ I_1 &= \frac{9}{4}A \end{aligned}$$

Since the answer is positive, this confirms that the current flows counter clockwise.

Example 2.8.2 Solving for Current

Applying Kirchhoff's Law to the diagram below, determine the value for I_1.

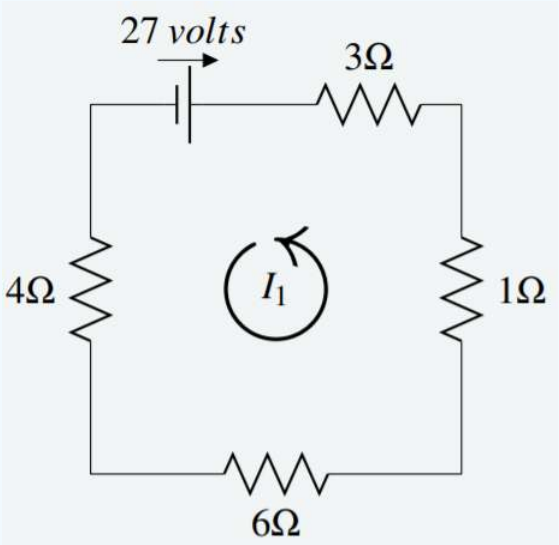


Figure 2.8.3

Solution

Begin in the top left corner this time, and trace the circuit in the counter clockwise direction. At the first resistor, multiplying resistance and current gives $4I_1$. Continuing in this way through the four resistors gives $4I_1+6I_1+1I_1+3I_1$. This must

equal the voltage source in the same direction. Notice that the direction of the voltage source is opposite to the counter clockwise direction, so the voltage is negative.

Therefore the equation and solution are given by

$$\begin{aligned} 4I_1+6I_1+1I_1+3I_1 &= -27 \\ 14I_1 &= -27 \\ I_1 &= -\frac{27}{14}A \end{aligned}$$

Since the answer is negative, this tells us that the current flows clockwise.

A more complicated example follows. Two of the circuits below may be familiar; they were examined in the examples above. However as they are now part of a larger system of circuits, the answers will differ.

Example 2.8.3: Unknown Currents

The diagram below consists of four circuits. The current (I_k) in the four circuits is denoted by I_1, I_2, I_3, I_4. Using Kirchhoff's Law, write an equation for each circuit and solve for each current.

Solution

The circuits are given in the following diagram.

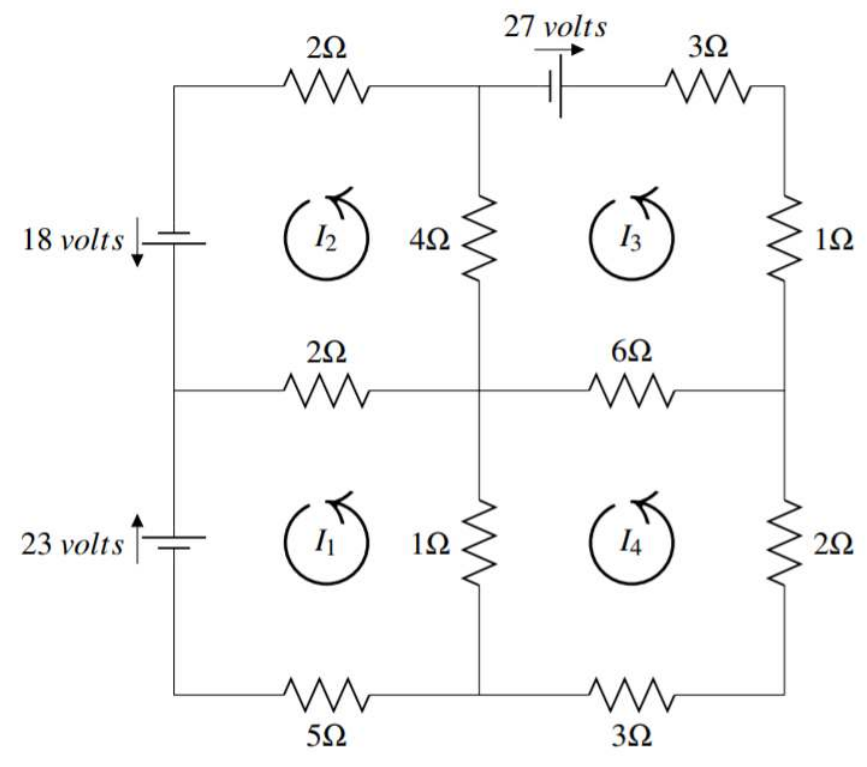


Figure 2.8.4

Starting with the top left circuit, multiply the resistance by the amps and sum the resulting products. Specifically, consider the resistor labeled 2Ω that is part of the circuits of I_1 and I_2. Notice that current I_2 runs through this in a positive (counter clockwise) direction, and I_1 runs through in the opposite (negative) direction. The product of resistance and amps is then $2(I_2-I_1)=2I_2-2I_1$. Continue in this way for each resistor, and set the sum of the products equal to the voltage source to write the equation:

$$2I_2-2I_1+4I_2-4I_3+2I_2=18$$

The above process is used on each of the other three circuits, and the resulting equations are:

Upper right circuit:

$$4I_3 - 4I_2 + 6I_3 - 6I_4 + I_3 + 3I_3 = -27$$

Lower right circuit:

$$3I_4 + 2I_4 + 6I_4 - 6I_3 + I_4 - I_1 = 0$$

Lower left circuit:

$$5I_1 + I_1 - I_4 + 2I_1 - 2I_2 = -23$$

Notice that the voltage for the upper right and lower left circuits are negative due to the clockwise direction they indicate.

The resulting system of four equations in four unknowns is

$$\begin{aligned} 2I_2 - 2I_1 + 4I_2 - 4I_3 + 2I_2 &= 18 \\ 4I_3 - 4I_2 + 6I_3 - 6I_4 + I_3 + I_3 &= -27 \\ 2I_4 + 3I_4 + 6I_4 - 6I_3 + I_4 - I_1 &= 0 \\ 5I_1 + I_1 - I_4 + 2I_1 - 2I_2 &= -23 \end{aligned}$$

Simplifying and rearranging with variables in order, we have:

$$\begin{aligned} -2I_1 + 8I_2 - 4I_3 &= 18 \\ -4I_2 + 14I_3 - 6I_4 &= -27 \\ -I_1 - 6I_3 + 12I_4 &= 0 \\ 8I_1 - 2I_2 - I_4 &= -23 \end{aligned}$$

The augmented matrix is

$$\left[\begin{array}{cccc|c} -2 & 8 & -4 & 0 & 18 \\ 0 & -4 & 14 & -6 & -27 \\ -1 & 0 & -6 & 12 & 0 \\ 8 & -2 & 0 & -1 & -23 \end{array}\right]$$

The solution to this matrix is

$$\begin{aligned} I_1 &= -3A \\ I_2 &= \frac{1}{4}A \\ I_3 &= -\frac{5}{2}A \\ I_4 &= -\frac{3}{2}A \end{aligned}$$

This tells us that currents I_1, I_3, and I_4 travel clockwise while I_2 travels counter clockwise.

This page titled 2.8: An Application to Resistor Networks is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.8: An Application to Resistor Networks** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

2.9: Exercises

Exercises

Exercise 2.9.1

Graphically, find the point (x_1, y_1) which lies on both lines, $x+3y=1$ and $4x-y=3$. That is, graph each line and see where they intersect.

Answer

$\begin{matrix} x+3y=1 \\ 4x-y=3 \end{matrix}$, Solution is: $\left[x=\frac{10}{13}, y=\frac{1}{13}\right]$.

Exercise 2.9.2

Graphically, find the point of intersection of the two lines $3x+y=3$ and $x+2y=1$. That is, graph each line and see where they intersect.

Answer

$\begin{matrix} 3x+y=3 \\ x+2y=1 \end{matrix}$, Solution is: $[x=1, y=0]$

Exercise 2.9.3

Do the three lines, $x+2y=1$, $2x-y=1$, and $4x+3y=3$ have a common point of intersection? If so, find the point and if not, tell why they don't have such a common point of intersection.

Answer

$\begin{matrix} x+2y=1 \\ 2x-y=1 \\ 4x+3y=3 \end{matrix}$, Solution is: $\left[x=\frac{3}{5}, y=\frac{1}{5}\right]$

Exercise 2.9.4

Do the three planes, $x+y-3z=2$, $2x+y+z=1$, and $3x+2y-2z=0$ have a common point of intersection? If so, find one and if not, tell why there is no such point.

Answer

No solution exists. You can see this by writing the augmented matrix and doing row operations. $\begin{bmatrix} 1 & 1 & -3 & 2 \\ 2 & 1 & 1 & 1 \\ 3 & 2 & -2 & 0 \end{bmatrix}$, row echelon form: $\begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & -7 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. Thus one of the equation says $0=1$ in an equivalent system of equations.

Exercise 2.9.5

Four times the weight of Gaston is 150 pounds more than the weight of Ichabod. Four times the weight of Ichabod is 660 pounds less than seventeen times the weight of Gaston. Four times the weight of Gaston plus the weight of Siegfried equals 290 pounds. Brunhilde would balance all three of the others. Find the weights of the four people.

Answer

$$\begin{matrix} 4g-I=150 \\ 4I-17g=-660 \\ 4g+s=290 \\ g+I+s-b=0 \end{matrix}$$, Solution is: $\{g=60, I=90, b=200, s=50\}$

? Exercise 2.9.6

Consider the following augmented matrix in which $*$ denotes an arbitrary number and ■ denotes a nonzero number. Determine whether the given augmented matrix is consistent. If consistent, is the solution unique?

$$\left[\begin{array}{ccccc|c} ■ & * & * & * & * & * \\ 0 & ■ & * & * & 0 & * \\ 0 & 0 & ■ & * & * & * \\ 0 & 0 & 0 & 0 & ■ & * \end{array}\right]$$

Answer

The solution exists but is not unique.

? Exercise 2.9.7

Consider the following augmented matrix in which $*$ denotes an arbitrary number and ■ denotes a nonzero number. Determine whether the given augmented matrix is consistent. If consistent, is the solution unique?

$$\left[\begin{array}{ccc|c} ■ & * & * & * \\ 0 & ■ & * & * \\ 0 & 0 & ■ & * \end{array}\right]$$

Answer

A solution exists and is unique.

? Exercise 2.9.8

Consider the following augmented matrix in which $*$ denotes an arbitrary number and ■ denotes a nonzero number. Determine whether the given augmented matrix is consistent. If consistent, is the solution unique?

$$\left[\begin{array}{ccccc|c} ■ & * & * & * & * & * \\ 0 & ■ & 0 & * & 0 & * \\ 0 & 0 & 0 & ■ & * & * \\ 0 & 0 & 0 & 0 & ■ & * \end{array}\right]$$

? Exercise 2.9.9

Consider the following augmented matrix in which $*$ denotes an arbitrary number and ■ denotes a nonzero number. Determine whether the given augmented matrix is consistent. If consistent, is the solution unique?

$$\left[\begin{array}{ccccc|c} ■ & * & * & * & * & * \\ 0 & ■ & * & * & 0 & * \\ 0 & 0 & 0 & 0 & ■ & 0 \\ 0 & 0 & 0 & 0 & * & ■ \end{array}\right]$$

Answer

There might be a solution. If so, there are infinitely many.

Exercise 2.9.10

Suppose a system of equations has fewer equations than variables. Will such a system necessarily be consistent? If so, explain why and if not, give an example which is not consistent.

Answer

No. Consider $x+y+z=2$ and $x+y+z=1$.

Exercise 2.9.11

If a system of equations has more equations than variables, can it have a solution? If so, give an example and if not, tell why not.

Answer

These can have a solution. For example, $x+y=1, 2x+2y=2, 3x+3y=3$ even has an infinite set of solutions.

Exercise 2.9.12

Find h such that

$$\left[\begin{array}{cc|c} 2 & h & 4 \\ 3 & 6 & 7 \end{array}\right]$$

is the augmented matrix of an *inconsistent* system.

Answer

$h=4$

Exercise 2.9.13

Find h such that

$$\left[\begin{array}{cc|c} 1 & h & 3 \\ 2 & 4 & 6 \end{array}\right]$$

is the augmented matrix of a *consistent* system.

Answer

Any h will work.

Exercise 2.9.14

Find h such that

$$\left[\begin{array}{cc|c} 1 & 1 & 4 \\ 3 & h & 12 \end{array}\right]$$

is the augmented matrix of a *consistent* system.

Answer

Any h will work.

Exercise 2.9.15

Choose h and k such that the augmented matrix shown has each of the following:

a. one solution
b. no solution
c. infinitely many solutions

$$\left[\begin{array}{cc|c} 1 & h & 2 \\ 2 & 4 & k \end{array}\right]$$

Answer

If $h \neq 2$ there will be a unique solution for any k. If $h = 2$ and $k \neq 4$, there are no solutions. If $h = 2$ and $k = 4$, then there are infinitely many solutions.

Exercise 2.9.16

Choose h and k such that the augmented matrix shown has each of the following:

a. one solution
b. no solution
c. infinitely many solutions

$$\left[\begin{array}{cc|c} 1 & 2 & 2 \\ 2 & h & k \end{array}\right]$$

Answer

If $h \neq 4$, then there is exactly one solution. If $h = 4$ and $k \neq 4$, then there are no solutions. If $h = 4$ and $k = 4$, then there are infinitely many solutions.

Exercise 2.9.17

Determine if the system is consistent. If so, is the solution unique?

$$\begin{aligned} x + 2y + z - w &= 2 \\ x - y + z + w &= 1 \\ 2x + y - z &= 1 \\ 4x + 2y + z &= 5 \end{aligned}$$

Answer

There is no solution. The system is inconsistent. You can see this from the augmented matrix. $\begin{bmatrix} 1 & 2 & 1 & -1 & 2 \\ 1 & -1 & 1 & 1 & 1 \\ 2 & 1 & -1 & 0 & 1 \\ 4 & 2 & 1 & 0 & 5 \end{bmatrix}$,

reduced row-echelon form: $\begin{bmatrix} 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{2}{3} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$.

Exercise 2.9.18

Determine if the system is consistent. If so, is the solution unique?

$$\begin{aligned} x+2y+z-w &= 2 \\ x-y+z+w &= 0 \\ 2x+y-z &= 1 \\ 4x+2y+z &= 3 \end{aligned}$$

Answer

Solution is: $\left[w=\frac{3}{2}y-1, x=\frac{2}{3}-\frac{1}{2}y, z=\frac{1}{3}\right]$

Exercise 2.9.19

Determine which matrices are in reduced row-echelon form.

a. $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 7 \end{bmatrix}$

b. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 5 \\ 0 & 0 & 1 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 3 \end{bmatrix}$

Answer

a. This one is not.
b. This one is.
c. This one is.

Exercise 2.9.20

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 2 & -1 & 3 & -1 \\ 1 & 0 & 2 & 1 \\ 1 & -1 & 1 & -2 \end{bmatrix}$$

Exercise 2.9.21

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 0 & 0 & -1 & -1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & -1 \end{bmatrix}$$

Exercise 2.9.22

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 3 & -6 & -7 & -8 \\ 1 & -2 & -2 & -2 \\ 1 & -2 & -3 & -4 \end{bmatrix}$$

Exercise 2.9.23

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 2 & 4 & 5 & 15 \\ 1 & 2 & 3 & 9 \\ 1 & 2 & 2 & 6 \end{bmatrix}$$

Exercise 2.9.24

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 4 & -1 & 7 & 10 \\ 1 & 0 & 3 & 3 \\ 1 & -1 & -2 & 1 \end{bmatrix}$$

Exercise 2.9.25

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} 3 & 5 & -4 & 2 \\ 1 & 2 & -1 & 1 \\ 1 & 1 & -2 & 0 \end{bmatrix}$$

Exercise 2.9.26

Row reduce the following matrix to obtain the row-echelon form. Then continue to obtain the reduced row-echelon form.

$$\begin{bmatrix} -2 & 3 & -8 & 7 \\ 1 & -2 & 5 & -5 \\ 1 & -3 & 7 & -8 \end{bmatrix}$$

Exercise 2.9.27

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 1 & 3 & 4 & 2 \\ 1 & 0 & 2 & 1 \end{array}\right]$$

Exercise 2.9.28

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 2 & 0 & 1 & 1 \\ 3 & 2 & 1 & 3 \end{array}\right]$$

Answer

The reduced row-echelon form is $\left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & -\frac{1}{4} & \frac{3}{4} \\ 0 & 0 & 0 & 0 \end{array}\right]$. Therefore, the solution is of the form $z = t, y = \frac{3}{4} + t\left(\frac{1}{4}\right), x = \frac{1}{2} - \frac{1}{2}t$ where $t \in \mathbb{R}$.

Exercise 2.9.29

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 1 & 0 & 4 & 2 \end{array}\right]$$

Answer

The reduced row-echelon form is $\left[\begin{array}{ccc|c} 1 & 0 & 4 & 2 \\ 0 & 1 & -4 & -1 \end{array}\right]$ and so the solution is $z = t, y = 4t, x = 2 - 4t$.

Exercise 2.9.30

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccccc|c} 1 & 0 & 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 1 & 2 & 0 & 0 & 1 & 3 \\ 1 & 0 & 1 & 0 & 2 & 2 \end{array}\right]$$

Answer

The reduced row-echelon form is $\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 9 & 3 \\ 0 & 1 & 0 & 0 & -4 & 0 \\ 0 & 0 & 1 & 0 & -7 & -1 \\ 0 & 0 & 0 & 1 & 6 & 1 \end{array}\right]$ and so $x_5 = t, x_4 = 1 - 6t, x_3 = -1 + 7t, x_2 = 4t, x_1 = 3 - 9t$.

Exercise 2.9.31

Find the solution of the system whose augmented matrix is

$$\left[\begin{array}{ccccc|c} 1 & 0 & 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 0 & 2 & 0 & 0 & 1 & 3 \\ 1 & -1 & 2 & 2 & 2 & 0 \end{array}\right]$$

Answer

The reduced row-echelon form is $\left[\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & -\frac{1}{2} & \frac{5}{2} \\ 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{3}{2} \\ 0 & 0 & 0 & 1 & \frac{3}{2} & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$. Therefore, let $x_5 = t, x_3 = s$. Then the other variables are given by $x_4 = -\frac{1}{2} - \frac{3}{2}t, x_2 = \frac{3}{2} - t\frac{1}{2}, , x_1 = \frac{5}{2} + \frac{1}{2}t - 2s$.

Exercise 2.9.32

Find the solution to the system of equations, $7x + 14y + 15z = 22$, $2x + 4y + 3z = 5$, and $3x + 6y + 10z = 13$.

Answer

Solution is: $[x = 1 - 2t, z = 1, y = t]$

Exercise 2.9.33

Find the solution to the system of equations, $3x - y + 4z = 6,\ y + 8z = 0,$ and $-2x + y = -4.$

Answer

Solution is: $[x = 2 - 4t, y = -8t, z = t]$

Exercise 2.9.34

Find the solution to the system of equations, $9x - 2y + 4z = -17,\ 13x - 3y + 6z = -25,$ and $-2x - z = 3.$

Answer

Solution is: $[x = -1, y = 2, z = -1]$

Exercise 2.9.35

Find the solution to the system of equations, $65x + 84y + 16z = 546,\ 81x + 105y + 20z = 682,$ and $84x + 110y + 21z = 713.$

Answer

Solution is: $[x = 2, y = 4, z = 5]$

Exercise 2.9.36

Find the solution to the system of equations, $8x + 2y + 3z = -3, 8x + 3y + 3z = -1,$ and $4x + y + 3z = -9.$

Answer

Solution is: $[x = 1, y = 2, z = -5]$

Exercise 2.9.37

Find the solution to the system of equations, $-8x + 2y + 5z = 18, -8x + 3y + 5z = 13,$ and $-4x + y + 5z = 19.$

Answer

Solution is: $[x = -1, y = -5, z = 4]$

Exercise 2.9.38

Find the solution to the system of equations, $3x - y - 2z = 3,\ y - 4z = 0,$ and $-2x + y = -2.$

Answer

Solution is: $[x = 2t + 1, y = 4t, z = t]$

Exercise 2.9.39

Find the solution to the system of equations, $-9x + 15y = 66, -11x + 18y = 79, -x + y = 4$, and $z = 3$.

Answer

Solution is: $[x=1, y=5, z=3]$

Exercise 2.9.40

Find the solution to the system of equations, $-19x+8y=-108$, $-71x+30y=-404$, $-2x+y=-12$, $4x+z=14$.

Answer

Solution is: $[x=4, y=-4, z=-2]$

Exercise 2.9.41

Suppose a system of equations has fewer equations than variables and you have found a solution to this system of equations. Is it possible that your solution is the only one? Explain.

Answer

No. Consider $x+y+z=2$ and $x+y+z=1$.

Exercise 2.9.42

Suppose a system of linear equations has a 2×4 augmented matrix and the last column is a pivot column. Could the system of linear equations be consistent? Explain.

Answer

No. This would lead to $0=1$.

Exercise 2.9.43

Suppose the coefficient matrix of a system of n equations with n variables has the property that every column is a pivot column. Does it follow that the system of equations must have a solution? If so, must the solution be unique? Explain.

Answer

Yes. It has a unique solution.

Exercise 2.9.44

Suppose there is a unique solution to a system of linear equations. What must be true of the pivot columns in the augmented matrix?

Answer

The last column must not be a pivot column. The remaining columns must each be pivot columns.

Exercise 2.9.45

The steady state temperature, u, of a plate solves Laplace's equation, $\Delta u=0$. One way to approximate the solution is to divide the plate into a square mesh and require the temperature at each node to equal the average of the temperature at the four adjacent nodes. In the following picture, the numbers represent the observed temperature at the indicated nodes. Find the temperature at the interior nodes, indicated by $x, y, z,$ and w. One of the equations is $z=0.05in\frac{1}{4}(10+0+w+x)$.

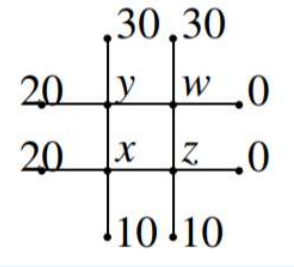


Figure 2.9.1

Answer

You need $\begin{matrix} \frac{1}{4}(20+30+w+x)-y=0 \\ \frac{1}{4}(y+30+0+z)-w=0 \\ \frac{1}{4}(20+y+z+10)-x=0 \\ \frac{1}{4}(x+w+0+10)-z=0 \end{matrix}$, Solution is: $[w=15, x=15, y=20, z=10]$.

Exercise 2.9.46

Find the rank of the following matrix.

$$\begin{bmatrix} 4 & -16 & -1 & -5 \\ 1 & -4 & 0 & -1 \\ 1 & -4 & -1 & -2 \end{bmatrix}$$

Exercise 2.9.47

Find the rank of the following matrix.

$$\begin{bmatrix} 3 & 6 & 5 & 12 \\ 1 & 2 & 2 & 5 \\ 1 & 2 & 1 & 2 \end{bmatrix}$$

Exercise 2.9.48

Find the rank of the following matrix.

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 3 \\ 1 & 4 & 1 & 0 & -8 \\ 1 & 4 & 0 & 1 & 2 \\ -1 & -4 & 0 & -1 & -2 \end{bmatrix}$$

Exercise 2.9.49

Find the rank of the following matrix.

$$\begin{bmatrix} 4 & -4 & 3 & -9 \\ 1 & -1 & 1 & -2 \\ 1 & -1 & 0 & -3 \end{bmatrix}$$

Exercise 2.9.50

Find the rank of the following matrix.

$$\begin{bmatrix} 2 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 7 \\ 1 & 0 & 0 & 1 & 7 \end{bmatrix}$$

Exercise 2.9.51

Find the rank of the following matrix.

$$\begin{bmatrix} 4 & 15 & 29 \\ 1 & 4 & 8 \\ 1 & 3 & 5 \\ 3 & 9 & 15 \end{bmatrix}$$

Exercise 2.9.52

Find the rank of the following matrix.

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 1 \\ 1 & 2 & 3 & -2 & -18 \\ 1 & 2 & 2 & -1 & -11 \\ -1 & -2 & -2 & 1 & 11 \end{bmatrix}$$

Exercise 2.9.53

Find the rank of the following matrix.

$$\begin{bmatrix} 1 & -2 & 0 & 3 & 11 \\ 1 & -2 & 0 & 4 & 15 \\ 1 & -2 & 0 & 3 & 11 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Exercise 2.9.54

Find the rank of the following matrix.

$$\begin{bmatrix} -2 & -3 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \\ -3 & 0 & -3 \end{bmatrix}$$

Exercise 2.9.55

Find the rank of the following matrix.

$$\begin{bmatrix} 4 & 4 & 20 & -1 & 17 \\ 1 & 1 & 5 & 0 & 5 \\ 1 & 1 & 5 & -1 & 2 \\ 3 & 3 & 15 & -3 & 6 \end{bmatrix}$$

Exercise 2.9.56

Find the rank of the following matrix.

$$\begin{bmatrix} -1 & 3 & 4 & -3 & 8 \\ 1 & -3 & -4 & 2 & -5 \\ 1 & -3 & -4 & 1 & -2 \\ -2 & 6 & 8 & -2 & 4 \end{bmatrix}$$

Exercise 2.9.57

Suppose A is an $m \times n$ matrix. Explain why the rank of A is always no larger than $\min(m, n)$.

Answer

It is because you cannot have more than $\min(m, n)$ nonzero rows in the reduced row-echelon form. Recall that the number of pivot columns is the same as the number of nonzero rows from the description of this reduced row-echelon form.

Exercise 2.9.58

State whether each of the following sets of data are possible for the matrix equation $AX = B$. If possible, describe the solution set. That is, tell whether there exists a unique solution, no solution or infinitely many solutions. Here, $[A|B]$ denotes the augmented matrix.

a. A is a 5×6 matrix, $rank(A) = 4$ and $rank[A|B] = 4$.
b. A is a 3×4 matrix, $rank(A) = 3$ and $rank[A|B] = 2$.
c. A is a 4×2 matrix, $rank(A) = 4$ and $rank[A|B] = 4$.
d. A is a 5×5 matrix, $rank(A) = 4$ and $rank[A|B] = 5$.
e. A is a 4×2 matrix, $rank(A) = 2$ and $rank[A|B] = 2$.

Answer

a. This says B is in the span of four of the columns. Thus the columns are not independent. Infinite solution set.
b. This surely can't happen. If you add in another column, the rank does not get smaller.
c. This says B is in the span of the columns and the columns must be independent. You can't have the rank equal 4 if you only have two columns.
d. This says B is not in the span of the columns. In this case, there is no solution to the system of equations represented by the augmented matrix.
e. In this case, there is a unique solution since the columns of A are independent.

Exercise 2.9.59

Consider the system $-5x + 2y - z = 0$ and $-5x - 2y - z = 0$. Both equations equal zero and so $-5x + 2y - z = -5x - 2y - z$ which is equivalent to $y = 0$. Does it follow that x and z can equal anything? Notice that when $x = 1$, $z = -4$, and $y = 0$ are plugged in to the equations, the equations do not equal 0. Why?

Answer

These are not legitimate row operations. They do not preserve the solution set of the system.

Exercise 2.9.60

Balance the following chemical reactions.

a. $KNO_3 + H_2CO_3 \rightarrow K_2CO_3 + HNO_3$
b. $AgI + Na_2S \rightarrow Ag_2S + NaI$

c. $Ba_3N_2 + H_2O \rightarrow Ba(OH)_2 + NH_3$
d. $CaCl_2 + Na_3PO_4 \rightarrow Ca_3(PO_4)_2 + NaCl$

Exercise 2.9.61

In the section on dimensionless variables it was observed that $\rho V^2 AB$ has the units of force. Describe a systematic way to obtain such combinations of the variables which will yield something which has the units of force.

Exercise 2.9.62

Consider the following diagram of four circuits.

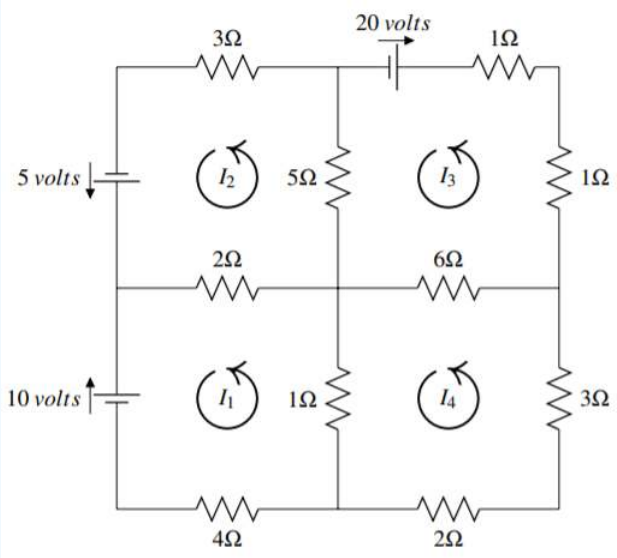


Figure 2.9.2

The current in amps in the four circuits is denoted by I_1, I_2, I_3, I_4 and it is understood that the motion is in the counter clockwise direction. If I_k ends up being negative, then it just means the current flows in the clockwise direction.

In the above diagram, the top left circuit should give the equation

$$2I_2 - 2I_1 + 5I_2 - 5I_3 + 3I_2 = 5$$

For the circuit on the lower left, you should have

$$4I_1 + I_1 - I_4 + 2I_1 - 2I_2 = -10$$

Write equations for each of the other two circuits and then give a solution to the resulting system of equations.

Answer

The other two equations are

$$\begin{aligned} 6I_3 - 6I_4 + I_3 + I_3 + 5I_3 - 5I_2 &= -20 \\ 2I_4 + 3I_4 + 6I_4 - 6I_3 + I_4 - I_1 &= 0 \end{aligned}$$

Then the system is

$$\begin{aligned} 2I_2 - 2I_1 + 5I_2 - 5I_3 + 3I_2 &= 5 \\ 4I_1 + I_1 - I_4 + 2I_1 - 2I_2 &= -10 \\ 6I_3 - 6I_4 + I_3 + I_3 + 5I_3 - 5I_2 &= -20 \\ 2I_4 + 3I_4 + 6I_4 - 6I_3 + I_4 - I_1 &= 0 \end{aligned}$$

The solution is:

$$I_1 = -\frac{750}{373}$$
$$I_2 = -\frac{1421}{1119}$$
$$I_3 = -\frac{3061}{1119}$$
$$I_4 = -\frac{1718}{1119}$$

? Exercise 2.9.63

Consider the following diagram of three circuits.

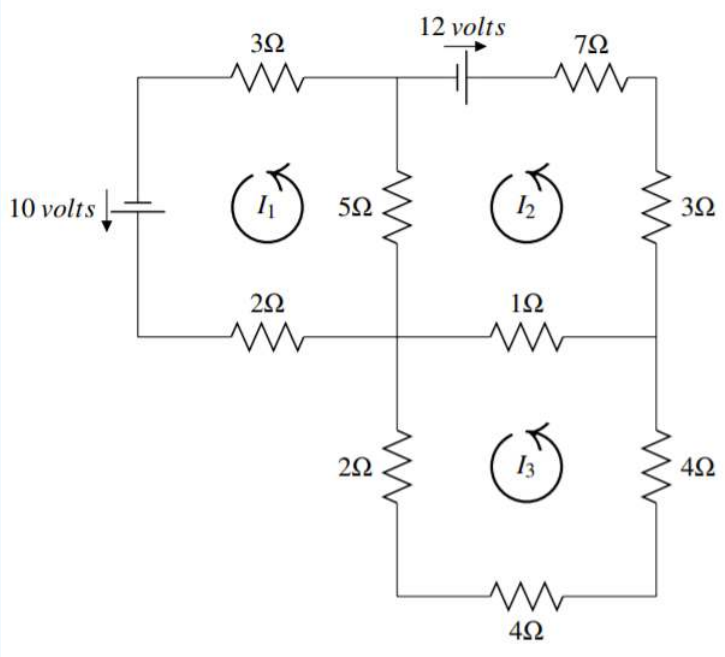


Figure 2.9.3

The current in amps in the four circuits is denoted by I_1, I_2, I_3 and it is understood that the motion is in the counter clockwise direction. If I_k ends up being negative, then it just means the current flows in the clockwise direction.

Find I_1, I_2, I_3.

Answer

You have

$$\begin{aligned} 2I_1 + 5I_1 + 3I_1 - 5I_2 &= 10 \\ I_2 - I_3 + 3I_2 + 7I_2 + 5I_2 - 5I_1 &= -12 \\ 2I_3 + 4I_3 + 4I_3 + I_3 - I_2 &= 0 \end{aligned}$$

Simplifying this yields

$$\begin{aligned} 10I_1 - 5I_2 &= 10 \\ -5I_1 + 16I_2 - I_3 &= -12 \\ -I_2 + 11I_3 &= 0 \end{aligned}$$

The solution is given by

$$I_1 = \frac{218}{295}, I_2 = -\frac{154}{295}, I_3 = -\frac{14}{295}$$

This page titled 2.9: Exercises is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **1.9: Exercises** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

CHAPTER OVERVIEW

3: Determinants

3.1: Basic Techniques
3.2: Properties of Determinants
3.3: Finding Determinants using Row Operations
3.4: Applications of the Determinant
3.E: Exercises

This page titled 3: Determinants is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

3.1: Basic Techniques

Cofactors and 2 x 2 Determinants

Outcomes

A. Evaluate the determinant of a square matrix using either Laplace Expansion or row operations.
B. Demonstrate the effects that row operations have on determinants.
C. Verify the following:
   a. The determinant of a product of matrices is the product of the determinants.
   b. The determinant of a matrix is equal to the determinant of its transpose.

Let A be an $n \times n$ matrix. That is, let A be a square matrix. The **determinant** of A, denoted by $\det(A)$ is a very important number which we will explore throughout this section.

If A is a 2×2 matrix, the determinant is given by the following formula.

Definition 3.1.1: Determinant of a Two By Two Matrix

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then

$$\det(A) = ad - cb$$

The determinant is also often denoted by enclosing the matrix with two vertical lines. Thus

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The following is an example of finding the determinant of a 2×2 matrix.

Example 3.1.1: A Two by Two Determinant

Find $\det(A)$ for the matrix $A = \begin{bmatrix} 2 & 4 \\ -1 & 6 \end{bmatrix}$.

Solution

From Definition 3.1.1,

$$\det(A) = (2)(6) - (-1)(4) = 12 + 4 = 16$$

Below is a video on the determinant of a 2x2 matrix.



The 2×2 determinant can be used to find the determinant of larger matrices. We will now explore how to find the determinant of a 3×3 matrix, using several tools including the 2×2 determinant.

We begin with the following definition.

Definition 3.1.2: The ij^{th} Minor of a Matrix

Let A be a 3×3 matrix. The ij^{th} **minor** of A, denoted as $minor(A)_{ij}$, is the determinant of the 2×2 matrix which results from deleting the ith row and the jth column of A.

In general, if A is an $n \times n$ matrix, then the ij^{th} minor of A is the determinant of the $n-1 \times n-1$ matrix which results from deleting the ith row and the jth column of A.

Hence, there is a minor associated with each entry of A. Consider the following example which demonstrates this definition.

Example 3.1.2: Finding Minors of a Matrix

Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 3 & 2 \\ 3 & 2 & 1 \end{bmatrix}$$

Find $minor(A)_{12}$ and $minor(A)_{23}$.

Solution

First we will find $minor(A)_{12}$. By Definition 3.1.2, this is the determinant of the 2×2 matrix which results when you delete the first row and the second column. This minor is given by

$$minor(A)_{12} = \det \begin{bmatrix} 4 & 2 \\ 3 & 1 \end{bmatrix}$$

Using Definition 3.1.1, we see that

$$\det \begin{bmatrix} 4 & 2 \\ 3 & 1 \end{bmatrix} = (4)(1) - (3)(2) = 4 - 6 = -2$$

Therefore $minor(A)_{12} = -2$.

Similarly, $minor(A)_{23}$ is the determinant of the 2×2 matrix which results when you delete the second row and the third column. This minor is therefore

$$minor(A)_{23} = \det \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} = -4$$

Finding the other minors of A is left as an exercise.

The ij^{th} minor of a matrix A is used in another important definition, given next.

Definition 3.1.3: The ij^{th} Cofactor of a Matrix

Suppose A is an $n \times n$ matrix. The ij^{th} **cofactor**, denoted by $\mathrm{cof}(A)_{ij}$ is defined to be

$$\mathrm{cof}(A)_{ij} = (-1)^{i+j} minor(A)_{ij}$$

It is also convenient to refer to the cofactor of an entry of a matrix as follows. If a_{ij} is the ij^{th} entry of the matrix, then its cofactor is just $\mathrm{cof}(A)_{ij}$.

Example 3.1.3: Finding Cofactors of a Matrix

Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 3 & 2 \\ 3 & 2 & 1 \end{bmatrix}$$

Find $\text{cof}(A)_{12}$ and $\text{cof}(A)_{23}$.

Solution

We will use Definition 3.1.3 to compute these cofactors.

First, we will compute $\text{cof}(A)_{12}$. Therefore, we need to find $minor(A)_{12}$. This is the determinant of the 2×2 matrix which results when you delete the first row and the second column. Thus $minor(A)_{12}$ is given by

$$\det \begin{bmatrix} 4 & 2 \\ 3 & 1 \end{bmatrix} = -2$$

Then,

$$\text{cof}(A)_{12} = (-1)^{1+2} minor(A)_{12} = (-1)^{1+2}(-2) = 2$$

Hence, $\text{cof}(A)_{12} = 2$.

Similarly, we can find $\text{cof}(A)_{23}$. First, find $minor(A)_{23}$, which is the determinant of the 2×2 matrix which results when you delete the second row and the third column. This minor is therefore

$$\det \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} = -4$$

Hence,

$$\text{cof}(A)_{23} = (-1)^{2+3} minor(A)_{23} = (-1)^{2+3}(-4) = 4$$

Below is a video on minors and cofactors of a 2x2 matrix.



You may wish to find the remaining cofactors for the above matrix. Remember that there is a cofactor for every entry in the matrix.

We have now established the tools we need to find the determinant of a 3×3 matrix.

Definition 3.1.4: The Determinant of a Three By Three Matrix

Let A be a 3×3 matrix. Then, $\det(A)$ is calculated by picking a row (or column) and taking the product of each entry in that row (column) with its cofactor and adding these products together.

This process when applied to the ith row (column) is known as **expanding along the ith row (column)** as is given by

$$\det(A) = a_{i1}\text{cof}(A)_{i1} + a_{i2}\text{cof}(A)_{i2} + a_{i3}\text{cof}(A)_{i3}$$

When calculating the determinant, you can choose to expand any row or any column. Regardless of your choice, you will always get the same number which is the determinant of the matrix A. This method of evaluating a determinant by expanding along a row or a column is called **Laplace Expansion** or **Cofactor Expansion**.

Consider the following example.

Example 3.1.4: Finding the Determinant of a Three by Three Matrix

Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 3 & 2 \\ 3 & 2 & 1 \end{bmatrix}$$

Find $\det(A)$ using the method of Laplace Expansion.

Solution

First, we will calculate $\det(A)$ by expanding along the first column. Using Definition 3.1.4, we take the 1 in the first column and multiply it by its cofactor,

$$1(-1)^{1+1}\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} = (1)(1)(-1) = -1$$

Similarly, we take the 4 in the first column and multiply it by its cofactor, as well as with the 3 in the first column. Finally, we add these numbers together, as given in the following equation.

$$\det(A) = \overbrace{1(-1)^{1+1}\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix}}^{\text{cof}(A)_{11}} + \overbrace{4(-1)^{2+1}\begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix}}^{\text{cof}(A)_{21}} + \overbrace{3(-1)^{3+1}\begin{vmatrix} 2 & 3 \\ 3 & 2 \end{vmatrix}}^{\text{cof}(A)_{31}}$$

Calculating each of these, we obtain

$$\det(A) = 1(1)(-1) + 4(-1)(-4) + 3(1)(-5) = -1 + 16 + -15 = 0$$

Hence, $\det(A) = 0$.

As mentioned in Definition 3.1.4, we can choose to expand along any row or column. Let's try now by expanding along the second row. Here, we take the 4 in the second row and multiply it to its cofactor, then add this to the 3 in the second row multiplied by its cofactor, and the 2 in the second row multiplied by its cofactor. The calculation is as follows.

$$\det(A) = \overbrace{4(-1)^{2+1}\begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix}}^{\text{cof}(A)_{21}} + \overbrace{3(-1)^{2+2}\begin{vmatrix} 1 & 3 \\ 3 & 1 \end{vmatrix}}^{\text{cof}(A)_{22}} + \overbrace{2(-1)^{2+3}\begin{vmatrix} 1 & 2 \\ 3 & 2 \end{vmatrix}}^{\text{cof}(A)_{23}}$$

Calculating each of these products, we obtain

$$\det(A) = 4(-1)(-2) + 3(1)(-8) + 2(-1)(-4) = 0$$

You can see that for both methods, we obtained $\det(A) = 0$.

Below is a video on minors and cofactors of a 3x3 matrix.



As mentioned above, we will always come up with the same value for $\det(A)$ regardless of the row or column we choose to expand along. You should try to compute the above determinant by expanding along other rows and columns. This is a good way to check your work, because you should come up with the same number each time!

We present this idea formally in the following theorem.

Theorem 3.1.1: The Determinant is Well Defined

Expanding the $n \times n$ matrix along any row or column always gives the same answer, which is the determinant.

Below is a video on finding the determinant of a 3x3 matrix using the diagonal method.



We have now looked at the determinant of 2×2 and 3×3 matrices. It turns out that the method used to calculate the determinant of a 3×3 matrix can be used to calculate the determinant of any sized matrix. Notice that Definition 3.1.2, Definition 3.1.3 and Definition 3.1.4 can all be applied to a matrix of any size.

Below is a video on the determinant of a 3x3 matrix using cofactor expansion.



For example, the ij^{th} minor of a 4×4 matrix is the determinant of the 3×3 matrix you obtain when you delete the ith row and the jth column. Just as with the 3×3 determinant, we can compute the determinant of a 4×4 matrix by Laplace Expansion, along any row or column

Consider the following example.

Example 3.1.5: Determinant of a Four by Four Matrix

Find $\det(A)$ where

$$A=\begin{bmatrix}1&2&3&4\\5&4&2&3\\1&3&4&5\\3&4&3&2\end{bmatrix}$$

Solution

As in the case of a 3×3 matrix, you can expand this along any row or column. Lets pick the third column. Then, using Laplace Expansion,

$$\det(A)=3(-1)^{1+3}\begin{vmatrix}5&4&3\\1&3&5\\3&4&2\end{vmatrix}+2(-1)^{2+3}\begin{vmatrix}1&2&4\\1&3&5\\3&4&2\end{vmatrix}+$$

$$4(-1)^{3+3}\begin{vmatrix}1&2&4\\5&4&3\\3&4&2\end{vmatrix}+3(-1)^{4+3}\begin{vmatrix}1&2&4\\5&4&3\\1&3&5\end{vmatrix}$$

Now, you can calculate each 3×3 determinant using Laplace Expansion, as we did above. You should complete these as an exercise and verify that $\det(A)=-12$.

The following provides a formal definition for the determinant of an $n\times n$ matrix. You may wish to take a moment and consider the above definitions for 2×2 and 3×3 determinants in context of this definition.

Definition 3.1.5: The Determinant of an $n\times n$ Matrix

Let A be an $n\times n$ matrix where $n\geq 2$ and suppose the determinant of an $(n-1)\times(n-1)$ has been defined. Then

$$\det(A)=\sum_{j=1}^{n}a_{ij}\text{cof}(A)_{ij}=\sum_{i=1}^{n}a_{ij}\text{cof}(A)_{ij}$$

The first formula consists of expanding the determinant along the ith row and the second expands the determinant along the jth column.

In the following sections, we will explore some important properties and characteristics of the determinant.

The Determinant of a Triangular Matrix

There is a certain type of matrix for which finding the determinant is a very simple procedure. Consider the following definition.

Definition 3.1.6: Triangular Matrices

A matrix A is upper triangular if $a_{ij} = 0$ whenever $i > j$. Thus the entries of such a matrix below the main diagonal equal 0, as shown. Here, $*$ refers to any nonzero number.

$$\begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & \vdots \\ \vdots & \vdots & \ddots & * \\ 0 & \cdots & 0 & * \end{bmatrix}$$

A lower triangular matrix is defined similarly as a matrix for which all entries above the main diagonal are equal to zero.

The following theorem provides a useful way to calculate the determinant of a triangular matrix.

Theorem 3.1.2: Determinant of a Triangular Matrix

Let A be an upper or lower triangular matrix. Then $\det(A)$ is obtained by taking the product of the entries on the main diagonal.

The verification of this Theorem can be done by computing the determinant using Laplace Expansion along the first row or column.

Consider the following example.

Example 3.1.6: Determinant of a Triangular Matrix

Let

$$A = \begin{bmatrix} 1 & 2 & 3 & 77 \\ 0 & 2 & 6 & 7 \\ 0 & 0 & 3 & 33.7 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Find $\det(A)$.

Solution

From Theorem 3.1.2, it suffices to take the product of the elements on the main diagonal. Thus $\det(A) = 1 \times 2 \times 3 \times (-1) = -6$.

Without using Theorem 3.1.2, you could use Laplace Expansion. We will expand along the first column. This gives

$$\det(A) = 1\begin{vmatrix} 2 & 6 & 7 \\ 0 & 3 & 33.7 \\ 0 & 0 & -1 \end{vmatrix} + 0(-1)^{2+1}\begin{vmatrix} 2 & 3 & 77 \\ 0 & 3 & 33.7 \\ 0 & 0 & -1 \end{vmatrix} + 0(-1)^{3+1}\begin{vmatrix} 2 & 3 & 77 \\ 2 & 6 & 7 \\ 0 & 0 & -1 \end{vmatrix} + 0(-1)^{4+1}\begin{vmatrix} 2 & 3 & 77 \\ 2 & 6 & 7 \\ 0 & 3 & 33.7 \end{vmatrix}$$

and the only nonzero term in the expansion is

$$1\begin{vmatrix} 2 & 6 & 7 \\ 0 & 3 & 33.7 \\ 0 & 0 & -1 \end{vmatrix}$$

Now find the determinant of this 3×3 matrix, by expanding along the first column to obtain

$$\det(A) = 1\times\left(2\times\begin{vmatrix} 3 & 33.7 \\ 0 & -1 \end{vmatrix} + 0(-1)^{2+1}\begin{vmatrix} 6 & 7 \\ 0 & -1 \end{vmatrix} + 0(-1)^{3+1}\begin{vmatrix} 6 & 7 \\ 3 & 33.7 \end{vmatrix}\right)$$

$$= 1\times 2\times\begin{vmatrix} 3 & 33.7 \\ 0 & -1 \end{vmatrix}$$

Next use Definition 3.1.1 to find the determinant of this 2×2 matrix, which is just $3\times -1 - 0\times 33.7 = -3$. Putting all these steps together, we have

$$\det(A) = 1\times 2\times 3\times(-1) = -6$$

which is just the product of the entries down the main diagonal of the original matrix!

You can see that while both methods result in the same answer, Theorem 3.1.2 provides a much quicker method.

Below is a video on finding the determinant of a matrix using the TI 83/84.



In the next section, we explore some important properties of determinants.

This page titled 3.1: Basic Techniques is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **3.1: Basic Techniques** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

3.2: Properties of Determinants

Properties of Determinants I: Examples

There are many important properties of determinants. Since many of these properties involve the row operations discussed in Chapter 1, we recall that definition now.

Definition 3.2.1: Row Operations

The row operations consist of the following

1. Switch two rows.
2. Multiply a row by a nonzero number.
3. Replace a row by a multiple of another row added to itself.

We will now consider the effect of row operations on the determinant of a matrix. In future sections, we will see that using the following properties can greatly assist in finding determinants. This section will use the theorems as motivation to provide various examples of the usefulness of the properties.

The first theorem explains the affect on the determinant of a matrix when two rows are switched.

Theorem 3.2.1: Switching Rows

Let A be an $n \times n$ matrix and let B be a matrix which results from switching two rows of A. Then $\det(B) = -\det(A)$.

When we switch two rows of a matrix, the determinant is multiplied by -1. Consider the following example.

Example 3.2.1: Switching Two Rows

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and let $B = \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}$. Knowing that $\det(A) = -2$, find $\det(B)$.

Solution

By Definition 3.1.1, $\det(A) = 1 \times 4 - 3 \times 2 = -2$. Notice that the rows of B are the rows of A but switched. By Theorem 3.2.1 since two rows of A have been switched, $\det(B) = -\det(A) = -(-2) = 2$. You can verify this using Definition 3.1.1.

The next theorem demonstrates the effect on the determinant of a matrix when we multiply a row by a scalar.

Theorem 3.2.2: Multiplying a Row by a Scalar

Let A be an $n \times n$ matrix and let B be a matrix which results from multiplying some row of A by a scalar k. Then $\det(B) = k\det(A)$.

Notice that this theorem is true when we multiply *one* row of the matrix by k. If we were to multiply *two* rows of A by k to obtain B, we would have $\det(B) = k^2 \det(A)$. Suppose we were to multiply all n rows of A by k to obtain the matrix B, so that $B = kA$. Then, $\det(B) = k^n \det(A)$. This gives the next theorem.

Theorem 3.2.3: Scalar Multiplication

Let A and B be $n \times n$ matrices and k a scalar, such that $B = kA$. Then $\det(B) = k^n \det(A)$.

Consider the following example.

Example 3.2.2: Multiplying a Row by 5

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 10 \\ 3 & 4 \end{bmatrix}$. Knowing that $\det(A) = -2$, find $\det(B)$.

Solution

By Definition 3.1.1, $\det(A) = -2$. We can also compute $\det(B)$ using Definition 3.1.1, and we see that $\det(B) = -10$.

Now, let's compute $\det(B)$ using Theorem 3.2.2 and see if we obtain the same answer. Notice that the first row of B is 5 times the first row of A, while the second row of B is equal to the second row of A. By Theorem 3.2.2, $\det(B) = 5 \times \det(A) = 5 \times -2 = -10$.

You can see that this matches our answer above.

Finally, consider the next theorem for the last row operation, that of adding a multiple of a row to another row.

Theorem 3.2.4: Adding a Multiple of a Row to Another Row

Let A be an $n \times n$ matrix and let B be a matrix which results from adding a multiple of a row to another row. Then $\det(A) = \det(B)$.

Therefore, when we add a multiple of a row to another row, the determinant of the matrix is unchanged. Note that if a matrix A contains a row which is a multiple of another row, $\det(A)$ will equal 0. To see this, suppose the first row of A is equal to -1 times the second row. By Theorem 3.2.4, we can add the first row to the second row, and the determinant will be unchanged. However, this row operation will result in a row of zeros. Using Laplace Expansion along the row of zeros, we find that the determinant is 0.

Consider the following example.

Example 3.2.3: Adding a Row to Another Row

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and let $B = \begin{bmatrix} 1 & 2 \\ 5 & 8 \end{bmatrix}$. Find $\det(B)$.

Solution

By Definition 3.1.1, $\det(A) = -2$. Notice that the second row of B is two times the first row of A added to the second row. By Theorem 3.2.1, $\det(B) = \det(A) = -2$. As usual, you can verify this answer using Definition 3.1.1.

Example 3.2.4: Multiple of a Row

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$. Show that $\det(A) = 0$.

Solution

Using Definition 3.1.1, the determinant is given by

$$\det(A) = 1 \times 4 - 2 \times 2 = 0$$

However notice that the second row is equal to 2 times the first row. Then by the discussion above following Theorem 3.2.4 the determinant will equal 0.

Until now, our focus has primarily been on row operations. However, we can carry out the same operations with columns, rather than rows. The three operations outlined in Definition 3.2.1 can be done with columns instead of rows. In this case, in Theorems 3.2.1, 3.2.2, and 3.2.4 you can replace the word, "row" with the word "column".

There are several other major properties of determinants which do not involve row (or column) operations. The first is the determinant of a product of matrices.

Theorem 3.2.5: Determinant of a Product

Let A and B be two $n \times n$ matrices. Then

$$\det(AB) = \det(A)\det(B)$$

In order to find the determinant of a product of matrices, we can simply take the product of the determinants.

Consider the following example.

Example 3.2.5: The Determinant of a Product

Compare $\det(AB)$ and $\det(A)\det(B)$ for

$$A = \begin{bmatrix} 1 & 2 \\ -3 & 2 \end{bmatrix}, B = \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix}$$

Solution

First compute AB, which is given by

$$AB = \begin{bmatrix} 1 & 2 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 4 \\ -1 & -4 \end{bmatrix}$$

and so by Definition 3.1.1

$$\det(AB) = \det \begin{bmatrix} 11 & 4 \\ -1 & -4 \end{bmatrix} = -40$$

Now

$$\det(A) = \det \begin{bmatrix} 1 & 2 \\ -3 & 2 \end{bmatrix} = 8$$

and

$$\det(B) = \det \begin{bmatrix} 3 & 2 \\ 4 & 1 \end{bmatrix} = -5$$

Computing $\det(A) \times \det(B)$ we have $8 \times -5 = -40$. This is the same answer as above and you can see that $\det(A)\det(B) = 8 \times (-5) = -40 = \det(AB)$.

Consider the next important property.

Theorem 3.2.6: Determinant of the Transpose

Let A be a matrix where A^T is the transpose of A. Then,

$$\det\left(A^T\right) = \det(A)$$

This theorem is illustrated in the following example.

Example 3.2.6: Determinant of the Transpose

Let

$$A = \begin{bmatrix} 2 & 5 \\ 4 & 3 \end{bmatrix}$$

Find $\det\left(A^T\right)$.

Solution

First, note that

$$A^T = \begin{bmatrix} 2 & 4 \\ 5 & 3 \end{bmatrix}$$

Using Definition 3.1.1, we can compute $\det(A)$ and $\det\left(A^T\right)$. It follows that $\det(A) = 2\times 3 - 4\times 5 = -14$ and $\det\left(A^T\right) = 2\times 3 - 5\times 4 = -14$. Hence, $\det(A) = \det\left(A^T\right)$.

The following provides an essential property of the determinant, as well as a useful way to determine if a matrix is invertible.

Theorem 3.2.7: Determinant of the Inverse

Let A be an $n\times n$ matrix. Then A is invertible if and only if $\det(A) \neq 0$. If this is true, it follows that

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Consider the following example.

Example 3.2.7: Determinant of an Invertible Matrix

Let $A = \begin{bmatrix} 3 & 6 \\ 2 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 3 \\ 5 & 1 \end{bmatrix}$. For each matrix, determine if it is invertible. If so, find the determinant of the inverse.

Solution

Consider the matrix A first. Using Definition 3.1.1 we can find the determinant as follows:

$$\det(A) = 3\times 4 - 2\times 6 = 12 - 12 = 0$$

By Theorem 3.2.7 A is not invertible.

Now consider the matrix B. Again by Definition 3.1.1 we have

$$\det(B) = 2\times 1 - 5\times 3 = 2 - 15 = -13$$

By Theorem 3.2.7 B is invertible and the determinant of the inverse is given by

$$\begin{aligned} \det\left(A^{-1}\right) &= \frac{1}{\det(A)} \\ &= \frac{1}{-13} \\ &= -\frac{1}{13} \end{aligned}$$

Properties of Determinants II: Some Important Proofs

This section includes some important proofs on determinants and cofactors.

First we recall the definition of a determinant. If $A = [a_{ij}]$ is an $n\times n$ matrix, then $\det A$ is defined by computing the expansion along the first row:

$$\det A = \sum_{i=1}^{n} a_{1,i}\mathrm{cof}(A)_{1,i}. \tag{3.2.1}$$

If $n = 1$ then $\det A = a_{1,1}$.

The following example is straightforward and strongly recommended as a means for getting used to definitions.

Example 3.2.8:

(1) Let E_{ij} be the elementary matrix obtained by interchanging ith and jth rows of I. Then $\det E_{ij} = -1$.

(2) Let E_{ik} be the elementary matrix obtained by multiplying the ith row of I by k. Then $\det E_{ik} = k$.

(3) Let E_{ijk} be the elementary matrix obtained by multiplying ith row of I by k and adding it to its jth row. Then $\det E_{ijk} = 1$.

(4) If C and B are such that CB is defined and the ith row of C consists of zeros, then the ith row of CB consists of zeros.

(5) If E is an elementary matrix, then $\det E = \det E^T$.

Many of the proofs in section use the Principle of Mathematical Induction. This concept is discussed in Appendix A.2 and is reviewed here for convenience. First we check that the assertion is true for $n = 2$ (the case $n = 1$ is either completely trivial or meaningless).

Next, we assume that the assertion is true for $n-1$ (where $n \geq 3$) and prove it for n. Once this is accomplished, by the Principle of Mathematical Induction we can conclude that the statement is true for all $n \times n$ matrices for every $n \geq 2$.

If A is an $n \times n$ matrix and $1 \leq j \leq n$, then the matrix obtained by removing 1st column and jth row from A is an $n-1 \times n-1$ matrix (we shall denote this matrix by $A(j)$ below). Since these matrices are used in computation of cofactors $\operatorname{cof}(A)_{1,i}$, for $1 \leq i \neq n$, the inductive assumption applies to these matrices.

Consider the following lemma.

Lemma 3.2.1:

If A is an $n \times n$ matrix such that one of its rows consists of zeros, then $\det A = 0$.

Proof

We will prove this lemma using Mathematical Induction.

If $n = 2$ this is easy (check!).

Let $n \geq 3$ be such that every matrix of size $n-1 \times n-1$ with a row consisting of zeros has determinant equal to zero. Let i be such that the ith row of A consists of zeros. Then we have $a_{ij} = 0$ for $1 \leq j \leq n$.

Fix $j \in \{1, 2, \ldots, n\}$ such that $j \neq i$. Then matrix $A(j)$ used in computation of $\operatorname{cof}(A)_{1,j}$ has a row consisting of zeros, and by our inductive assumption $\operatorname{cof}(A)_{1,j} = 0$.

On the other hand, if $j = i$ then $a_{1,j} = 0$. Therefore $a_{1,j}\operatorname{cof}(A)_{1,j} = 0$ for all j and by (3.2.1) we have

$$\det A = \sum_{j=1}^{n} a_{1,j}\operatorname{cof}(A)_{1,j} = 0$$

as each of the summands is equal to 0.

Lemma 3.2.2:

Assume A, B and C are $n \times n$ matrices that for some $1 \leq i \leq n$ satisfy the following.

1. jth rows of all three matrices are identical, for $j \neq i$.
2. Each entry in the jth row of A is the sum of the corresponding entries in jth rows of B and C.

Then $\det A = \det B + \det C$.

Proof

This is not difficult to check for $n = 2$ (do check it!).

Now assume that the statement of Lemma is true for $n-1 \times n-1$ matrices and fix A, B and C as in the statement. The assumptions state that we have $a_{l,j} = b_{l,j} = c_{l,j}$ for $j \neq i$ and for $1 \leq l \leq n$ and $a_{l,i} = b_{l,i} + c_{l,i}$ for all $1 \leq l \leq n$.

Therefore $A(i) = B(i) = C(i)$, and $A(j)$ has the property that its ith row is the sum of ith rows of $B(j)$ and $C(j)$ for $j \neq i$ while the other rows of all three matrices are identical. Therefore by our inductive assumption we have $\operatorname{cof}(A)_{1j} = \operatorname{cof}(B)_{1j} + \operatorname{cof}(C)_{1j}$ for $j \neq i$.

By (3.2.1) we have (using all equalities established above)

$$\begin{aligned}\det A &= \sum_{l=1}^{n} a_{1,l}\operatorname{cof}(A)_{1,l} \\ &= \sum_{l\neq i} a_{1,l}(\operatorname{cof}(B)_{1,l} + \operatorname{cof}(C)_{1,l}) + (b_{1,i} + c_{1,i})\operatorname{cof}(A)_{1,i} \\ &= \det B + \det C\end{aligned}$$

This proves that the assertion is true for all n and completes the proof.

Theorem 3.2.8:

Let A and B be $n \times n$ matrices.

1. If A is obtained by interchanging ith and jth rows of B (with $i \neq j$), then $\det A = -\det B$.
2. If A is obtained by multiplying ith row of B by k then $\det A = k \det B$.
3. If two rows of A are identical then $\det A = 0$.
4. If A is obtained by multiplying ith row of B by k and adding it to jth row of B ($i \neq j$) then $\det A = \det B$.

Proof

We prove all statements by induction. The case $n = 2$ is easily checked directly (and it is strongly suggested that you do check it).

We assume $n \geq 3$ and (1)–(4) are true for all matrices of size $n-1 \times n-1$.

(1) We prove the case when $j = i+1$, i.e., we are interchanging two consecutive rows.

Let $l \in \{1, \ldots, n\} \setminus \{i, j\}$. Then $A(l)$ is obtained from $B(l)$ by interchanging two of its rows (draw a picture) and by our assumption

$$\operatorname{cof}(A)_{1,l} = -\operatorname{cof}(B)_{1,l}. \tag{3.2.2}$$

Now consider $a_{1,i}\operatorname{cof}(A)_{1,l}$. We have that $a_{1,i} = b_{1,j}$ and also that $A(i) = B(j)$. Since $j = i+1$, we have

$$(-1)^{1+j} = (-1)^{1+i+1} = -(-1)^{1+i}$$

and therefore $a_{1i}\operatorname{cof}(A)_{1i} = -b_{1j}\operatorname{cof}(B)_{1j}$ and $a_{1j}\operatorname{cof}(A)_{1j} = -b_{1i}\operatorname{cof}(B)_{1i}$. Putting this together with (3.2.2) into (3.2.1) we see that if in the formula for $\det A$ we change the sign of each of the summands we obtain the formula for $\det B$.

$$\det A = \sum_{l=1}^{n} a_{1l}\operatorname{cof}(A)_{1l} = -\sum_{l=1}^{n} b_{1l}B_{1l} = \det B.$$

We have therefore proved the case of (1) when $j = i+1$. In order to prove the general case, one needs the following fact. If $i < j$, then in order to interchange ith and jth row one can proceed by interchanging two adjacent rows $2(j-i)+1$ times: First swap ith and $i+1$ st, then $i+1$ st and $i+2$ nd, and so on. After one interchanges $j-1$ st and jth row, we have ith row in position of jth and lth row in position of $l-1$ st for $i+1 \leq l \leq j$. Then proceed backwards swapping adjacent rows until everything is in place.

Since $2(j-i)+1$ is an odd number $(-1)^{2(j-i)+1} = -1$ and we have that $\det A = -\det B$.

(2) This is like (1)… but much easier. Assume that (2) is true for all $n-1 \times n-1$ matrices. We have that $a_{ji} = kb_{ji}$ for $1 \leq j \leq n$. In particular $a_{1i} = kb_{1i}$, and for $l \neq i$ matrix $A(l)$ is obtained from $B(l)$ by multiplying one of its rows by k. Therefore $\operatorname{cof}(A)_{1l} = k\operatorname{cof}(B)_{1l}$ for $l \neq i$, and for all l we have $a_{1l}\operatorname{cof}(A)_{1l} = kb_{1l}\operatorname{cof}(B)_{1l}$. By (3.2.1), we have $\det A = k \det B$.

(3) This is a consequence of (1). If two rows of A are identical, then A is equal to the matrix obtained by interchanging those two rows and therefore by (1) $\det A = -\det A$. This implies $\det A = 0$.

(4) Assume (4) is true for all $n-1 \times n-1$ matrices and fix A and B such that A is obtained by multiplying ith row of B by k and adding it to jth row of B $(i \neq j)$ then $\det A = \det B$. If $k = 0$ then $A = B$ and there is nothing to prove, so we may assume $k \neq 0$.

Let C be the matrix obtained by replacing the jth row of B by the ith row of B multiplied by k. By Lemma 3.2.2, we have that

$$\det A = \det B + \det C$$

and we 'only' need to show that $\det C = 0$. But ith and jth rows of C are proportional. If D is obtained by multiplying the jth row of C by $\frac{1}{k}$ then by (2) we have $\det C = \frac{1}{k} \det D$ (recall that $k \neq 0$!). But ith and jth rows of D are identical, hence by (3) we have $\det D = 0$ and therefore $\det C = 0$.

Theorem 3.2.9:

Let A and B be two $n \times n$ matrices. Then

$$\det(AB) = \det(A) \det(B)$$

Proof

If A is an elementary matrix of either type, then multiplying by A on the left has the same effect as performing the corresponding elementary row operation. Therefore the equality $\det(AB) = \det A \det B$ in this case follows by Example 3.2.8and Theorem 3.2.8

If C is the reduced row-echelon form of A then we can write $A = E_1 \cdot E_2 \cdot \cdots \cdot E_m \cdot C$ for some elementary matrices $E_1, \ldots, E_m$.

Now we consider two cases.

Assume first that $C = I$. Then $A = E_1 \cdot E_2 \cdot \cdots \cdot E_m$ and $AB = E_1 \cdot E_2 \cdot \cdots \cdot E_m B$. By applying the above equality m times, and then $m-1$ times, we have that

$$\begin{aligned} \det AB &= \det E_1 \det E_2 \cdot \det E_m \cdot \det B \\ &= \det(E_1 \cdot E_2 \cdot \cdots \cdot E_m) \det B \\ &= \det A \det B. \end{aligned}$$

Now assume $C \neq I$. Since it is in reduced row-echelon form, its last row consists of zeros and by (4) of Example 3.2.8the last row of CB consists of zeros. By Lemma 3.2.1we have $\det C = \det(CB) = 0$ and therefore

$$\det A = \det(E_1 \cdot E_2 \cdot E_m) \cdot \det(C) = \det(E_1 \cdot E_2 \cdot E_m) \cdot 0 = 0$$

and also

$$\det AB = \det(E_1 \cdot E_2 \cdot E_m) \cdot \det(CB) = \det(E_1 \cdot E_2 \cdot \cdots \cdot E_m) 0 = 0$$

hence $\det AB = 0 = \det A \det B$.

The same 'machine' used in the previous proof will be used again.

Theorem 3.2.10:

Let A be a matrix where A^T is the transpose of A. Then,

$$\det\left(A^T\right) = \det(A)$$

Proof

Note first that the conclusion is true if A is elementary by (5) of Example 3.2.8

Let C be the reduced row-echelon form of A. Then we can write $A = E_1 \cdot E_2 \cdot \cdots \cdot E_m C$. Then $A^T = C^T \cdot E_m^T \cdot \cdots \cdot E_2^T \cdot E_1$. By Theorem 3.2.9we have

$$\det(A^T) = \det(C^T) \cdot \det(E_m^T) \cdot \cdots \cdot \det(E_2^T) \cdot \det(E_1).$$

By (5) of Example 3.2.8we have that $\det E_j = \det E_j^T$ for all j. Also, $\det C$ is either 0 or 1 (depending on whether $C = I$ or not) and in either case $\det C = \det C^T$. Therefore $\det A = \det A^T$.

Below is a video on using properties of matrix determinants.



Below is another video on using properties of matrix determinants.



The above discussions allow us to now prove Theorem 3.1.1. It is restated below.

Theorem 3.2.11:

Expanding an $n \times n$ matrix along any row or column always gives the same result, which is the determinant.

Proof

We first show that the determinant can be computed along any row. The case $n = 1$ does not apply and thus let $n \geq 2$.

Let Abe an $n \times n$ matrix and fix $j > 1$. We need to prove that

$$\det A = \sum_{i=1}^{n} a_{j,i} \text{cof}(A)_{j,i}.$$

Let us prove the case when $j = 2$.

Let B be the matrix obtained from A by interchanging its 1st and 2nd rows. Then by Theorem 3.2.8we have

$$\det A = -\det B.$$

Now we have

$$\det B = \sum_{i=1}^{n} b_{1,i}\text{cof}(B)_{1,i}.$$

Since B is obtained by interchanging the 1st and 2nd rows of A we have that $b_{1,i} = a_{2,i}$ for all i and one can see that $minor(B)_{1,i} = minor(A)_{2,i}$.

Further,

$$\text{cof}(B)_{1,i} = (-1)^{1+i} minor B_{1,i} = -(-1)^{2+i} minor(A)_{2,i} = -\text{cof}(A)_{2,i}$$

hence $\det B = -\sum_{i=1}^{n} a_{2,i}\text{cof}(A)_{2,i}$, and therefore $\det A = -\det B = \sum_{i=1}^{n} a_{2,i}\text{cof}(A)_{2,i}$ as desired.

The case when $j > 2$ is very similar; we still have $minor(B)_{1,i} = minor(A)_{j,i}$ but checking that $\det B = -\sum_{i=1}^{n} a_{j,i}\text{cof}(A)_{j,i}$ is slightly more involved.

Now the cofactor expansion along column j of A is equal to the cofactor expansion along row j of A^T, which is by the above result just proved equal to the cofactor expansion along row 1 of A^T, which is equal to the cofactor expansion along column 1 of A. Thus the cofactor cofactor along any column yields the same result.

Finally, since $\det A = \det A^T$ by Theorem 3.2.10 we conclude that the cofactor expansion along row 1 of A is equal to the cofactor expansion along row 1 of A^T, which is equal to the cofactor expansion along column 1 of A. Thus the proof is complete.

This page titled 3.2: Properties of Determinants is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **3.2: Properties of Determinants** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

3.3: Finding Determinants using Row Operations

Theorems 3.2.1, 3.2.2 and 3.2.4 illustrate how row operations affect the determinant of a matrix. In this section, we look at two examples where row operations are used to find the determinant of a large matrix. Recall that when working with large matrices, Laplace Expansion is effective but timely, as there are many steps involved. This section provides useful tools for an alternative method. By first applying row operations, we can obtain a simpler matrix to which we apply Laplace Expansion.

While working through questions such as these, it is useful to record your row operations as you go along. Keep this in mind as you read through the next example.

Example 3.3.1: Finding a Determinant

Find the determinant of the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 1 & 2 & 3 \\ 4 & 5 & 4 & 3 \\ 2 & 2 & -4 & 5 \end{bmatrix}$$

Solution

We will use the properties of determinants outlined above to find $\det(A)$. First, add -5 times the first row to the second row. Then add -4 times the first row to the third row, and -2 times the first row to the fourth row. This yields the matrix

$$B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -9 & -13 & -17 \\ 0 & -3 & -8 & -13 \\ 0 & -2 & -10 & -3 \end{bmatrix}$$

Notice that the only row operation we have done so far is adding a multiple of a row to another row. Therefore, by Theorem 3.2.4, $\det(B) = \det(A)$.

At this stage, you could use Laplace Expansion to find $\det(B)$. However, we will continue with row operations to find an even simpler matrix to work with.

Add -3 times the third row to the second row. By Theorem 3.2.4 this does not change the value of the determinant. Then, multiply the fourth row by -3. This results in the matrix

$$C = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 11 & 22 \\ 0 & -3 & -8 & -13 \\ 0 & 6 & 30 & 9 \end{bmatrix}$$

Here, $\det(C) = -3\det(B)$, which means that $\det(B) = \left(-\frac{1}{3}\right)\det(C)$

Since $\det(A) = \det(B)$, we now have that $\det(A) = \left(-\frac{1}{3}\right)\det(C)$. Again, you could use Laplace Expansion here to find $\det(C)$. However, we will continue with row operations.

Now replace the add 2 times the third row to the fourth row. This does not change the value of the determinant by Theorem 3.2.4. Finally switch the third and second rows. This causes the determinant to be multiplied by -1. Thus $\det(C) = -\det(D)$ where

$$D = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -3 & -8 & -13 \\ 0 & 0 & 11 & 22 \\ 0 & 0 & 14 & -17 \end{bmatrix}$$

Hence, $\det(A) = \left(-\frac{1}{3}\right)\det(C) = \left(\frac{1}{3}\right)\det(D)$

You could do more row operations or you could note that this can be easily expanded along the first column. Then, expand the resulting 3×3 matrix also along the first column. This results in

$$\det(D) = 1(-3)\begin{vmatrix} 11 & 22 \\ 14 & -17 \end{vmatrix} = 1485$$

and so $\det(A) = \left(\frac{1}{3}\right)(1485) = 495$.

You can see that by using row operations, we can simplify a matrix to the point where Laplace Expansion involves only a few steps. In Example 3.3.1, we also could have continued until the matrix was in upper triangular form, and taken the product of the entries on the main diagonal. Whenever computing the determinant, it is useful to consider all the possible methods and tools.

Consider the next example.

Example 3.3.2: Find the Determinant

Find the determinant of the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 1 & -3 & 2 & 1 \\ 2 & 1 & 2 & 5 \\ 3 & -4 & 1 & 2 \end{bmatrix}$$

Solution

Once again, we will simplify the matrix through row operations. Add -1 times the first row to the second row. Next add -2 times the first row to the third and finally take -3 times the first row and add to the fourth row. This yields

$$B = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & -5 & -1 & -1 \\ 0 & -3 & -4 & 1 \\ 0 & -10 & -8 & -4 \end{bmatrix}$$

By Theorem 3.2.4, $\det(A) = \det(B)$.

Remember you can work with the columns also. Take -5 times the fourth column and add to the second column. This yields

$$C = \begin{bmatrix} 1 & -8 & 3 & 2 \\ 0 & 0 & -1 & -1 \\ 0 & -8 & -4 & 1 \\ 0 & 10 & -8 & -4 \end{bmatrix}$$

By Theorem 3.2.4 $\det(A) = \det(C)$.

Now take -1 times the third row and add to the top row. This gives.

$$D = \begin{bmatrix} 1 & 0 & 7 & 1 \\ 0 & 0 & -1 & -1 \\ 0 & -8 & -4 & 1 \\ 0 & 10 & -8 & -4 \end{bmatrix}$$

which by Theorem 3.2.4 has the same determinant as A.

Now, we can find $\det(D)$ by expanding along the first column as follows. You can see that there will be only one non zero term.

$$\det(D) = 1 \det \begin{bmatrix} 0 & -1 & -1 \\ -8 & -4 & 1 \\ 10 & -8 & -4 \end{bmatrix} + 0 + 0 + 0$$

Expanding again along the first column, we have

$$\det(D) = 1\left(0 + 8\det\begin{bmatrix} -1 & -1 \\ -8 & -4 \end{bmatrix} + 10\det\begin{bmatrix} -1 & -1 \\ -4 & 1 \end{bmatrix}\right) = -82$$

Now since $\det(A) = \det(D)$, it follows that $\det(A) = -82$.

Remember that you can verify these answers by using Laplace Expansion on A. Similarly, if you first compute the determinant using Laplace Expansion, you can use the row operation method to verify.

This page titled 3.3: Finding Determinants using Row Operations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **3.3: Finding Determinants using Row Operations** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

3.4: Applications of the Determinant

Outcomes

A. Use determinants to determine whether a matrix has an inverse, and evaluate the inverse using cofactors.
B. Apply Cramer's Rule to solve a 2×2 or a 3×3 linear system.
C. Given data points, find an appropriate interpolating polynomial and use it to estimate points.

A Formula for the Inverse

The determinant of a matrix also provides a way to find the inverse of a matrix. Recall the definition of the inverse of a matrix in Definition 2.6.1. We say that A^{-1}, an $n \times n$ matrix, is the inverse of A, also $n \times n$, if $AA^{-1} = I$ and $A^{-1}A = I$.

We now define a new matrix called the **cofactor matrix** of A. The cofactor matrix of A is the matrix whose ij^{th} entry is the ij^{th} cofactor of A. The formal definition is as follows.

Definition 3.4.1: The Cofactor Matrix

Let $A = [a_{ij}]$ be an $n \times n$ matrix. Then the **cofactor matrix of A**, denoted $\mathrm{cof}(A)$, is defined by $\mathrm{cof}(A) = \left[\mathrm{cof}(A)_{ij}\right]$ where $\mathrm{cof}(A)_{ij}$ is the ij^{th} cofactor of A.

Note that $\mathrm{cof}(A)_{ij}$ denotes the ij^{th} entry of the cofactor matrix.

We will use the cofactor matrix to create a formula for the inverse of A. First, we define the **adjugate** of A to be the transpose of the cofactor matrix. We can also call this matrix the **classical adjoint** of A, and we denote it by $adj(A)$.

In the specific case where A is a 2×2 matrix given by

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then $adj(A)$ is given by

$$adj(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

In general, $adj(A)$ can always be found by taking the transpose of the cofactor matrix of A. The following theorem provides a formula for A^{-1} using the determinant and adjugate of A.

Theorem 3.4.1: The Inverse and the Determinant

Let A be an $n \times n$ matrix. Then

$$A\ adj(A) = adj(A)\ A = \det(A) I$$

Moreover A is invertible if and only if $\det(A) \neq 0$. In this case we have:

$$A^{-1} = \frac{1}{\det(A)} adj(A)$$

Notice that the first formula holds for any $n \times n$ matrix A, and in the case A is invertible we actually have a formula for A^{-1}.

Consider the following example.

Example 3.4.1: Find Inverse Using the Determinant

Find the inverse of the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix}$$

using the formula in Theorem 3.4.1.

Solution

According to Theorem 3.4.1,

$$A^{-1} = \frac{1}{\det(A)} adj(A)$$

First we will find the determinant of this matrix. Using Theorems 3.2.1, 3.2.2, and 3.2.4, we can first simplify the matrix through row operations. First, add -3 times the first row to the second row. Then add -1 times the first row to the third row to obtain

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -6 & -8 \\ 0 & 0 & -2 \end{bmatrix}$$

By Theorem 3.2.4, $\det(A) = \det(B)$. By Theorem 3.1.2, $\det(B) = 1 \times -6 \times -2 = 12$. Hence, $\det(A) = 12$.

Now, we need to find $adj(A)$. To do so, first we will find the cofactor matrix of A. This is given by

$$\text{cof}(A) = \begin{bmatrix} -2 & -2 & 6 \\ 4 & -2 & 0 \\ 2 & 8 & -6 \end{bmatrix}$$

Here, the ij^{th} entry is the ij^{th} cofactor of the original matrix A which you can verify. Therefore, from Theorem 3.4.1, the inverse of A is given by

$$A^{-1} = \frac{1}{12} \begin{bmatrix} -2 & -2 & 6 \\ 4 & -2 & 0 \\ 2 & 8 & -6 \end{bmatrix}^T = \begin{bmatrix} -\frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{2}{3} \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}$$

Remember that we can always verify our answer for A^{-1}. Compute the product AA^{-1} and $A^{-1}A$ and make sure each product is equal to I.

Compute $A^{-1}A$ as follows

$$A^{-1}A = \begin{bmatrix} -\frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ -\frac{1}{6} & -\frac{1}{6} & \frac{2}{3} \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 3 & 0 & 1 \\ 1 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

You can verify that $AA^{-1} = I$ and hence our answer is correct.

We will look at another example of how to use this formula to find A^{-1}.

Example 3.4.2: Find the Inverse From a Formula

Find the inverse of the matrix

$$A = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{6} & \frac{1}{3} & -\frac{1}{2} \\ -\frac{5}{6} & \frac{2}{3} & -\frac{1}{2} \end{bmatrix}$$

using the formula given in Theorem 3.4.1.

Solution

First we need to find $\det(A)$. This step is left as an exercise and you should verify that $\det(A) = \frac{1}{6}$. The inverse is therefore equal to

$$A^{-1} = \frac{1}{(1/6)}\, adj(A) = 6\, adj(A)$$

We continue to calculate as follows. Here we show the 2×2 determinants needed to find the cofactors.

$$A^{-1} = 6 \begin{bmatrix} \begin{vmatrix} \frac{1}{3} & -\frac{1}{2} \\ \frac{2}{3} & -\frac{1}{2} \end{vmatrix} & -\begin{vmatrix} -\frac{1}{6} & -\frac{1}{2} \\ -\frac{5}{6} & -\frac{1}{2} \end{vmatrix} & \begin{vmatrix} -\frac{1}{6} & \frac{1}{3} \\ -\frac{5}{6} & \frac{2}{3} \end{vmatrix} \\ -\begin{vmatrix} 0 & \frac{1}{2} \\ \frac{2}{3} & -\frac{1}{2} \end{vmatrix} & \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{5}{6} & -\frac{1}{2} \end{vmatrix} & -\begin{vmatrix} \frac{1}{2} & 0 \\ -\frac{5}{6} & \frac{2}{3} \end{vmatrix} \\ \begin{vmatrix} 0 & \frac{1}{2} \\ \frac{1}{3} & -\frac{1}{2} \end{vmatrix} & -\begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{6} & -\frac{1}{2} \end{vmatrix} & \begin{vmatrix} \frac{1}{2} & 0 \\ -\frac{1}{6} & \frac{1}{3} \end{vmatrix} \end{bmatrix}^T$$

Expanding all the 2×2 determinants, this yields

$$A^{-1} = 6 \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}^T = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ 1 & -2 & 1 \end{bmatrix}$$

Again, you can always check your work by multiplying $A^{-1}A$ and AA^{-1} and ensuring these products equal I.

$$A^{-1}A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 1 & 1 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{6} & \frac{1}{3} & -\frac{1}{2} \\ -\frac{5}{6} & \frac{2}{3} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This tells us that our calculation for A^{-1} is correct. It is left to the reader to verify that $AA^{-1} = I$.

The verification step is very important, as it is a simple way to check your work! If you multiply $A^{-1}A$ and AA^{-1} and these products are not both equal to I, be sure to go back and double check each step. One common error is to forget to take the transpose of the cofactor matrix, so be sure to complete this step.

We will now prove Theorem 3.4.1.

Theorem 3.4.1: The Inverse and the Determinant

Proof

(of Theorem 3.4.1) Recall that the (i, j)-entry of $adj(A)$ is equal to $\text{cof}(A)_{ji}$. Thus the (i, j)-entry of $B = A \cdot adj(A)$ is :

$$B_{ij} = \sum_{k=1}^{n} a_{ik} adj(A)_{kj} = \sum_{k=1}^{n} a_{ik} \text{cof}(A)_{jk}$$

By the cofactor expansion theorem, we see that this expression for B_{ij} is equal to the determinant of the matrix obtained from A by replacing its jth row by $a_{i1}, a_{i2}, \ldots a_{in}$ — i.e., its ith row.

If $i = j$ then this matrix is A itself and therefore $B_{ii} = \det A$. If on the other hand $i \neq j$, then this matrix has its ith row equal to its jth row, and therefore $B_{ij} = 0$ in his case. Thus we obtain:

$$A\, adj(A) = \det(A) I$$

Similarly we can verify that:

$$adj(A)\, A = \det(A) I$$

And this proves the first part of the theorem.

Further if A is invertible, then by Theorem 3.2.5 we have:

$$1 = \det(I) = \det\left(AA^{-1}\right) = \det(A)\det\left(A^{-1}\right)$$

and thus $\det(A) \neq 0$. Equivalently, if $\det(A) = 0$, then A is not invertible.

Finally if $\det(A) \neq 0$, then the above formula shows that A is invertible and that:

$$A^{-1} = \frac{1}{\det(A)} adj(A)$$

This completes the proof.

This method for finding the inverse of A is useful in many contexts. In particular, it is useful with complicated matrices where the entries are functions, rather than numbers.

Consider the following example.

Example 3.4.3: Inverse for Non-Constant Matrix

Suppose

$$A(t) = \begin{bmatrix} e^t & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{bmatrix}$$

Show that $A(t)^{-1}$ exists and then find it.

Solution

First note $\det(A(t)) = e^t(\cos^2 t + \sin^2 t) = e^t \neq 0$ so $A(t)^{-1}$ exists.

The cofactor matrix is

$$C(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^t \cos t & e^t \sin t \\ 0 & -e^t \sin t & e^t \cos t \end{bmatrix}$$

and so the inverse is

$$\frac{1}{e^t}\begin{bmatrix} 1 & 0 & 0 \\ 0 & e^t \cos t & e^t \sin t \\ 0 & -e^t \sin t & e^t \cos t \end{bmatrix}^T = \begin{bmatrix} e^{-t} & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{bmatrix}$$

Cramer's Rule

Another context in which the formula given in Theorem 3.4.1 is important is **Cramer's Rule**. Recall that we can represent a system of linear equations in the form $AX = B$, where the solutions to this system are given by X. Cramer's Rule gives a formula for the solutions X in the special case that A is a square invertible matrix. Note this rule does not apply if you have a system of equations in which there is a different number of equations than variables (in other words, when A is not square), or when A is not invertible.

Suppose we have a system of equations given by $AX = B$, and we want to find solutions X which satisfy this system. Then recall that if A^{-1} exists,

$$\begin{aligned} AX &= B \\ A^{-1}(AX) &= A^{-1}B \\ \left(A^{-1}A\right)X &= A^{-1}B \\ IX &= A^{-1}B \\ X &= A^{-1}B \end{aligned}$$

Hence, the solutions X to the system are given by $X = A^{-1}B$. Since we assume that A^{-1} exists, we can use the formula for A^{-1} given above. Substituting this formula into the equation for X, we have

$$X = A^{-1}B = \frac{1}{\det(A)} adj(A)\, B$$

Let x_i be the ith entry of X and b_j be the jth entry of B. Then this equation becomes

$$x_i = \sum_{j=1}^{n} [a_{ij}]^{-1} b_j = \sum_{j=1}^{n} \frac{1}{\det(A)} adj(A)_{ij} b_j$$

where $adj(A)_{ij}$ is the ij^{th} entry of $adj(A)$.

By the formula for the expansion of a determinant along a column,

$$x_i = \frac{1}{\det(A)} \det \begin{bmatrix} * & \cdots & b_1 & \cdots & * \\ \vdots & & \vdots & & \vdots \\ * & \cdots & b_n & \cdots & * \end{bmatrix}$$

where here the ith column of A is replaced with the column vector $[b_1 \cdots, b_n]^T$. The determinant of this modified matrix is taken and divided by $\det(A)$. This formula is known as Cramer's rule.

We formally define this method now.

Procedure 3.4.1: Using Cramer's Rule

Suppose A is an $n \times n$ invertible matrix and we wish to solve the system $AX = B$ for $X = [x_1, \cdots, x_n]^T$. Then Cramer's rule says

$$x_i = \frac{\det(A_i)}{\det(A)}$$

where A_i is the matrix obtained by replacing the ith column of A with the column matrix

$$B = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Below is a video on using Cramer's Rule to solve a system of equations.



We illustrate this procedure in the following example.

Example 3.4.4: Using Cramer's Rule

Find x, y, z if

$$\begin{bmatrix} 1 & 2 & 1 \\ 3 & 2 & 1 \\ 2 & -3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Solution

We will use method outlined in Procedure 3.4.1 to find the values for x, y, z which give the solution to this system. Let

$$B = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

In order to find x, we calculate

$$x = \frac{\det(A_1)}{\det(A)}$$

where A_1 is the matrix obtained from replacing the first column of A with B.

Hence, A_1 is given by

$$A_1 = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 2 & 1 \\ 3 & -3 & 2 \end{bmatrix}$$

Therefore,

$$x = \frac{\det(A_1)}{\det(A)} = \frac{\begin{vmatrix} 1 & 2 & 1 \\ 2 & 2 & 1 \\ 3 & -3 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 2 & 1 \\ 3 & 2 & 1 \\ 2 & -3 & 2 \end{vmatrix}} = \frac{1}{2}$$

Similarly, to find y we construct A_2 by replacing the second column of A with B. Hence, A_2 is given by

$$A_2 = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 3 & 2 \end{bmatrix}$$

Therefore,

$$y = \frac{\det(A_2)}{\det(A)} = \frac{\begin{vmatrix} 1 & 1 & 1 \\ 3 & 2 & 1 \\ 2 & 3 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 2 & 1 \\ 3 & 2 & 1 \\ 2 & -3 & 2 \end{vmatrix}} = -\frac{1}{7}$$

Similarly, A_3 is constructed by replacing the third column of A with B. Then, A_3 is given by

$$A_3 = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 2 & 2 \\ 2 & -3 & 3 \end{bmatrix}$$

Therefore, z is calculated as follows.

$$z = \frac{\det(A_3)}{\det(A)} = \frac{\begin{vmatrix} 1 & 2 & 1 \\ 3 & 2 & 2 \\ 2 & -3 & 3 \end{vmatrix}}{\begin{vmatrix} 1 & 2 & 1 \\ 3 & 2 & 1 \\ 2 & -3 & 2 \end{vmatrix}} = \frac{11}{14}$$

Cramer's Rule gives you another tool to consider when solving a system of linear equations.

We can also use Cramer's Rule for systems of non linear equations. Consider the following system where the matrix A has functions rather than numbers for entries.

Using Cramer's Rule

Example 3.4.5: Use Cramer's Rule for Non-Constant Matrix

Solve for z if

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & e^t \cos t & e^t \sin t \\ 0 & -e^t \sin t & e^t \cos t \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ t \\ 0.05int^2 \end{bmatrix}$$

Solution

We are asked to find the value of z in the solution. We will solve using Cramer's rule. Thus

$$z = .05in \frac{\begin{vmatrix} 1 & 0 & 1 \\ 0 & e^t \cos t & t \\ 0 & -e^t \sin t & t^2 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & 0 \\ 0 & e^t \cos t & e^t \sin t \\ 0 & -e^t \sin t & e^t \cos t \end{vmatrix}} = t\left((\cos t)\, t + \sin t\right) e^{-t}$$

Polynomial Interpolation

In studying a set of data that relates variables x and y, it may be the case that we can use a polynomial to "fit" to the data. If such a polynomial can be established, it can be used to estimate values of x and y which have not been provided.

Consider the following example.

Example 3.4.6: Polynomial Interpolation

Given data points $(1,4),(2,9),(3,12)$ find an interpolating polynomial $p(x)$ of degree at most 2 and then estimate the value corresponding to $x = \frac{1}{2}$.

Solution

We want to find a polynomial given by

$$p(x) = r_0 + r_1 x_1 + r_2 x_2^2$$

such that $p(1) = 4, p(2) = 9$ and $p(3) = 12$. To find this polynomial, substitute the known values in for x and solve for r_0, r_1, and r_2.

$$\begin{aligned} p(1) &= r_0 + r_1 + r_2 = 4 \\ p(2) &= r_0 + 2r_1 + 4r_2 = 9 \\ p(3) &= r_0 + 3r_1 + 9r_2 = 12 \end{aligned}$$

Writing the augmented matrix, we have

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 1 & 2 & 4 & 9 \\ 1 & 3 & 9 & 12 \end{array}\right]$$

After row operations, the resulting matrix is

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & 8 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

Therefore the solution to the system is $r_0 = -3, r_1 = 8, r_2 = -1$ and the required interpolating polynomial is

$$p(x) = -3 + 8x - x^2$$

To estimate the value for $x = \frac{1}{2}$, we calculate $p(\frac{1}{2})$:

$$\begin{aligned} p(\frac{1}{2}) &= -3 + 8(\frac{1}{2}) - (\frac{1}{2})^2 \\ &= -3 + 4 - \frac{1}{4} \\ &= \frac{3}{4} \end{aligned}$$

This procedure can be used for any number of data points, and any degree of polynomial. The steps are outlined below.

Procedure 3.4.2: Finding an Interpolation Polynomial

Suppose that values of x and corresponding values of y are given, such that the actual relationship between x and y is unknown. Then, values of y can be estimated using an **interpolating polynomial $p(x)$**. If given $x_1, \ldots, x_n$ and the corresponding $y_1, \ldots, y_n$, the procedure to find $p(x)$ is as follows:

1. The desired polynomial $p(x)$ is given by

$$p(x) = r_0 + r_1 x + r_2 x^2 + \ldots + r_{n-1} x^{n-1}$$

2. $p(x_i) = y_i$ for all $i = 1, 2, \ldots, n$ so that

$$\begin{aligned} r_0 + r_1 x_1 + r_2 x_1^2 + \ldots + r_{n-1} x_1^{n-1} &= y_1 \\ r_0 + r_1 x_2 + r_2 x_2^2 + \ldots + r_{n-1} x_2^{n-1} &= y_2 \\ &\vdots \\ r_0 + r_1 x_n + r_2 x_n^2 + \ldots + r_{n-1} x_n^{n-1} &= y_n \end{aligned}$$

3. Set up the augmented matrix of this system of equations

$$\left[\begin{array}{ccccc|c} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} & y_1 \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} & y_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} & y_n \end{array}\right]$$

4. Solving this system will result in a unique solution $r_0, r_1, \cdots, r_{n-1}$. Use these values to construct $p(x)$, and estimate the value of $p(a)$ for any $x = a$.

This procedure motivates the following theorem.

Theorem 3.4.2: Polynomial Interpolation

Given n data points $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ with the x_i distinct, there is a unique polynomial $p(x) = r_0 + r_1 x + r_2 x^2 + \cdots + r_{n-1} x^{n-1}$ such that $p(x_i) = y_i$ for $i = 1, 2, \cdots, n$. The resulting polynomial $p(x)$ is called the **interpolating polynomial** for the data points.

We conclude this section with another example.

Example 3.4.7: Polynomial Interpolation

Consider the data points $(0, 1), (1, 2), (3, 22), (5, 66)$. Find an interpolating polynomial $p(x)$ of degree at most three, and estimate the value of $p(2)$.

Solution

The desired polynomial $p(x)$ is given by:

$$p(x) = r_0 + r_1 x + r_2 x^2 + r_3 x^3$$

Using the given points, the system of equations is

$$\begin{aligned} p(0) &= r_0 = 1 \\ p(1) &= r_0 + r_1 + r_2 + r_3 = 2 \\ p(3) &= r_0 + 3r_1 + 9r_2 + 27r_3 = 22 \\ p(5) &= r_0 + 5r_1 + 25r_2 + 125r_3 = 66 \end{aligned}$$

The augmented matrix is given by:

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 3 & 9 & 27 & 22 \\ 1 & 5 & 25 & 125 & 66 \end{array}\right]$$

The resulting matrix is

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

Therefore, $r_0 = 1, r_1 = -2, r_2 = 3, r_3 = 0$ and $p(x) = 1 - 2x + 3x^2$. To estimate the value of $p(2)$, we compute \(p(2) = 1 -2(2) + 3(2

^2) = 1 - 4 + 12 = 9\).

Below is a video on using the determinants to find the area of a triangle.



Below is a video on using the determinants to determine if a matrix is inverable.



This page titled 3.4: Applications of the Determinant is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **3.4: Applications of the Determinant** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

3.E: Exercises

Exercise 3.E.1

Find the determinants of the following matrices.

a. $\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$

b. $\begin{bmatrix} 0 & 3 \\ 0 & 2 \end{bmatrix}$

c. $\begin{bmatrix} 4 & 3 \\ 6 & 2 \end{bmatrix}$

Exercise 3.E.2

Let $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 3 \\ -2 & 5 & 1 \end{bmatrix}$. Find the following.

a. $minor(A)_{11}$

b. $minor(A)_{21}$

c. $minor(A)_{32}$

d. $cof(A)_{11}$

e. $cof(A)_{21}$

f. $cof(A)_{32}$

Exercise 3.E.3

Find the determinants of the following matrices.

a. $\begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 2 \\ 0 & 9 & 8 \end{bmatrix}$

b. $\begin{bmatrix} 4 & 3 & 2 \\ 1 & 7 & 8 \\ 3 & -9 & 3 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 2 & 3 & 2 \\ 1 & 3 & 2 & 3 \\ 4 & 1 & 5 & 0 \\ 1 & 2 & 1 & 2 \end{bmatrix}$

Answer

a. The answer is 31.

b. The answer is 375.

c. The answer is -2.

Exercise 3.E.4

Find the following determinant by expanding along the first row and second column.

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ 2 & 1 & 1 \end{vmatrix}$$

Answer

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ 2 & 1 & 1 \end{vmatrix} = 6$$

? Exercise $3.E.5$

Find the following determinant by expanding along the first column and third row.

$$\begin{vmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{vmatrix}$$

Answer

$$\begin{vmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{vmatrix} = 2$$

? Exercise $3.E.6$

Find the following determinant by expanding along the second row and first column.

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ 2 & 1 & 1 \end{vmatrix}$$

Answer

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \\ 2 & 1 & 1 \end{vmatrix} = 6$$

? Exercise $3.E.7$

Compute the determinant by cofactor expansion. Pick the easiest row or column to use.

$$\begin{vmatrix} 1 & 0 & 0 & 1 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 2 & 1 & 3 & 1 \end{vmatrix}$$

Answer

$$\begin{vmatrix} 1 & 0 & 0 & 1 \\ 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 2 & 1 & 3 & 1 \end{vmatrix} = -4$$

? Exercise $3.E.8$

Find the determinant of the following matrices.

a. $A = \begin{bmatrix} 1 & -34 \\ 0 & 2 \end{bmatrix}$

b. $A = \begin{bmatrix} 4 & 3 & 14 \\ 0 & -2 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

c. $A = \begin{bmatrix} 2 & 3 & 15 & 0 \\ 0 & 4 & 1 & 7 \\ 0 & 0 & -3 & 5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Exercise 3.E.9

An operation is done to get from the first matrix to the second. Identify what was done and tell how it will affect the value of the determinant.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \to \cdots \to \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

Answer

It does not change the determinant. This was just taking the transpose.

Exercise 3.E.10

An operation is done to get from the first matrix to the second. Identify what was done and tell how it will affect the value of the determinant.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \to \cdots \to \begin{bmatrix} c & d \\ a & b \end{bmatrix}$$

Answer

In this case two rows were switched and so the resulting determinant is -1 times the first

Exercise 3.E.11

An operation is done to get from the first matrix to the second. Identify what was done and tell how it will affect the value of the determinant.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \to \cdots \to \begin{bmatrix} a & b \\ a+c & b+d \end{bmatrix}$$

Answer

The determinant is unchanged. It was just the first row added to the second.

Exercise 3.E.12

An operation is done to get from the first matrix to the second. Identify what was done and tell how it will affect the value of the determinant.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \to \cdots \to \begin{bmatrix} a & b \\ 2c & 2d \end{bmatrix}$$

Answer

The second row was multiplied by 2 so the determinant of the result is 2 times the original determinant.

Exercise 3.E.13

An operation is done to get from the first matrix to the second. Identify what was done and tell how it will affect the value of the determinant.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \to \cdots \to \begin{bmatrix} b & a \\ d & c \end{bmatrix}$$

Answer

In this case the two columns were switched so the determinant of the second is -1 times the determinant of the first.

Exercise 3.E.14

Let A be an $r \times r$ matrix and suppose there are $r-1$ rows (columns) such that all rows (columns) are linear combinations of these $r-1$ rows (columns). Show $\det(A) = 0$.

Answer

If the determinant is nonzero, then it will remain nonzero with row operations applied to the matrix. However, by assumption, you can obtain a row of zeros by doing row operations. Thus the determinant must have been zero after all.

Exercise 3.E.15

Show $\det(aA) = a^n \det(A)$ for an $n \times n$ matrix A and scalar a.

Answer

$\det(aA) = \det(aIA) = \det(aI)\det(A) = a^n \det(A)$. The matrix which has a down the main diagonal has determinant equal to a^n.

Exercise 3.E.16

Construct 2×2 matrices A and B to show that the $\det A \det B = \det(AB)$.

Answer

$$\det\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ -5 & 6 \end{bmatrix}\right) = -8$$

$$\det\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \det\begin{bmatrix} -1 & 2 \\ -5 & 6 \end{bmatrix} = -2 \times 4 = -8$$

Exercise 3.E.17

Is it true that $\det(A+B) = \det(A) + \det(B)$? If this is so, explain why. If it is not so, give a counter example.

Answer

This is not true at all. Consider $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$.

Exercise 3.E.18

An $n \times n$ matrix is called **nilpotent** if for some positive integer, k it follows $A^k = 0$. If A is a nilpotent matrix and k is the smallest possible integer such that $A^k = 0$, what are the possible values of $\det(A)$?

Answer

It must be 0 because $0 = \det(0) = \det(A^k) = (\det(A))^k$.

Exercise 3.E.19

A matrix is said to be **orthogonal** if $A^T A = I$. Thus the inverse of an orthogonal matrix is just its transpose. What are the possible values of $\det(A)$ if A is an orthogonal matrix?

Answer

You would need $\det(AA^T) = \det(A)\det(A^T) = \det(A)^2 = 1$ and so $\det(A) = 1$, or -1.

Exercise 3.E.20

Let A and B be two $n \times n$ matrices. $A \sim B$ (A is **similar** to B) means there exists an invertible matrix P such that $A = P^{-1}BP$. Show that if $A \sim B$, then $\det(A) = \det(B)$.

Answer

$\det(A) = \det(S^{-1}BS) = \det(S^{-1})\det(B)\det(S) = \det(B)\det(S^{-1}S) = \det(B)$.

Exercise 3.E.21

Tell whether each statement is true or false. If true, provide a proof. If false, provide a counter example.

a. If A is a 3×3 matrix with a zero determinant, then one column must be a multiple of some other column.
b. If any two columns of a square matrix are equal, then the determinant of the matrix equals zero.
c. For two $n \times n$ matrices A and B, $\det(A+B) = \det(A) + \det(B)$.
d. For an $n \times n$ matrix A, $\det(3A) = 3\det(A)$
e. If A^{-1} exists then $\det(A^{-1}) = \det(A)^{-1}$.
f. If B is obtained by multiplying a single row of A by 4 then $\det(B) = 4\det(A)$.
g. For A an $n \times n$ matrix, $\det(-A) = (-1)^n \det(A)$.
h. If A is a real $n \times n$ matrix, then $\det(A^T A) \geq 0$.
i. If $A^k = 0$ for some positive integer k, then $\det(A) = 0$.
j. If $AX = 0$ for some $X \neq 0$, then $\det(A) = 0$.

Answer

a. False. Consider $\begin{bmatrix} 1 & 1 & 2 \\ -1 & 5 & 4 \\ 0 & 3 & 3 \end{bmatrix}$
b. True.
c. False.
d. False.
e. True.
f. False.
g. True.
h. True.
i. True.
j. True.

Exercise 3.E.22

Find the determinant using row operations to first simplify.

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ -4 & 1 & 2 \end{vmatrix}$$

Answer

$$\begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ -4 & 1 & 2 \end{vmatrix} = -6$$

Exercise 3.E.23

Find the determinant using row operations to first simplify.

$$\begin{vmatrix} 2 & 1 & 3 \\ 2 & 4 & 2 \\ 1 & 4 & -5 \end{vmatrix}$$

Answer

$$\begin{vmatrix} 2 & 1 & 3 \\ 2 & 4 & 2 \\ 1 & 4 & -5 \end{vmatrix} = -32$$

Exercise 3.E.24

Find the determinant using row operations to first simplify.

$$\begin{vmatrix} 1 & 2 & 1 & 2 \\ 3 & 1 & -2 & 3 \\ -1 & 0 & 3 & 1 \\ 2 & 3 & 2 & -2 \end{vmatrix}$$

Answer

One can row reduce this using only row operation 3 to

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & -5 & -5 & -3 \\ 0 & 0 & 2 & \frac{9}{5} \\ 0 & 0 & 0 & -\frac{63}{10} \end{bmatrix}$$

and therefore, the determinant is -63.

$$\begin{vmatrix} 1 & 2 & 1 & 2 \\ 3 & 1 & -2 & 3 \\ -1 & 0 & 3 & 1 \\ 2 & 3 & 2 & -2 \end{vmatrix} = 63$$

Exercise $3.E.25$

Find the determinant using row operations to first simplify.

$$\begin{vmatrix} 1 & 4 & 1 & 2 \\ 3 & 2 & -2 & 3 \\ -1 & 0 & 3 & 3 \\ 2 & 1 & 2 & -2 \end{vmatrix}$$

Answer

One can row reduce this using only row operation 3 to

$$\begin{bmatrix} 1 & 4 & 1 & 2 \\ 0 & -10 & -5 & -3 \\ 0 & 0 & 2 & \frac{19}{5} \\ 0 & 0 & 0 & -\frac{211}{20} \end{bmatrix}$$

Thus the determinant is given by

$$\begin{vmatrix} 1 & 4 & 1 & 2 \\ 3 & 2 & -2 & 3 \\ -1 & 0 & 3 & 3 \\ 2 & 1 & 2 & -2 \end{vmatrix} = 211$$

Exercise $3.E.26$

Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 3 & 1 & 0 \end{bmatrix}$$

Determine whether the matrix A has an inverse by finding whether the determinant is non zero. If the determinant is nonzero, find the inverse using the formula for the inverse which involves the cofactor matrix.

Answer

$\det \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 3 & 1 & 0 \end{bmatrix} = -13$ and so it has an inverse. This inverse is

$$\frac{1}{-13}\begin{bmatrix} \begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} & -\begin{vmatrix} 0 & 1 \\ 3 & 0 \end{vmatrix} & \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} \\ -\begin{vmatrix} 2 & 3 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 3 & 0 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 0 & 2 \end{vmatrix} \end{bmatrix}^T = \frac{1}{-13}\begin{bmatrix} -1 & 3 & -6 \\ 3 & -9 & 5 \\ -4 & -1 & 2 \end{bmatrix}^T$$

$$= \begin{bmatrix} \frac{1}{13} & -\frac{3}{13} & \frac{4}{13} \\ -\frac{3}{13} & \frac{9}{13} & \frac{1}{13} \\ \frac{6}{13} & -\frac{5}{13} & -\frac{2}{13} \end{bmatrix}$$

Exercise 3.E.27

Let

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix}$$

Determine whether the matrix A has an inverse by finding whether the determinant is non zero. If the determinant is nonzero, find the inverse using the formula for the inverse.

Answer

$\det \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 3 & 1 & 1 \end{bmatrix} = 7$ so it has an inverse. This inverse is $\frac{1}{7} \begin{bmatrix} 1 & 3 & -6 \\ -2 & 1 & 5 \\ 2 & -1 & 2 \end{bmatrix}^T = \begin{bmatrix} \frac{1}{7} & -\frac{2}{7} & \frac{2}{7} \\ \frac{3}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{6}{7} & \frac{5}{7} & \frac{2}{7} \end{bmatrix}$

Exercise 3.E.28

Let

$$A = \begin{bmatrix} 1 & 3 & 3 \\ 2 & 4 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

Determine whether the matrix A has an inverse by finding whether the determinant is non zero. If the determinant is nonzero, find the inverse using the formula for the inverse.

Answer

$$\det \begin{bmatrix} 1 & 3 & 3 \\ 2 & 4 & 1 \\ 0 & 1 & 1 \end{bmatrix} = 3$$

so it has an inverse which is

$$\begin{bmatrix} 1 & 0 & -3 \\ -\frac{2}{3} & \frac{1}{3} & \frac{5}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \end{bmatrix}$$

Exercise 3.E.29

Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 2 & 6 & 7 \end{bmatrix}$$

Determine whether the matrix A has an inverse by finding whether the determinant is non zero. If the determinant is nonzero, find the inverse using the formula for the inverse.

Exercise 3.E.30

Let

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix}$$

Determine whether the matrix A has an inverse by finding whether the determinant is non zero. If the determinant is nonzero, find the inverse using the formula for the inverse.

Answer

$$\det \begin{bmatrix} 1 & 0 & 3 \\ 1 & 0 & 1 \\ 3 & 1 & 0 \end{bmatrix} = 2$$

and so it has an inverse. The inverse turns out to equal

$$\begin{bmatrix} -\frac{1}{2} & \frac{3}{2} & 0 \\ \frac{3}{2} & -\frac{9}{2} & 1 \\ \frac{1}{2} & -\frac{1}{2} & 0 \end{bmatrix}$$

Exercise $3.E.31$

For the following matrices, determine if they are invertible. If so, use the formula for the inverse in terms of the cofactor matrix to find each inverse. If the inverse does not exist, explain why.

a. $\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$

b. $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 4 & 1 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \\ 0 & 1 & 2 \end{bmatrix}$

Answer

a. $\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1$

b. $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 4 & 1 & 1 \end{vmatrix} = -15$

c. $\begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \\ 0 & 1 & 2 \end{vmatrix} = 0$

Exercise $3.E.32$

Consider the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos t & -\sin t \\ 0 & \sin t & \cos t \end{bmatrix}$$

Does there exist a value of t for which this matrix fails to have an inverse? Explain.

Answer

No. It has a nonzero determinant for all t

Exercise 3.E.33

Consider the matrix

$$A = \begin{bmatrix} 1 & t & t^2 \\ 0 & 1 & 2t \\ t & 0 & 2 \end{bmatrix}$$

Does there exist a value of t for which this matrix fails to have an inverse? Explain.

Answer

$$\det \begin{bmatrix} 1 & t & t^2 \\ 0 & 1 & 2t \\ t & 0 & 2 \end{bmatrix} = t^3 + 2$$

and so it has no inverse when $t = -\sqrt[3]{2}$

Exercise 3.E.34

Consider the matrix

$$A = \begin{bmatrix} e^t & \cosh t & \sinh t \\ e^t & \sinh t & \cosh t \\ e^t & \cosh t & \sinh t \end{bmatrix}$$

Does there exist a value of t for which this matrix fails to have an inverse? Explain.

Answer

$$\det \begin{bmatrix} e^t & \cosh t & \sinh t \\ e^t & \sinh t & \cosh t \\ e^t & \cosh t & \sinh t \end{bmatrix} = 0$$

and so this matrix fails to have a nonzero determinant at any value of t.

Exercise 3.E.35

Consider the matrix

$$A = \begin{bmatrix} e^t & e^{-t}\cos t & e^{-t}\sin t \\ e^t & -e^{-t}\cos t - e^{-t}\sin t & -e^{-t}\sin t + e^{-t}\cos t \\ e^t & 2e^{-t}\sin t & -2e^{-t}\cos t \end{bmatrix}$$

Does there exist a value of t for which this matrix fails to have an inverse? Explain.

Answer

$$\det \begin{bmatrix} e^t & e^{-t}\cos t & e^{-t}\sin t \\ e^t & -e^{-t}\cos t - e^{-t}\sin t & -e^{-t}\sin t + e^{-t}\cos t \\ e^t & 2e^{-t}\sin t & -2e^{-t}\cos t \end{bmatrix} = 5e^{-t} \neq 0$$

and so this matrix is always invertible.

Exercise 3.E.36

Show that if $\det(A) \neq 0$ for A an $n \times n$ matrix, it follows that if $AX = 0$, then $X = 0$.

Answer

If $\det(A) \neq 0$, then A^{-1} exists and so you could multiply on both sides on the left by A^{-1} and obtain that $X = 0$.

Exercise 3.E.37

Suppose A, B are $n \times n$ matrices and that $AB = I$. Show that then $BA = I$. **Hint**: First explain why $\det(A)$, $\det(B)$ are both nonzero. Then $(AB)A = A$ and then show $BA(BA - I) = 0$. From this use what is given to conclude $A(BA - I) = 0$. Then use Exercise 3.E.36.

Answer

You have $1 = \det(A)\det(B)$. Hence both A and B have inverses. Letting X be given,

$$A(BA - I)X = (AB)AX - AX = AX - AX = 0$$

and so it follows from the above problem that $(BA - I)X = 0$. Since X is arbitrary, it follows that $BA = I$.

Exercise 3.E.38

Use the formula for the inverse in terms of the cofactor matrix to find the inverse of the matrix

$$A = \begin{bmatrix} e^t & 0 & 0 \\ 0 & e^t \cos t & e^t \sin t \\ 0 & e^t \cos t - e^t \sin t & e^t \cos t + e^t \sin t \end{bmatrix}$$

Answer

$$\det \begin{bmatrix} e^t & 0 & 0 \\ 0 & e^t \cos t & e^t \sin t \\ 0 & e^t \cos t - e^t \sin t & e^t \cos t + e^t \sin t \end{bmatrix} = e^{3t}.$$

Hence the inverse is

$$e^{-3t} \begin{bmatrix} e^{2t} & 0 & 0 \\ 0 & e^{2t} \cos t + e^{2t} \sin t & -(e^{2t} \cos t - e^{2t} \sin)t \\ 0 & -e^{2t} \sin t & e^{2t} \cos(t) \end{bmatrix}^T$$
$$= \begin{bmatrix} e^{-t} & 0 & 0 \\ 0 & e^{-t}(\cos t + \sin t) & -(\sin t)e^{-t} \\ 0 & -e^{-t}(\cos t - \sin t) & (\cos t)e^{-t} \end{bmatrix}$$

Exercise 3.E.39

Find the inverse, if it exists, of the matrix

$$A = \begin{bmatrix} e^t & \cos t & \sin t \\ e^t & -\sin t & \cos t \\ e^t & -\cos t & -\sin t \end{bmatrix}$$

Answer

$$\begin{bmatrix} e^t & \cos t & \sin t \\ e^t & -\sin t & \cos t \\ e^t & -\cos t & -\sin t \end{bmatrix}^{-1}$$
$$= \begin{bmatrix} \frac{1}{2}e^{-t} & 0 & \frac{1}{2}e^{-t} \\ \frac{1}{2}\cos t + \frac{1}{2}\sin t & -\sin t & \frac{1}{2}\sin t - \frac{1}{2}\cos t \\ \frac{1}{2}\sin t - \frac{1}{2}\cos t & \cos t & -\frac{1}{2}\cos t - \frac{1}{2}\sin t \end{bmatrix}$$

? Exercise 3.E. 40

Suppose A is an upper triangular matrix. Show that A^{-1} exists if and only if all elements of the main diagonal are non zero. Is it true that A^{-1} will also be upper triangular? Explain. Could the same be concluded for lower triangular matrices?

Answer

The given condition is what it takes for the determinant to be non zero. Recall that the determinant of an upper triangular matrix is just the product of the entries on the main diagonal

? Exercise 3.E. 41

If A, B, and C are each $n \times n$ matrices and ABC is invertible, show why each of A, B, and C are invertible.

Answer

This follows because $\det(ABC) = \det(A)\det(B)\det(C)$ and if this product is nonzero, then each determinant in the product is nonzero and so each of these matrices is invertible.

? Exercise 3.E. 42

Decide if this statement is true or false: Cramer's rule is useful for finding solutions to systems of linear equations in which there is an infinite set of solutions.

Answer

False.

? Exercise 3.E. 43

Use Cramer's rule to find the solution to

$$\begin{aligned} x + 2y &= 1 \\ 2x - y &= 2 \end{aligned}$$

Answer

Solution is: $[x = 1, y = 0]$

? Exercise 3.E. 44

Use Cramer's rule to find the solution to

$$\begin{aligned} x + 2y + z &= 1 \\ 2x - y - z &= 2 \\ x + z &= 1 \end{aligned}$$

Answer

Solution is: $[x=1, y=0, z=0]$. For example,

$$y = \frac{\begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & -1 \\ 1 & 1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & 2 & 1 \\ 2 & -1 & -1 \\ 1 & 0 & 1 \end{vmatrix}} = 0$$

This page titled 3.E: Exercises is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

CHAPTER OVERVIEW

4: R^n

4.1: Vectors in R^n
4.2: Vector Algebra
4.3: Geometric Meaning of Vector Addition
4.4: Length of a Vector
4.5: Geometric Meaning of Scalar Multiplication
4.6: Parametric Lines
4.7: The Dot Product
4.8: Planes in R^n
4.9: The Cross Product
4.10: Spanning, Linear Independence and Basis in R^n
4.11: Orthogonality
4.12: Applications
4.E: Exercises

This page titled 4: R^n is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

4.1: Vectors in Rⁿ

Outcomes

A. Find the position vector of a point in $\mathbb{R}^n$.

The notation $\mathbb{R}^n$ refers to the collection of ordered lists of n real numbers, that is

$$\mathbb{R}^n = \{(x_1 \cdots x_n) : x_j \in \mathbb{R} \text{ for } j = 1, \cdots, n\}$$

In this chapter, we take a closer look at vectors in $\mathbb{R}^n$. First, we will consider what $\mathbb{R}^n$ looks like in more detail. Recall that the point given by $0 = (0, \cdots, 0)$ is called the **origin**.

Now, consider the case of $\mathbb{R}^n$ for $n = 1$. Then from the definition we can identify $\mathbb{R}$ with points in $\mathbb{R}^1$ as follows:

$$\mathbb{R} = \mathbb{R}^1 = \{(x_1) : x_1 \in \mathbb{R}\}$$

Hence, $\mathbb{R}$ is defined as the set of all real numbers and geometrically, we can describe this as all the points on a line.

Now suppose $n = 2$. Then, from the definition,

$$\mathbb{R}^2 = \{(x_1, x_2) : x_j \in \mathbb{R} \text{ for } j = 1, 2\}$$

Consider the familiar coordinate plane, with an x axis and a y axis. Any point within this coordinate plane is identified by where it is located along the x axis, and also where it is located along the y axis. Consider as an example the following diagram.

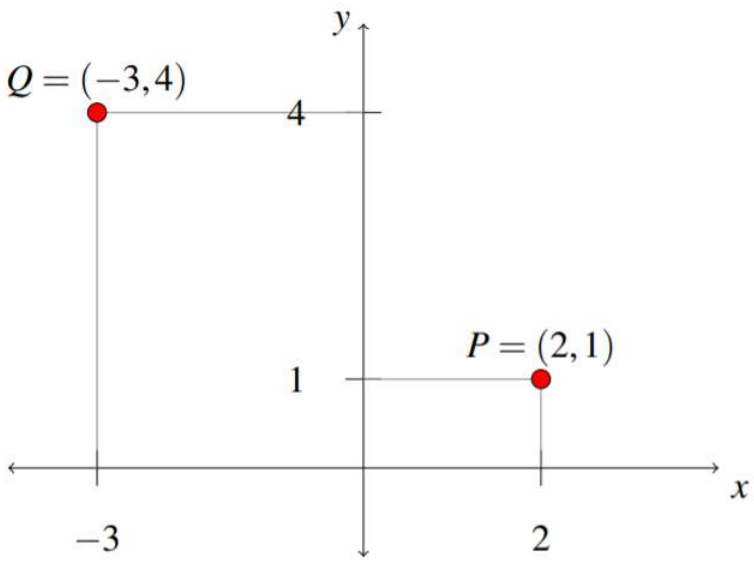


Figure 4.1.1

Hence, every element in $\mathbb{R}^2$ is identified by two components, x and y, in the usual manner. The coordinates x, y (or x_1,x_2) uniquely determine a point in the plan. Note that while the definition uses x_1 and x_2 to label the coordinates and you may be used to x and y, these notations are equivalent.

Now suppose $n = 3$. You may have previously encountered the 3-dimensional coordinate system, given by

$$\mathbb{R}^3 = \{(x_1, x_2, x_3) : x_j \in \mathbb{R} \text{ for } j = 1, 2, 3\}$$

Points in $\mathbb{R}^3$ will be determined by three coordinates, often written (x, y, z) which correspond to the x, y, and z axes. We can think as above that the first two coordinates determine a point in a plane. The third component determines the height above or below the plane, depending on whether this number is positive or negative, and all together this determines a point in space. You see that the ordered triples correspond to points in space just as the ordered pairs correspond to points in a plane and single real numbers correspond to points on a line.

The idea behind the more general $\mathbb{R}^n$ is that we can extend these ideas beyond $n = 3$. This discussion regarding points in $\mathbb{R}^n$ leads into a study of vectors in $\mathbb{R}^n$. While we consider $\mathbb{R}^n$ for all n, we will largely focus on $n = 2, 3$ in this section.

Consider the following definition.

Definition 4.1.1 The Position Vector

Let $P = (p_1, \cdots, p_n)$ be the coordinates of a point in $\mathbb{R}^n$. Then the vector $\overrightarrow{0P}$ with its tail at $0 = (0, \cdots, 0)$ and its tip at P is called the **position vector** of the point P. We write

$$\overrightarrow{0P} = \begin{bmatrix} p_1 \\ \vdots \\ p_n \end{bmatrix}$$

For this reason we may write both $P = (p_1, \cdots, p_n) \in \mathbb{R}^n$ and $\overrightarrow{0P} = [p_1 \cdots p_n]^T \in \mathbb{R}^n$.

This definition is illustrated in the following picture for the special case of $\mathbb{R}^3$.

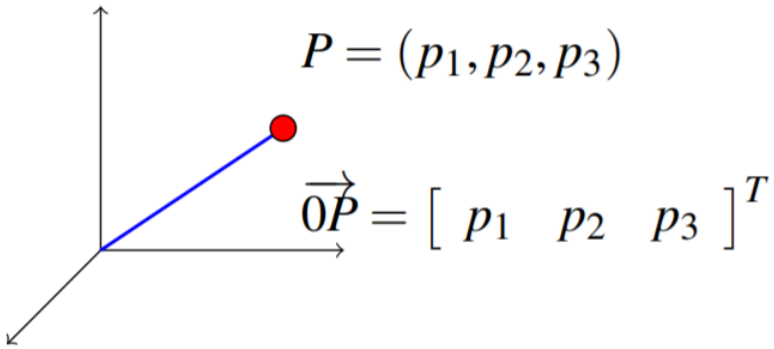


Figure 4.1.2

Thus every point P in $\mathbb{R}^n$ determines its position vector $\overrightarrow{0P}$. Conversely, every such position vector $\overrightarrow{0P}$ which has its tail at 0 and point at P determines the point P of $\mathbb{R}^n$.

Now suppose we are given two points, P, Q whose coordinates are $(p_1, \cdots, p_n)$ and $(q_1, \cdots, q_n)$ respectively. We can also determine the **position vector from P to Q** (also called the **vector from P to Q**) defined as follows.

$$\overrightarrow{PQ} = \begin{bmatrix} q_1 - p_1 \\ \vdots \\ q_n - p_n \end{bmatrix} = \overrightarrow{0Q} - \overrightarrow{0P}$$

Now, imagine taking a vector in $\mathbb{R}^n$ and moving it around, always keeping it pointing in the same direction as shown in the following picture.

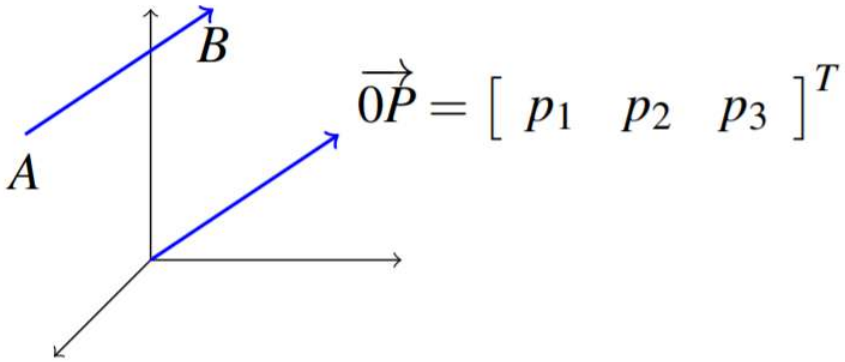


Figure 4.1.3

After moving it around, it is regarded as the same vector. Each vector, $\overrightarrow{0P}$ and $\overrightarrow{AB}$ has the same length (or magnitude) and direction. Therefore, they are equal.

Consider now the general definition for a vector in $\mathbb{R}^n$.

Definition 4.1.2 Vectors in $\mathbb{R}^n$

Let $\mathbb{R}^n = \{(x_1, \cdots, x_n) : x_j \in \mathbb{R} \text{ for } j = 1, \cdots, n\}$. Then,

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is called a **vector**. Vectors have both size (magnitude) and direction. The numbers x_j are called the **components** of $\vec{x}$.

Using this notation, we may use $\vec{p}$ to denote the position vector of point P. Notice that in this context, $\vec{p} = \overrightarrow{0P}$. These notations may be used interchangeably.

You can think of the components of a vector as directions for obtaining the vector. Consider $n = 3$. Draw a vector with its tail at the point $(0, 0, 0)$ and its tip at the point (a, b, c). This vector it is obtained by starting at $(0, 0, 0)$, moving parallel to the x axis to $(a, 0, 0)$ and then from here, moving parallel to the y axis to $(a, b, 0)$ and finally parallel to the z axis to (a, b, c). Observe that the same vector would result if you began at the point (d, e, f), moved parallel to the x axis to $(d+a, e, f)$, then parallel to the y axis to $(d+a, e+b, f)$, and finally parallel to the z axis to $(d+a, e+b, f+c)$. Here, the vector would have its tail sitting at the point determined by $A = (d, e, f)$ and its point at $B = (d+a, e+b, f+c)$. It is the **same vector** because it will point in the same direction and have the same length. It is like you took an actual arrow, and moved it from one location to another keeping it pointing the same direction.

We conclude this section with a brief discussion regarding notation. In previous sections, we have written vectors as columns, or $n \times 1$ matrices. For convenience in this chapter we may write vectors as the transpose of row vectors, or $1 \times n$ matrices. These are of course equivalent and we may move between both notations. Therefore, recognize that

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 \end{bmatrix}^T$$

Notice that two vectors $\vec{u} = [u_1 \cdots u_n]^T$ and $\vec{v} = [v_1 \cdots v_n]^T$ are equal if and only if all corresponding components are equal. Precisely,

$$\begin{gathered} \vec{u} = \vec{v} \text{ if and only if} \\ u_j = v_j \text{ for all } j = 1, \cdots, n \end{gathered}$$

Thus $\begin{bmatrix} 1 & 2 & 4 \end{bmatrix}^T \in \mathbb{R}^3$ and $\begin{bmatrix} 2 & 1 & 4 \end{bmatrix}^T \in \mathbb{R}^3$ but $\begin{bmatrix} 1 & 2 & 4 \end{bmatrix}^T \neq \begin{bmatrix} 2 & 1 & 4 \end{bmatrix}^T$ because, even though the same numbers are involved, the order of the numbers is different.

For the specific case of $\mathbb{R}^3$, there are three special vectors which we often use. They are given by

$$\vec{i} = [\,1 \quad 0 \quad 0\,]^T$$
$$\vec{j} = [\,0 \quad 1 \quad 0\,]^T$$
$$\vec{k} = [\,0 \quad 0 \quad 1\,]^T$$

We can write any vector $\vec{u} = [\,u_1 \quad u_2 \quad u_3\,]^T$ as a linear combination of these vectors, written as $\vec{u} = u_1\vec{i} + u_2\vec{j} + u_3\vec{k}$. This notation will be used throughout this chapter.

Below is a video on vectors in 2D.



Below is a video on finding a unit vector given the graph of a vector in 2D.



Below is a video on finding the component form of a vector from its graph.



Below is a video on plotting vectors in 3D.



This page titled 4.1: Vectors in R^n is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.1: Vectors in R^n** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.2: Vector Algebra

Outcomes

A. Understand vector addition and scalar multiplication, algebraically.
B. Introduce the notion of linear combination of vectors.

Addition and scalar multiplication are two important algebraic operations done with vectors. Notice that these operations apply to vectors in $\mathbb{R}^n$, for any value of n. We will explore these operations in more detail in the following sections.

Addition of Vectors in $\mathbb{R}^n$

Addition of vectors in $\mathbb{R}^n$ is defined as follows.

Definition 4.2.1: Addition of Vectors in $\mathbb{R}^n$

If $\vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$, $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$ then $\vec{u} + \vec{v} \in \mathbb{R}^n$ and is defined by

$$\begin{aligned} \vec{u} + \vec{v} &= \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \\ &= \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix} \end{aligned}$$

To add vectors, we simply add corresponding components. Therefore, in order to add vectors, they must be the same size.

Below is a video on finding the sum of two vectors as a linear combination of other vectors.



Addition of vectors satisfies some important properties which are outlined in the following theorem.

Theorem 4.2.1: Properties of Vector Addition

The following properties hold for vectors $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$.

- The Commutative Law of Addition

$$\vec{u} + \vec{v} = \vec{v} + \vec{u}$$

- The Associative Law of Addition

$$(\vec{u}+\vec{v})+\vec{w}=\vec{u}+(\vec{v}+\vec{w})$$

- The Existence of an Additive Identity

$$\vec{u}+\vec{0}=\vec{u} \tag{4.2.1}$$

- The Existence of an Additive Inverse

$$\vec{u}+(-\vec{u})=\vec{0}$$

The additive identity shown in Equation (4.2.1) is also called the **zero vector**, the $n\times 1$ vector in which all components are equal to 0. Further, $-\vec{u}$ is simply the vector with all components having same value as those of $\vec{u}$ but opposite sign; this is just $(-1)\vec{u}$. This will be made more explicit in the next section when we explore scalar multiplication of vectors. Note that subtraction is defined as $\vec{u}-\vec{v}=\vec{u}+(-\vec{v})$.

Scalar Multiplication of Vectors in $\mathbb{R}^n$

Scalar multiplication of vectors in $\mathbb{R}^n$ is defined as follows.

Definition 4.2.2: Scalar Multiplication of Vectors in $\mathbb{R}^n$

If $\vec{u}\in\mathbb{R}^n$ and $k\in\mathbb{R}$ is a scalar, then $k\vec{u}\in\mathbb{R}^n$ is defined by

$$k\vec{u}=k\begin{bmatrix}u_1\\ \vdots\\ u_n\end{bmatrix}=\begin{bmatrix}ku_1\\ \vdots\\ ku_n\end{bmatrix}$$

Just as with addition, scalar multiplication of vectors satisfies several important properties. These are outlined in the following theorem.

Theorem 4.2.2: Properties of Scalar Multiplication

The following properties hold for vectors $\vec{u},\vec{v}\in\mathbb{R}^n$ and k,p scalars.

- The Distributive Law over Vector Addition

$$k\left(\vec{u}+\vec{v}\right)=k\vec{u}+k\vec{v}$$

- The Distributive Law over Scalar Addition

$$\left(k+p\right)\vec{u}=k\vec{u}+p\vec{u}$$

- The Associative Law for Scalar Multiplication

$$k\left(p\vec{u}\right)=\left(kp\right)\vec{u}$$

- Rule for Multiplication by 1

$$1\vec{u}=\vec{u}$$

Proof

We will show the proof of:

$$k\left(\vec{u}+\vec{v}\right)=k\vec{u}+k\vec{v}$$

Note that:

$$\begin{aligned} k\left(\vec{u}+\vec{v}\right) &= k\left[u_1+v_1\cdots u_n+v_n\right]^T \\ &= \left[k\left(u_1+v_1\right)\cdots k\left(u_n+v_n\right)\right]^T \\ &= \left[ku_1+kv_1\cdots ku_n+kv_n\right]^T \\ &= \left[ku_1\cdots ku_n\right]^T+\left[kv_1\cdots kv_n\right]^T \\ &= k\vec{u}+k\vec{v} \end{aligned}$$

Below is a video on the dot product.



Below is a video on vector scalar multiplication.



We now present a useful notion you may have seen earlier combining vector addition and scalar multiplication

Definition 4.2.3: Linear Combination

A vector $\vec{v}$ is said to be a **linear combination** of the vectors $\vec{u}_1,\cdots,\vec{u}_n$ if there exist scalars, $a_1,\cdots,a_n$ such that

$$\vec{v}=a_1\vec{u}_1+\cdots+a_n\vec{u}_n$$

For example,

$$3\begin{bmatrix}-4\\1\\0\end{bmatrix}+2\begin{bmatrix}-3\\0\\1\end{bmatrix}=\begin{bmatrix}-18\\3\\2\end{bmatrix}.$$

Thus we can say that

$$\vec{v} = \begin{bmatrix} -18 \\ 3 \\ 2 \end{bmatrix}$$

is a linear combination of the vectors

$$\vec{u}_1 = \begin{bmatrix} -4 \\ 1 \\ 0 \end{bmatrix} \text{ and } \vec{u}_2 = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

This page titled 4.2: Vector Algebra is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.2: Vector Algebra** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.3: Geometric Meaning of Vector Addition

Learning Objectives

A. Understand vector addition, geometrically.

Recall that an element of $\mathbb{R}^n$ is an ordered list of numbers. For the specific case of $n = 2, 3$ this can be used to determine a point in two or three dimensional space. This point is specified relative to some coordinate axes.

Consider the case $n = 3$. Recall that taking a vector and moving it around without changing its length or direction does not change the vector. This is important in the geometric representation of vector addition.

Suppose we have two vectors, $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^3$. Each of these can be drawn geometrically by placing the tail of each vector at 0 and its point at (u_1, u_2, u_3) and (v_1, v_2, v_3) respectively. Suppose we slide the vector $\vec{v}$ so that its tail sits at the point of $\vec{u}$. We know that this does not change the vector $\vec{v}$. Now, draw a new vector from the tail of $\vec{u}$ to the point of $\vec{v}$. This vector is $\vec{u} + \vec{v}$.

The geometric significance of vector addition in $\mathbb{R}^n$ for any n is given in the following definition.

Definition 4.3.1: Geometry of Vector Addition

Let $\vec{u}$ and $\vec{v}$ be two vectors. Slide $\vec{v}$ so that the tail of $\vec{v}$ is on the point of $\vec{u}$. Then draw the arrow which goes from the tail of $\vec{u}$ to the point of $\vec{v}$. This arrow represents the vector $\vec{u} + \vec{v}$.

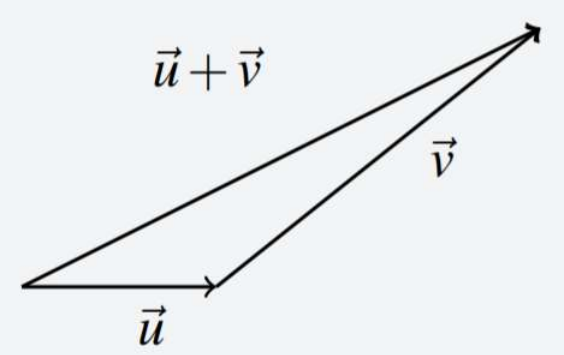


Figure 4.3.1

Below is a video on finding the sum of two vectors from a graph.



This definition is illustrated in the following picture in which $\vec{u} + \vec{v}$ is shown for the special case $n = 3$.

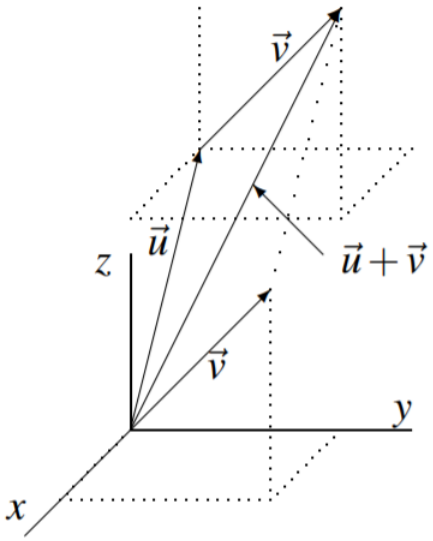


Figure 4.3.2

Notice the parallelogram created by $\vec{u}$ and $\vec{v}$ in the above diagram. Then $\vec{u}+\vec{v}$ is the directed diagonal of the parallelogram determined by the two vectors $\vec{u}$ and $\vec{v}$.

When you have a vector $\vec{v}$, its additive inverse $-\vec{v}$ will be the vector which has the same magnitude as $\vec{v}$ but the opposite direction. When one writes $\vec{u}-\overrightarrow{v}$, the meaning is $\vec{u}+(-\vec{v})$ as with real numbers. The following example illustrates these definitions and conventions.

Example 4.3.1: Graphing Vector Addition

Consider the following picture of vectors $\vec{u}$ and $\vec{v}$.



Figure 4.3.3

Sketch a picture of $\vec{u}+\vec{v}, \vec{u}-\vec{v}$.

Solution

We will first sketch $\vec{u}+\vec{v}$. Begin by drawing $\vec{u}$ and then at the point of $\vec{u}$, place the tail of $\vec{v}$ as shown. Then $\vec{u}+\vec{v}$ is the vector which results from drawing a vector from the tail of $\vec{u}$ to the tip of $\vec{v}$.

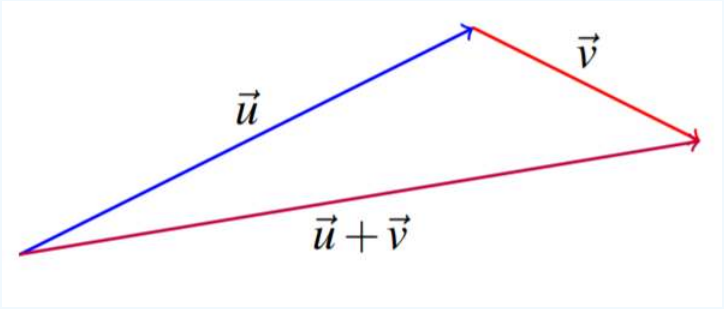


Figure 4.3.4

Next consider $\vec{u}-\vec{v}$. This means $\vec{u}+(-\vec{v})$. From the above geometric description of vector addition, $-\vec{v}$ is the vector which has the same length but which points in the opposite direction to $\vec{v}$. Here is a picture.

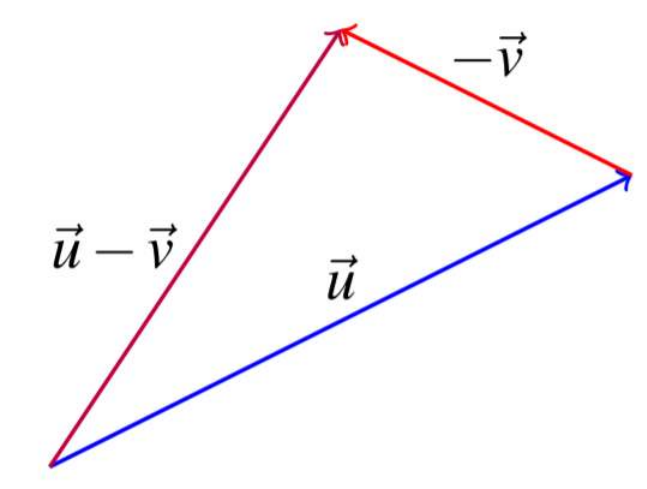


Figure 4.3.5

Below is a video on finding the difference of scalar multiples of vectors.



This page titled 4.3: Geometric Meaning of Vector Addition is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.3: Geometric Meaning of Vector Addition** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.4: Length of a Vector

Learning Objectives

- Find the length of a vector and the distance between two points in $\mathbb{R}^n$.
- Find the corresponding unit vector to a vector in $\mathbb{R}^n$.

In this section, we explore what is meant by the length of a vector in $\mathbb{R}^n$. We develop this concept by first looking at the distance between two points in $\mathbb{R}^n$.

First, we will consider the concept of distance for $\mathbb{R}$, that is, for points in $\mathbb{R}^1$. Here, the distance between two points P and Q is given by the absolute value of their difference. We denote the distance between P and Q by $d(P,Q)$ which is defined as

$$d(P,Q) = \sqrt{(P-Q)^2} \tag{4.4.1}$$

Consider now the case for $n = 2$, demonstrated by the following picture.

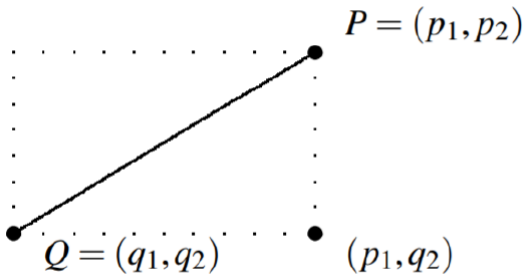


Figure 4.4.1

There are two points $P = (p_1, p_2)$ and $Q = (q_1, q_2)$ in the plane. The distance between these points is shown in the picture as a solid line. Notice that this line is the hypotenuse of a right triangle which is half of the rectangle shown in dotted lines. We want to find the length of this hypotenuse which will give the distance between the two points. Note the lengths of the sides of this triangle are $|p_1 - q_1|$ and $|p_2 - q_2|$, the absolute value of the difference in these values. Therefore, the Pythagorean Theorem implies the length of the hypotenuse (and thus the distance between P and Q) equals

$$\left(|p_1 - q_1|^2 + |p_2 - q_2|^2\right)^{1/2} = \left((p_1 - q_1)^2 + (p_2 - q_2)^2\right)^{1/2} \tag{4.4.2}$$

Now suppose $n = 3$ and let $P = (p_1, p_2, p_3)$ and $Q = (q_1, q_2, q_3)$ be two points in $\mathbb{R}^3$. Consider the following picture in which the solid line joins the two points and a dotted line joins the points (q_1, q_2, q_3) and (p_1, p_2, q_3).

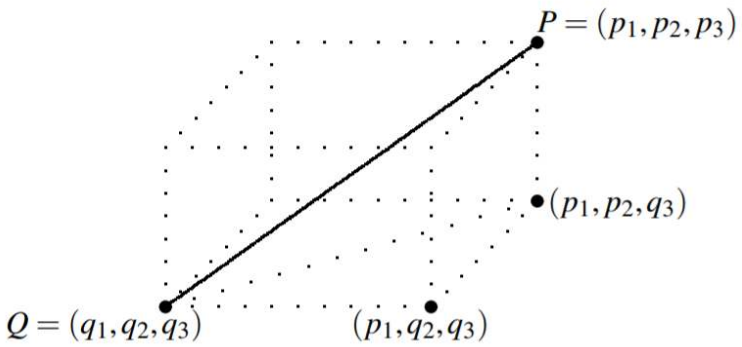


Figure 4.4.2

Here, we need to use Pythagorean Theorem twice in order to find the length of the solid line. First, by the Pythagorean Theorem, the length of the dotted line joining (q_1, q_2, q_3) and (p_1, p_2, q_3) equals

$$\left((p_1 - q_1)^2 + (p_2 - q_2)^2\right)^{1/2}$$

while the length of the line joining (p_1, p_2, q_3) to (p_1, p_2, p_3) is just $|p_3 - q_3|$. Therefore, by the Pythagorean Theorem again, the length of the line joining the points $P = (p_1, p_2, p_3)$ and $Q = (q_1, q_2, q_3)$ equals

$$\left(\left(\left((p_1 - q_1)^2 + (p_2 - q_2)^2 \right)^{1/2} \right)^2 + (p_3 - q_3)^2 \right)^{1/2}$$

$$= \left((p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2 \right)^{1/2} \tag{4.4.3}$$

This discussion motivates the following definition for the distance between points in $\mathbb{R}^n$.

Definition 4.4.1: Distance Between Points

Let $P = (p_1, \cdots, p_n)$ and $Q = (q_1, \cdots, q_n)$ be two points in $\mathbb{R}^n$. Then the distance between these points is defined as

$$\text{distance between } P \text{ and } Q \ = d(P, Q) = \left(\sum_{k=1}^{n} |p_k - q_k|^2 \right)^{1/2}$$

This is called the **distance formula**. We may also write $|P - Q|$ as the distance between P and Q.

From the above discussion, you can see that Definition 4.4.1 holds for the special cases $n = 1, 2, 3$, as in Equations (4.4.1), (4.4.2), (4.4.3). In the following example, we use Definition 4.4.1 to find the distance between two points in $\mathbb{R}^4$.

Example 4.4.1: Distance Between Points

Find the distance between the points P and Q in $\mathbb{R}^4$, where P and Q are given by

$$P = (1, 2, -4, 6)$$

and

$$Q = (2, 3, -1, 0)$$

Solution

We will use the formula given in Definition 4.4.1 to find the distance between P and Q. Use the distance formula and write

$$d(P, Q) = \left((1-2)^2 + (2-3)^2 + (-4-(-1))^2 + (6-0)^2 \right)^{\frac{1}{2}} = 47$$

Therefore, $d(P, Q) = \sqrt{47}$.

There are certain properties of the distance between points which are important in our study. These are outlined in the following theorem.

Theorem 4.4.1: Properties of Distance

Let P and Q be points in $\mathbb{R}^n$, and let the distance between them, $d(P, Q)$, be given as in Definition 4.4.1. Then, the following properties hold.

- $d(P, Q) = d(Q, P)$
- $d(P, Q) \geq 0$, and equals 0 exactly when $P = Q$.

There are many applications of the concept of distance. For instance, given two points, we can ask what collection of points are all the same distance between the given points. This is explored in the following example.

Example 4.4.2: The Plane Between Two Points

Describe the points in $\mathbb{R}^3$ which are at the same distance between $(1,2,3)$ and $(0,1,2)$.

Solution

Let $P=(p_1,p_2,p_3)$ be such a point. Therefore, P is the same distance from $(1,2,3)$ and $(0,1,2)$. Then by Definition 4.4.1,

$$\sqrt{(p_1-1)^2+(p_2-2)^2+(p_3-3)^2}=\sqrt{(p_1-0)^2+(p_2-1)^2+(p_3-2)^2}$$

Squaring both sides we obtain

$$(p_1-1)^2+(p_2-2)^2+(p_3-3)^2=p_1^2+(p_2-1)^2+(p_3-2)^2$$

and so

$$p_1^2-2p_1+14+p_2^2-4p_2+p_3^2-6p_3=p_1^2+p_2^2-2p_2+5+p_3^2-4p_3$$

Simplifying, this becomes

$$-2p_1+14-4p_2-6p_3=-2p_2+5-4p_3$$

which can be written as

$$2p_1+2p_2+2p_3=-9 \tag{4.4.4}$$

Therefore, the points $P=(p_1,p_2,p_3)$ which are the same distance from each of the given points form a plane whose equation is given by (4.4.4).

We can now use our understanding of the distance between two points to define what is meant by the length of a vector. Consider the following definition.

Definition 4.4.2: Length of a Vector

Let $\vec{u}=[u_1\cdots u_n]^T$ be a vector in $\mathbb{R}^n$. Then, the length of $\vec{u}$, written $\|\vec{u}\|$ is given by

$$\|\vec{u}\|=\sqrt{u_1^2+\cdots+u_n^2}$$

This definition corresponds to Definition 4.4.1, if you consider the vector $\vec{u}$ to have its tail at the point $0=(0,\cdots,0)$ and its tip at the point $U=(u_1,\cdots,u_n)$. Then the length of $\vec{u}$ is equal to the distance between 0 and U, $d(0,U)$. In general, $d(P,Q)=||\overrightarrow{PQ}||$.

Consider Example 4.4.1. By Definition 4.4.2, we could also find the distance between P and Q as the length of the vector connecting them. Hence, if we were to draw a vector $\overrightarrow{PQ}$ with its tail at P and its point at Q, this vector would have length equal to $\sqrt{47}$.

We conclude this section with a new definition for the special case of vectors of length 1.

Definition 4.4.3: Unit Vector

Let $\vec{u}$ be a vector in $\mathbb{R}^n$. Then, we call $\vec{u}$ a **unit vector** if it has length 1, that is if

$$\|\vec{u}\|=1$$

Let $\vec{v}$ be a vector in $\mathbb{R}^n$. Then, the vector $\vec{u}$ which has the same direction as $\vec{v}$ but length equal to 1 is the corresponding unit vector of $\vec{v}$. This vector is given by

$$\vec{u}=\frac{1}{\|\vec{v}\|}\vec{v}$$

We often use the term **normalize** to refer to this process. When we **normalize** a vector, we find the corresponding unit vector of length 1. Consider the following example.

Example 4.4.3: Finding a Unit Vector

Let $\vec{v}$ be given by

$$\vec{v} = \begin{bmatrix} 1 & -3 & 4 \end{bmatrix}^T$$

Find the unit vector $\vec{u}$ which has the same direction as $\vec{v}$.

Solution

We will use Definition 4.4.3 to solve this. Therefore, we need to find the length of $\vec{v}$ which, by Definition 4.4.2 is given by

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}$$

Using the corresponding values we find that

$$\begin{aligned} \|\vec{v}\| &= \sqrt{1^2 + (-3)^2 + 4^2} \\ &= \sqrt{1+9+16} \\ &= \sqrt{26} \end{aligned}$$

In order to find $\vec{u}$, we divide $\vec{v}$ by $\sqrt{26}$. The result is

$$\begin{aligned} \vec{u} &= \frac{1}{\|\vec{v}\|}\vec{v} \\ &= \frac{1}{\sqrt{26}}\begin{bmatrix} 1 & -3 & 4 \end{bmatrix}^T \\ &= \begin{bmatrix} \frac{1}{\sqrt{26}} & -\frac{3}{\sqrt{26}} & \frac{4}{\sqrt{26}} \end{bmatrix}^T \end{aligned}$$

You can verify using the Definition 4.4.1 that $\|\vec{u}\| = 1$.

Below is a video on finding the unit vector and drawing the vector and its unit vector.



Below is a video on finding the unit vector in the direction of a given vector.



This page titled 4.4: Length of a Vector is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.4: Length of a Vector** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.5: Geometric Meaning of Scalar Multiplication

Outcomes

A. Understand scalar multiplication, geometrically.

Recall that the point $P = (p_1, p_2, p_3)$ determines a vector $\vec{p}$ from 0 to P. The length of $\vec{p}$, denoted $\|\vec{p}\|$, is equal to $\sqrt{p_1^2 + p_2^2 + p_3^2}$ by Definition 4.4.1.

Now suppose we have a vector $\vec{u} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T$ and we multiply $\vec{u}$ by a scalar k. By Definition 4.2.2, $k\vec{u} = \begin{bmatrix} ku_1 & ku_2 & ku_3 \end{bmatrix}^T$. Then, by using Definition 4.4.1, the length of this vector is given by

$$\sqrt{\left((ku_1)^2 + (ku_2)^2 + (ku_3)^2 \right)} = |k| \sqrt{u_1^2 + u_2^2 + u_3^2}$$

Thus the following holds.

$$\|k\vec{u}\| = |k| \, \|\vec{u}\|$$

In other words, multiplication by a scalar magnifies or shrinks the length of the vector by a factor of $|k|$. If $|k| > 1$, the length of the resulting vector will be magnified. If $|k| < 1$, the length of the resulting vector will shrink. Remember that by the definition of the absolute value, $|k| > 0$.

What about the direction? Draw a picture of $\vec{u}$ and $k\vec{u}$ where k is negative. Notice that this causes the resulting vector to point in the opposite direction while if $k > 0$ it preserves the direction the vector points. Therefore the direction can either reverse, if $k < 0$, or remain preserved, if $k > 0$.

Consider the following example.

Example 4.5.1: Graphing Scalar Multiplication

Consider the vectors $\vec{u}$ and $\vec{v}$ drawn below.



Figure 4.5.1

Draw $-\vec{u}$, $2\vec{v}$, and $-\frac{1}{2}\vec{v}$.

Solution

In order to find $-\vec{u}$, we preserve the length of $\vec{u}$ and simply reverse the direction. For $2\vec{v}$, we double the length of $\vec{v}$, while preserving the direction. Finally $-\frac{1}{2}\vec{v}$ is found by taking half the length of $\vec{v}$ and reversing the direction. These vectors are shown in the following diagram.



Figure 4.5.2

Below is a video on vector scalar multiplication.

Now that we have studied both vector addition and scalar multiplication, we can combine the two actions. Recall Definition 9.2.2 of linear combinations of column matrices. We can apply this definition to vectors in $\mathbb{R}^n$. A linear combination of vectors in $\mathbb{R}^n$ is a sum of vectors multiplied by scalars.

In the following example, we examine the geometric meaning of this concept.

Example 4.5.2: Graphing a Linear Combination of Vectors

Consider the following picture of the vectors $\vec{u}$ and $\vec{v}$



Figure 4.5.3

Sketch a picture of $\vec{u}+2\vec{v}, \vec{u}-\frac{1}{2}\vec{v}$.

Solution

The two vectors are shown below.

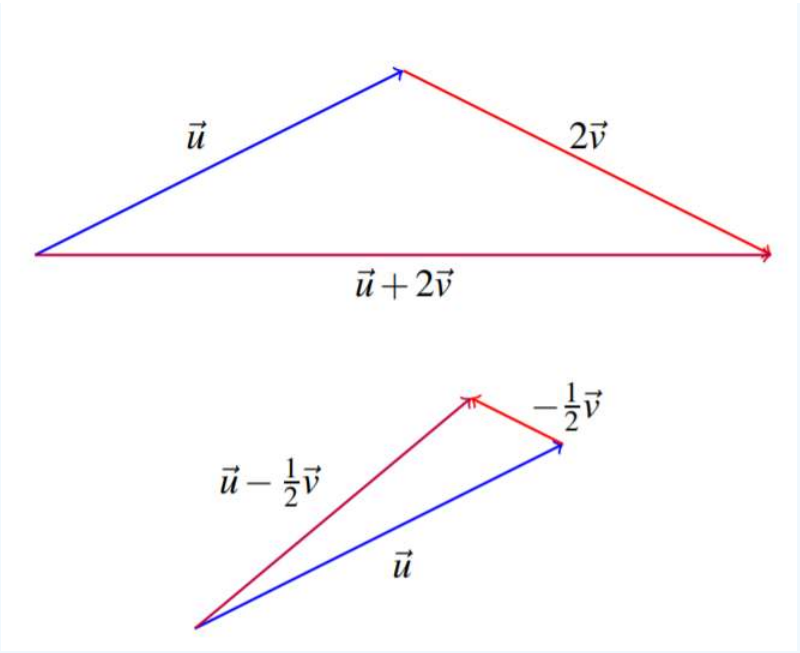


Figure 4.5.4

Below is a video on subtraction of two vectors.

Below is a video on the difference of scalar multiples of two vectors.

This page titled 4.5: Geometric Meaning of Scalar Multiplication is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.5: Geometric Meaning of Scalar Multiplication** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.6: Parametric Lines

Outcomes

A. Find the vector and parametric equations of a line.

We can use the concept of vectors and points to find equations for arbitrary lines in $\mathbb{R}^n$, although in this section the focus will be on lines in $\mathbb{R}^3$.

To begin, consider the case $n=1$ so we have $\mathbb{R}^1=\mathbb{R}$. There is only one line here which is the familiar number line, that is $\mathbb{R}$ itself. Therefore it is not necessary to explore the case of $n=1$ further.

Now consider the case where $n=2$, in other words $\mathbb{R}^2$. Let P and P_0 be two different points in $\mathbb{R}^2$ which are contained in a line L. Let $\vec{p}$ and $\overrightarrow{p_0}$ be the position vectors for the points P and P_0 respectively. Suppose that Q is an arbitrary point on L. Consider the following diagram.

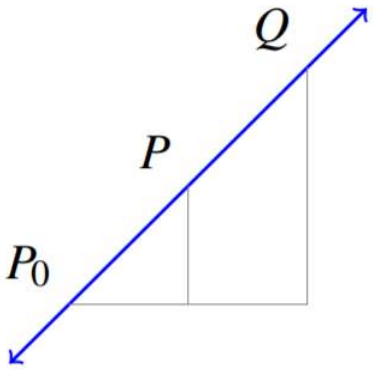


Figure 4.6.1

Our goal is to be able to define Q in terms of P and P_0. Consider the vector $\overrightarrow{P_0P}=\vec{p}-\overrightarrow{p_0}$ which has its tail at P_0 and point at P. If we add $\vec{p}-\overrightarrow{p_0}$ to the position vector $\overrightarrow{p_0}$ for P_0, the sum would be a vector with its point at P. In other words,

$$\vec{p}=\overrightarrow{p_0}+(\vec{p}-\overrightarrow{p_0})$$

Now suppose we were to add $t(\vec{p}-\overrightarrow{p_0})$ to $\vec{p}$ where t is some scalar. You can see that by doing so, we could find a vector with its point at Q. In other words, we can find t such that

$$\vec{q}=\overrightarrow{p_0}+t\left(\vec{p}-\overrightarrow{p_0}\right)$$

This equation determines the line L in $\mathbb{R}^2$. In fact, it determines a line L in $\mathbb{R}^n$. Consider the following definition.

Definition 4.6.1: Vector Equation of a Line

Suppose a line L in $\mathbb{R}^n$ contains the two different points P and P_0. Let $\vec{p}$ and $\overrightarrow{p_0}$ be the position vectors of these two points, respectively. Then, L is the collection of points Q which have the position vector $\vec{q}$ given by

$$\vec{q}=\overrightarrow{p_0}+t\left(\vec{p}-\overrightarrow{p_0}\right)$$

where $t\in\mathbb{R}$.

Let $\vec{d}=\vec{p}-\overrightarrow{p_0}$. Then $\vec{d}$ is the **direction vector for L** and the **vector equation** for L is given by

$$\vec{p}=\overrightarrow{p_0}+t\vec{d}, t\in\mathbb{R}$$

Note that this definition agrees with the usual notion of a line in two dimensions and so this is consistent with earlier concepts. Consider now points in $\mathbb{R}^3$. If a point $P \in \mathbb{R}^3$ is given by $P = (x, y, z)$, $P_0 \in \mathbb{R}^3$ by $P_0 = (x_0, y_0, z_0)$, then we can write

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + t \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

where $\vec{d} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. This is the vector equation of L written in **component form** .

The following theorem claims that such an equation is in fact a line.

Proposition 4.6.1: Algebraic Description of a Straight Line

Let $\vec{a}, \vec{b} \in \mathbb{R}^n$ with $\vec{b} \neq \vec{0}$. Then $\vec{x} = \vec{a} + t\vec{b},\ t \in \mathbb{R}$, is a line.

Proof

Let $\overrightarrow{x_1}, \overrightarrow{x_2} \in \mathbb{R}^n$. Define $\overrightarrow{x_1} = \vec{a}$ and let $\overrightarrow{x_2} - \overrightarrow{x_1} = \vec{b}$. Since $\vec{b} \neq \vec{0}$, it follows that $\overrightarrow{x_2} \neq \overrightarrow{x_1}$. Then $\vec{a} + t\vec{b} = \overrightarrow{x_1} + t\left(\overrightarrow{x_2} - \overrightarrow{x_1}\right)$. It follows that $\vec{x} = \vec{a} + t\vec{b}$ is a line containing the two different points X_1 and X_2 whose position vectors are given by $\vec{x}_1$ and $\vec{x}_2$ respectively.

We can use the above discussion to find the equation of a line when given two distinct points. Consider the following example.

Example 4.6.1: A Line From Two Points

Find a vector equation for the line through the points $P_0 = (1, 2, 0)$ and $P = (2, -4, 6)$.

Solution

We will use the definition of a line given above in Definition 4.6.1 to write this line in the form

$$\vec{q} = \overrightarrow{p_0} + t\left(\vec{p} - \overrightarrow{p_0}\right)$$

Let $\vec{q} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} B$. Then, we can find $\vec{p}$ and $\overrightarrow{p_0}$ by taking the position vectors of points P and P_0 respectively. Then,

$$\vec{q} = \overrightarrow{p_0} + t\left(\vec{p} - \overrightarrow{p_0}\right)$$

can be written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} B = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} B + t \begin{bmatrix} 1 \\ -6 \\ 6 \end{bmatrix} B,\ t \in \mathbb{R}$$

Here, the direction vector $\begin{bmatrix} 1 \\ -6 \\ 6 \end{bmatrix} B$ is obtained by $\vec{p} - \overrightarrow{p_0} = \begin{bmatrix} 2 \\ -4 \\ 6 \end{bmatrix} B - \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} B$ as indicated above in Definition 4.6.1.

Notice that in the above example we said that we found “a” vector equation for the line, not “the” equation. The reason for this terminology is that there are infinitely many different vector equations for the same line. To see this, replace t with another parameter, say $3s$. Then you obtain a different vector equation for the same line because the same set of points is obtained.

In Example 4.6.1, the vector given by $\begin{bmatrix} 1 \\ -6 \\ 6 \end{bmatrix} B$ is the direction vector defined in Definition 4.6.1. If we know the direction vector of a line, as well as a point on the line, we can find the vector equation.

Consider the following example.

Example 4.6.2: A Line From a Point and a Direction Vector

Find a vector equation for the line which contains the point $P_0 = (1, 2, 0)$ and has direction vector $\vec{d} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} B$

Solution

We will use Definition 4.6.1 to write this line in the form $\vec{p} = \overrightarrow{p_0} + t\vec{d},\ t \in \mathbb{R}$. We are given the direction vector $\vec{d}$. In order to find $\overrightarrow{p_0}$, we can use the position vector of the point P_0. This is given by $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} B$. Letting $\vec{p} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} B$, the equation for the line is given by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} B = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} B + t \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} B,\ t \in \mathbb{R} \tag{4.6.1}$$

Below is a video on finding parametric equations of a line in 3D.



We sometimes elect to write a line such as the one given in (4.6.1) in the form

$$\left.\begin{array}{l} x = 1 + t \\ y = 2 + 2t \\ z = t \end{array}\right\} \quad \text{where } t \in \mathbb{R} \tag{4.6.2}$$

This set of equations give the same information as (4.6.1), and is called the **parametric equation of the line**.

Consider the following definition.

Definition 4.6.2: Parametric Equation of a Line

Let L be a line in $\mathbb{R}^3$ which has direction vector $\vec{d} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} B$ and goes through the point $P_0 = (x_0, y_0, z_0)$. Then, letting t be a parameter, we can write L as

$$\left.\begin{array}{l} x = x_0 + ta \\ y = y_0 + tb \\ z = z_0 + tc \end{array}\right\} \quad \text{where } t \in \mathbb{R}$$

This is called a **parametric equation** of the line L.

You can verify that the form discussed following Example 4.6.2 in equation (4.6.2) is of the form given in Definition 4.6.2.

There is one other form for a line which is useful, which is the **symmetric form**. Consider the line given by (4.6.2). You can solve for the parameter t to write

$$\begin{array}{l} t = x - 1 \\ t = \frac{y-2}{2} \\ t = z \end{array}$$

Therefore,

$$x - 1 = \frac{y-2}{2} = z$$

This is the **symmetric form** of the line.

In the following example, we look at how to take the equation of a line from symmetric form to parametric form.

Example 4.6.3: Change Symmetric Form to Parametric Form

Suppose the **symmetric form of a line** is

$$\frac{x-2}{3} = \frac{y-1}{2} = z+3$$

Write the line in parametric form as well as vector form.

Solution

We want to write this line in the form given by Definition 4.6.2. This is of the form

$$\left.\begin{array}{l} x = x_0 + ta \\ y = y_0 + tb \\ z = z_0 + tc \end{array}\right\} \quad \text{where } t \in \mathbb{R}$$

Let $t = \frac{x-2}{3}, t = \frac{y-1}{2}$ and $t = z+3$, as given in the symmetric form of the line. Then solving for x, y, z, yields

$$\left.\begin{array}{l} x = 2 + 3t \\ y = 1 + 2t \\ z = -3 + t \end{array}\right\} \quad \text{with } t \in \mathbb{R}$$

This is the parametric equation for this line.

Now, we want to write this line in the form given by Definition 4.6.1. This is the form

$$\vec{p} = \overrightarrow{p_0} + t\vec{d}$$

where $t \in \mathbb{R}$. This equation becomes

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} B = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix} B + t \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} B, \ t \in \mathbb{R}$$

This page titled 4.6: Parametric Lines is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.6: Parametric Lines** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.7: The Dot Product

Outcomes

A. Compute the dot product of vectors, and use this to compute vector projections.

The Dot Product

There are two ways of multiplying vectors which are of great importance in applications. The first of these is called the **dot product**. When we take the dot product of vectors, the result is a scalar. For this reason, the dot product is also called the **scalar product** and sometimes the **inner product**. The definition is as follows.

Definition 4.7.1: Dot Product

Let $\vec{u}, \vec{v}$ be two vectors in $\mathbb{R}^n$. Then we define the **dot product** $\vec{u} \bullet \vec{v}$ as

$$\vec{u} \bullet \vec{v} = \sum_{k=1}^{n} u_k v_k$$

The dot product $\vec{u} \bullet \vec{v}$ is sometimes denoted as $(\vec{u}, \vec{v})$ where a comma replaces $\bullet$. It can also be written as $\langle \vec{u}, \vec{v} \rangle$. If we write the vectors as column or row matrices, it is equal to the matrix product $\vec{v}\vec{w}^T$.

Consider the following example.

Example 4.7.1: Compute a Dot Product

Find $\vec{u} \bullet \vec{v}$ for

$$\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix}, \vec{v} = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}$$

Solution

By Definition 4.7.1, we must compute

$$\vec{u} \bullet \vec{v} = \sum_{k=1}^{4} u_k v_k$$

This is given by

$$\begin{aligned} \vec{u} \bullet \vec{v} &= (1)(0) + (2)(1) + (0)(2) + (-1)(3) \\ &= 0 + 2 + 0 + -3 \\ &= -1 \end{aligned}$$

Below is a video on finding the dot product of 2 vectors in 2D.



Below is a video on finding the dot product of 2 vectors from a given graph.



Below is a video on finding the dot product of 2 vectors in 3D.



With this definition, there are several important properties satisfied by the dot product.

Proposition 4.7.1: Properties of the Dot Product

Let k and p denote scalars and $\vec{u}, \vec{v}, \vec{w}$ denote vectors. Then the dot product $\vec{u} \bullet \vec{v}$ satisfies the following properties.

- $\vec{u} \bullet \vec{v} = \vec{v} \bullet \vec{u}$
- $\vec{u} \bullet \vec{u} \geq 0$ and equals zero if and only if $\vec{u} = \vec{0}$
- $(k\vec{u} + p\vec{v}) \bullet \vec{w} = k(\vec{u} \bullet \vec{w}) + p(\vec{v} \bullet \vec{w})$

- $\vec{u} \bullet (k\vec{v} + p\vec{w}) = k(\vec{u} \bullet \vec{v}) + p(\vec{u} \bullet \vec{w})$
- $\|\vec{u}\|^2 = \vec{u} \bullet \vec{u}$

Proof

The proof is left as an exercise.

This proposition tells us that we can also use the dot product to find the length of a vector.

Example 4.7.2: Length of a Vector

Find the length of

$$\vec{u} = \begin{bmatrix} 2 \\ 1 \\ 4 \\ 2 \end{bmatrix}$$

That is, find $\|\vec{u}\|$.

Solution

By Proposition 4.7.1, $\|\vec{u}\|^2 = \vec{u} \bullet \vec{u}$. Therefore, $\|\vec{u}\| = \sqrt{\vec{u} \bullet \vec{u}}$. First, compute $\vec{u} \bullet \vec{u}$.

This is given by

$$\begin{aligned} \vec{u} \bullet \vec{u} &= (2)(2) + (1)(1) + (4)(4) + (2)(2) \\ &= 4 + 1 + 16 + 4 \\ &= 25 \end{aligned}$$

Then,

$$\begin{aligned} \|\vec{u}\| &= \sqrt{\vec{u} \bullet \vec{u}} \\ &= \sqrt{25} \\ &= 5 \end{aligned}$$

Below is a video on finding the dot product of 2 vectors given their magnitudes and directions.



You may wish to compare this to our previous definition of length, given in Definition 4.4.2.

The **Cauchy Schwarz inequality** is a fundamental inequality satisfied by the dot product. It is given in the following theorem.

Theorem 4.7.1: Cauchy Schwarz Inequality

The dot product satisfies the inequality

$$|\vec{u} \bullet \vec{v}| \leq \|\vec{u}\| \|\vec{v}\| \tag{4.7.1}$$

Furthermore equality is obtained if and only if one of $\vec{u}$ or $\vec{v}$ is a scalar multiple of the other.

Proof

First note that if $\vec{v} = \vec{0}$ both sides of (4.7.1) equal zero and so the inequality holds in this case. Therefore, it will be assumed in what follows that $\vec{v} \neq \vec{0}$.

Define a function of $t \in \mathbb{R}$ by

$$f(t) = (\vec{u} + t\vec{v}) \bullet (\vec{u} + t\vec{v})$$

Then by Proposition 4.7.1, $f(t) \geq 0$ for all $t \in \mathbb{R}$. Also from Proposition 4.7.1

$$\begin{aligned} f(t) &= \vec{u} \bullet (\vec{u} + t\vec{v}) + t\vec{v} \bullet (\vec{u} + t\vec{v}) \\ &= \vec{u} \bullet \vec{u} + t(\vec{u} \bullet \vec{v}) + t\vec{v} \bullet \vec{u} + t^2 \vec{v} \bullet \vec{v} \\ &= \|\vec{u}\|^2 + 2t(\vec{u} \bullet \vec{v}) + \|\vec{v}\|^2 t^2 \end{aligned}$$

Now this means the graph of $y = f(t)$ is a parabola which opens up and either its vertex touches the t axis or else the entire graph is above the t axis. In the first case, there exists some t where $f(t) = 0$ and this requires $\vec{u} + t\vec{v} = \vec{0}$ so one vector is a multiple of the other. Then clearly equality holds in (4.7.1). In the case where $\vec{v}$ is not a multiple of $\vec{u}$, it follows $f(t) > 0$ for all t which says $f(t)$ has no real zeros and so from the quadratic formula,

$$(2(\vec{u} \bullet \vec{v}))^2 - 4\|\vec{u}\|^2 \|\vec{v}\|^2 < 0$$

which is equivalent to $|\vec{u} \bullet \vec{v}| < \|\vec{u}\| \|\vec{v}\|$.

Notice that this proof was based only on the properties of the dot product listed in Proposition 4.7.1. This means that whenever an operation satisfies these properties, the Cauchy Schwarz inequality holds. There are many other instances of these properties besides vectors in $\mathbb{R}^n$.

The Cauchy Schwarz inequality provides another proof of the **triangle inequality** for distances in $\mathbb{R}^n$.

Theorem 4.7.2: Triangle Inequality

For $\vec{u}, \vec{v} \in \mathbb{R}^n$

$$\|\vec{u} + \vec{v}\| \leq \|\vec{u}\| + \|\vec{v}\| \tag{4.7.2}$$

and equality holds if and only if one of the vectors is a non-negative scalar multiple of the other.

Also

$$\|\|\vec{u}\| - \|\vec{v}\|\| \leq \|\vec{u} - \vec{v}\| \tag{4.7.3}$$

Proof

By properties of the dot product and the Cauchy Schwarz inequality,

$$\begin{aligned} \|\vec{u} + \vec{v}\|^2 &= (\vec{u} + \vec{v}) \bullet (\vec{u} + \vec{v}) \\ &= (\vec{u} \bullet \vec{u}) + (\vec{u} \bullet \vec{v}) + (\vec{v} \bullet \vec{u}) + (\vec{v} \bullet \vec{v}) \\ &= \|\vec{u}\|^2 + 2(\vec{u} \bullet \vec{v}) + \|\vec{v}\|^2 \\ &\leq \|\vec{u}\|^2 + 2|\vec{u} \bullet \vec{v}| + \|\vec{v}\|^2 \\ &\leq \|\vec{u}\|^2 + 2\|\vec{u}\| \|\vec{v}\| + \|\vec{v}\|^2 = (\|\vec{u}\| + \|\vec{v}\|)^2 \end{aligned}$$

Hence,

$$\|\vec{u}+\vec{v}\|^2 \leq (\|\vec{u}\|+\|\vec{v}\|)^2$$

Taking square roots of both sides you obtain (4.7.2).

It remains to consider when equality occurs. Suppose $\vec{u}=\vec{0}$. Then, $\vec{u}=0\vec{v}$ and the claim about when equality occurs is verified. The same argument holds if $\vec{v}=\vec{0}$. Therefore, it can be assumed both vectors are nonzero. To get equality in (4.7.2) above, Theorem 4.7.1 implies one of the vectors must be a multiple of the other. Say $\vec{v}=k\vec{u}$. If $k<0$ then equality cannot occur in (4.7.2) because in this case

$$\vec{u}\bullet\vec{v}=k\|\vec{u}\|^2<0<|k|\,\|\vec{u}\|^2=|\vec{u}\bullet\vec{v}|$$

Therefore, $k\geq 0$.

To get the other form of the triangle inequality write

$$\vec{u}=\vec{u}-\vec{v}+\vec{v}$$

so

$$\begin{aligned}\|\vec{u}\| &= \|\vec{u}-\vec{v}+\vec{v}\| \\ &\leq \|\vec{u}-\vec{v}\|+\|\vec{v}\|\end{aligned}$$

Therefore,

$$\|\vec{u}\|-\|\vec{v}\|\leq\|\vec{u}-\vec{v}\| \tag{4.7.4}$$

Similarly,

$$\|\vec{v}\|-\|\vec{u}\|\leq\|\vec{v}-\vec{u}\|=\|\vec{u}-\vec{v}\| \tag{4.7.5}$$

It follows from (4.7.4) and (4.7.5) that (4.7.3) holds. This is because $|\|\vec{u}\|-\|\vec{v}\||$ equals the left side of either (4.7.4) or (4.7.5) and either way, $|\|\vec{u}\|-\|\vec{v}\||\leq\|\vec{u}-\vec{v}\|$.

The Geometric Significance of the Dot Product

Given two vectors, $\vec{u}$ and $\vec{v}$, the **included angle** is the angle between these two vectors which is given by θ such that $0\leq\theta\leq\pi$. The dot product can be used to determine the included angle between two vectors. Consider the following picture where θ gives the included angle.

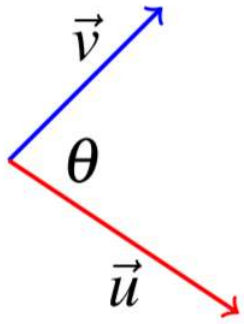


Figure 4.7.1

Proposition 4.7.2: The Dot Product and the Included Angle

Let $\vec{u}$ and $\vec{v}$ be two vectors in $\mathbb{R}^n$, and let θ be the included angle. Then the following equation holds.

$$\vec{u}\bullet\vec{v}=\|\vec{u}\|\|\vec{v}\|\cos\theta$$

In words, the dot product of two vectors equals the product of the magnitude (or length) of the two vectors multiplied by the cosine of the included angle. Note this gives a geometric description of the dot product which does not depend explicitly on the

coordinates of the vectors.

Consider the following example.

Example 4.7.3: Find the Angle Between Two Vectors

Find the angle between the vectors given by

$$\vec{u} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \vec{v} = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$$

Solution

By Proposition 4.7.2,

$$\vec{u} \bullet \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos\theta$$

Hence,

$$\cos\theta = \frac{\vec{u} \bullet \vec{v}}{\|\vec{u}\| \|\vec{v}\|}$$

First, we can compute $\vec{u} \bullet \vec{v}$. By Definition 4.7.1, this equals

$$\vec{u} \bullet \vec{v} = (2)(3) + (1)(4) + (-1)(1) = 9$$

Then,

$$\|\vec{u}\| = \sqrt{(2)(2) + (1)(1) + (1)(1)} = \sqrt{6}$$
$$\|\vec{v}\| = \sqrt{(3)(3) + (4)(4) + (1)(1)} = \sqrt{26}$$

Therefore, the cosine of the included angle equals

$$\cos\theta = \frac{9}{\sqrt{26}\sqrt{6}} = 0.7205766\ldots$$

With the cosine known, the angle can be determined by computing the inverse cosine of that angle, giving approximately $\theta = 0.76616$ radians.

Another application of the geometric description of the dot product is in finding the angle between two lines. Typically one would assume that the lines intersect. In some situations, however, it may make sense to ask this question when the lines do not intersect, such as the angle between two object trajectories. In any case we understand it to mean the smallest angle between (any of) their direction vectors. The only subtlety here is that if $\vec{u}$ is a direction vector for a line, then so is any multiple $k\vec{u}$, and thus we will find complementary angles among all angles between direction vectors for two lines, and we simply take the smaller of the two.

Example 4.7.4: Find the Angle Between Two Lines

Find the angle between the two lines

$$L_1 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

and

$$L_2 : \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \\ -3 \end{bmatrix} + s \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$

Solution

You can verify that these lines do not intersect, but as discussed above this does not matter and we simply find the smallest angle between any directions vectors for these lines.

To do so we first find the angle between the direction vectors given above:

$$\vec{u} = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \ \vec{v} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$$

In order to find the angle, we solve the following equation for θ

$$\vec{u} \bullet \vec{v} = \|\vec{u}\| \|\vec{v}\| \cos\theta$$

to obtain $\cos\theta = -\frac{1}{2}$ and since we choose included angles between 0 and π we obtain $\theta = \frac{2\pi}{3}$.

Now the angles between any two direction vectors for these lines will either be $\frac{2\pi}{3}$ or its complement $\phi = \pi - \frac{2\pi}{3} = \frac{\pi}{3}$. We choose the smaller angle, and therefore conclude that the angle between the two lines is $\frac{\pi}{3}$.

We can also use Proposition 4.7.2 to compute the dot product of two vectors.

Example 4.7.5: Using Geometric Description to Find a Dot Product

Let $\vec{u}, \vec{v}$ be vectors with $\|\vec{u}\| = 3$ and $\|\vec{v}\| = 4$. Suppose the angle between $\vec{u}$ and $\vec{v}$ is $\pi/3$. Find $\vec{u} \bullet \vec{v}$.

Solution

From the geometric description of the dot product in Proposition 4.7.2

$$\vec{u} \bullet \vec{v} = (3)(4)\cos(\pi/3) = 3 \times 4 \times 1/2 = 6$$

Two nonzero vectors are said to be **perpendicular**, sometimes also called **orthogonal**, if the included angle is $\pi/2$ radians (90°).

Consider the following proposition.

Proposition 4.7.3:Perpendicular Vectors

Let $\vec{u}$ and $\vec{v}$ be nonzero vectors in $\mathbb{R}^n$. Then, $\vec{u}$ and $\vec{v}$ are said to be **perpendicular** exactly when

$$\vec{u} \bullet \vec{v} = 0$$

Proof

This follows directly from Proposition 4.7.2 First if the dot product of two nonzero vectors is equal to 0, this tells us that $\cos\theta = 0$ (this is where we need nonzero vectors). Thus $\theta = \pi/2$ and the vectors are perpendicular.

If on the other hand $\vec{v}$ is perpendicular to $\vec{u}$, then the included angle is $\pi/2$ radians. Hence $\cos\theta = 0$ and $\vec{u} \bullet \vec{v} = 0$.

Consider the following example.

Example 4.7.6: Determine if Two Vectors are Perpendicular

Determine whether the two vectors,

$$\vec{u} = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \vec{v} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$$

are perpendicular.

Solution

In order to determine if these two vectors are perpendicular, we compute the dot product. This is given by

$$\vec{u} \bullet \vec{v} = (2)(1) + (1)(3) + (-1)(5) = 0$$

Therefore, by Proposition 4.7.3 these two vectors are perpendicular.

Projections

In some applications, we wish to write a vector as a sum of two related vectors. Through the concept of projections, we can find these two vectors. First, we explore an important theorem. The result of this theorem will provide our definition of a vector projection.

Theorem 4.7.3: Vector Projections

Let $\vec{v}$ and $\vec{u}$ be nonzero vectors. Then there exist unique vectors $\vec{v}_{||}$ and $\vec{v}_{\perp}$ such that

$$\vec{v} = \vec{v}_{||} + \vec{v}_{\perp} \tag{4.7.6}$$

where $\vec{v}_{||}$ is a scalar multiple of $\vec{u}$, and $\vec{v}_{\perp}$ is perpendicular to $\vec{u}$.

Proof

Suppose (4.7.6) holds and $\vec{v}_{||} = k\vec{u}$. Taking the dot product of both sides of (4.7.6) with $\vec{u}$ and using $\vec{v}_{\perp} \bullet \vec{u} = 0$, this yields

$$\begin{aligned} \vec{v} \bullet \vec{u} &= (\vec{v}_{||} + \vec{v}_{\perp}) \bullet \vec{u} \\ &= k\vec{u} \bullet \vec{u} + \vec{v}_{\perp} \bullet \vec{u} \\ &= k\|\vec{u}\|^2 \end{aligned}$$

which requires $k = \vec{v} \bullet \vec{u} / \|\vec{u}\|^2$. Thus there can be no more than one vector $\vec{v}_{||}$. It follows $\vec{v}_{\perp}$ must equal $\vec{v} - \vec{v}_{||}$. This verifies there can be no more than one choice for both $\vec{v}_{||}$ and $\vec{v}_{\perp}$ and proves their uniqueness.

Now let

$$\vec{v}_{||} = \frac{\vec{v} \bullet \vec{u}}{\|\vec{u}\|^2} \vec{u}$$

and let

$$\vec{v}_{\perp} = \vec{v} - \vec{v}_{||} = \vec{v} - \frac{\vec{v} \bullet \vec{u}}{\|\vec{u}\|^2} \vec{u}$$

Then $\vec{v}_{||} = k\vec{u}$ where $k = \frac{\vec{v} \bullet \vec{u}}{\|\vec{u}\|^2}$. It only remains to verify $\vec{v}_{\perp} \bullet \vec{u} = 0$. But

$$\begin{aligned} \vec{v}_{\perp} \bullet \vec{u} &= \vec{v} \bullet \vec{u} - \frac{\vec{v} \bullet \vec{u}}{\|\vec{u}\|^2} \vec{u} \bullet \vec{u} \\ &= \vec{v} \bullet \vec{u} - \vec{v} \bullet \vec{u} \\ &= 0 \end{aligned}$$

The vector $\vec{v}_{||}$ in Theorem 4.7.3 is called the **projection** of $\vec{v}$ onto $\vec{u}$ and is denoted by

$$\vec{v}_{||} = \text{proj}_{\vec{u}} (\vec{v})$$

We now make a formal definition of the vector projection.

Definition 4.7.2: Vector Projection

Let $\vec{u}$ and $\vec{v}$ be vectors. Then, the **projection of $\vec{v}$ onto $\vec{u}$** is given by

$$\text{proj}_{\vec{u}} (\vec{v}) = \left(\frac{\vec{v} \bullet \vec{u}}{\vec{u} \bullet \vec{u}} \right) \vec{u} = \frac{\vec{v} \bullet \vec{u}}{\|\vec{u}\|^2} \vec{u}$$

Below is a video on proving the vector projection formula.



Consider the following example of a projection.

Example 4.7.7: Find the Projection of One Vector Onto Another

Find $\text{proj}_{\vec{u}}(\vec{v})$ if

$$\vec{u} = \begin{bmatrix} 2 \\ 3 \\ -4 \end{bmatrix}, \vec{v} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Solution

We can use the formula provided in Definition 4.7.2 to find $\text{proj}_{\vec{u}}(\vec{v})$. First, compute $\vec{v} \bullet \vec{u}$. This is given by

$$\begin{aligned} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 2 \\ 3 \\ -4 \end{bmatrix} &= (2)(1) + (3)(-2) + (-4)(1) \\ &= 2 - 6 - 4 \\ &= -8 \end{aligned}$$

Similarly, $\vec{u} \bullet \vec{u}$ is given by

$$\begin{aligned} \begin{bmatrix} 2 \\ 3 \\ -4 \end{bmatrix} \bullet \begin{bmatrix} 2 \\ 3 \\ -4 \end{bmatrix} &= (2)(2) + (3)(3) + (-4)(-4) \\ &= 4 + 9 + 16 \\ &= 29 \end{aligned}$$

Therefore, the projection is equal to

$$\begin{aligned} \text{proj}_{\vec{u}}(\vec{v}) &= -\frac{8}{29} \begin{bmatrix} 2 \\ 3 \\ -4 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{16}{29} \\ -\frac{24}{29} \\ \frac{32}{29} \end{bmatrix} \end{aligned}$$

Below is a video on vector projection.



Below is another video on vector projection.



Below is a video on vector projection in 3D.



We will conclude this section with an important application of projections. Suppose a line L and a point P are given such that P is not contained in L. Through the use of projections, we can determine the shortest distance from P to L.

Example 4.7.8: Shortest Distance from a Point to a Line

Let $P = (1, 3, 5)$ be a point in $\mathbb{R}^3$, and let L be the line which goes through point $P_0 = (0, 4, -2)$ with direction vector $\vec{d} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$. Find the shortest distance from P to the line L, and find the point Q on L that is closest to P.

Solution

In order to determine the shortest distance from P to L, we will first find the vector $\overrightarrow{P_0P}$ and then find the projection of this vector onto L. The vector $\overrightarrow{P_0P}$ is given by

$$\begin{bmatrix}1\\3\\5\end{bmatrix}-\begin{bmatrix}0\\4\\-2\end{bmatrix}=\begin{bmatrix}1\\-1\\7\end{bmatrix}$$

Then, if Q is the point on L closest to P, it follows that

$$\begin{aligned}\overrightarrow{P_0Q}&=\text{proj}_{\vec{d}}\overrightarrow{P_0P}\\&=\left(\frac{\overrightarrow{P_0P}\bullet\vec{d}}{\|\vec{d}\|^2}\right)\vec{d}\\&=\frac{15}{9}\begin{bmatrix}2\\1\\2\end{bmatrix}\\&=\frac{5}{3}\begin{bmatrix}2\\1\\2\end{bmatrix}\end{aligned}$$

Now, the distance from P to L is given by

$$\|\overrightarrow{QP}\|=\|\overrightarrow{P_0P}-\overrightarrow{P_0Q}\|=\sqrt{26}$$

The point Q is found by adding the vector $\overrightarrow{P_0Q}$ to the position vector $\overrightarrow{0P_0}$ for P_0 as follows

$$\begin{bmatrix}0\\4\\-2\end{bmatrix}+\frac{5}{3}\begin{bmatrix}2\\1\\2\end{bmatrix}=\begin{bmatrix}\frac{10}{3}\\\frac{17}{3}\\\frac{4}{3}\end{bmatrix}$$

Therefore, $Q=(\frac{10}{3},\frac{17}{3},\frac{4}{3})$.

This page titled 4.7: The Dot Product is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.7: The Dot Product** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.8: Planes in Rⁿ

Outcomes

A. Find the vector and scalar equations of a plane.

Much like the above discussion with lines, vectors can be used to determine planes in $\mathbb{R}^n$. Given a vector $\vec{n}$ in $\mathbb{R}^n$ and a point P_0, it is possible to find a **unique** plane which contains P_0 and is perpendicular to the given vector.

Below is a video on graphing a plane using traces.



Definition 4.8.1: Normal Vector

Let $\vec{n}$ be a nonzero vector in $\mathbb{R}^n$. Then $\vec{n}$ is called a **normal vector** to a plane if and only if

$$\vec{n} \bullet \vec{v} = 0$$

for every vector $\vec{v}$ in the plane.

In other words, we say that $\vec{n}$ is orthogonal (perpendicular) to every vector in the plane.

Consider now a plane with normal vector given by $\vec{n}$, and containing a point P_0. Notice that this plane is unique. If P is an arbitrary point on this plane, then by definition the normal vector is orthogonal to the vector between P_0 and P. Letting $\overrightarrow{0P}$ and $\overrightarrow{0P_0}$ be the position vectors of points P and P_0 respectively, it follows that

$$\vec{n} \bullet (\overrightarrow{0P} - \overrightarrow{0P_0}) = 0$$

or

$$\vec{n} \bullet \overrightarrow{P_0P} = 0$$

The first of these equations gives the **vector equation** of the plane.

Definition 4.8.2: Vector Equation of a Plane

Let $\vec{n}$ be the normal vector for a plane which contains a point P_0. If P is an arbitrary point on this plane, then the **vector equation** of the plane is given by

$$\vec{n} \bullet (\overrightarrow{0P} - \overrightarrow{0P_0}) = 0$$

Below is a video on determining the equation of a plane using a normal vector.



Notice that this equation can be used to determine if a point P is contained in a certain plane.

Example 4.8.1: A Point in a Plane

Let $\vec{n} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ be the normal vector for a plane which contains the point $P_0 = (2, 1, 4)$. Determine if the point $P = (5, 4, 1)$ is contained in this plane.

Solution

By Definition 4.8.2, P is a point in the plane if it satisfies the equation

$$\vec{n} \bullet (\overrightarrow{0P} - \overrightarrow{0P_0}) = 0$$

Given the above $\vec{n}$, P_0, and P, this equation becomes

$$\begin{aligned} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \bullet \left(\begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix} \right) &= \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \bullet \left(\begin{bmatrix} 3 \\ 3 \\ -3 \end{bmatrix} \right) \\ &= 3 + 6 - 9 = 0 \end{aligned}$$

Therefore $P = (5, 4, 1)$ is contained in the plane.

Suppose $\vec{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$, $P = (x, y, z)$ and $P_0 = (x_0, y_0, z_0)$.

Then

$$\begin{aligned} \vec{n} \bullet (\overrightarrow{0P} - \overrightarrow{0P_0}) &= 0 \\ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \bullet \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \right) &= 0 \\ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \bullet \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix} &= 0 \\ a(x - x_0) + b(y - y_0) + c(z - z_0) &= 0 \end{aligned}$$

We can also write this equation as

$$ax + by + cz = ax_0 + by_0 + cz_0$$

Notice that since P_0 is given, $ax_0+by_0+cz_0$ is a known scalar, which we can call d. This equation becomes

$$ax+by+cz=d$$

Definition 4.8.3: Scalar Equation of a Plane

Let $\vec{n}=\begin{bmatrix}a\\b\\c\end{bmatrix}$ be the normal vector for a plane which contains the point $P_0=(x_0,y_0,z_0)$.Then if $P=(x,y,z)$ is an arbitrary point on the plane, the **scalar equation** of the plane is given by

$$ax+by+cz=d$$

where $a,b,c,d\in\mathbb{R}$ and $d=ax_0+by_0+cz_0$.

Consider the following equation.

Example 4.8.2: Finding the Equation of a Plane

Find an equation of the plane containing $P_0=(3,-2,5)$ and orthogonal to $\vec{n}=\begin{bmatrix}-2\\4\\1\end{bmatrix}$.

Solution

The above vector $\vec{n}$ is the normal vector for this plane. Using Definition 4.8.2, we can determine the vector equation for this plane.

$$\begin{aligned}\vec{n}\bullet(\overrightarrow{0P}-\overrightarrow{0P_0})&=0\\ \begin{bmatrix}-2\\4\\1\end{bmatrix}\bullet\left(\begin{bmatrix}x\\y\\z\end{bmatrix}-\begin{bmatrix}3\\-2\\5\end{bmatrix}\right)&=0\\ \begin{bmatrix}-2\\4\\1\end{bmatrix}\bullet\begin{bmatrix}x-3\\y+2\\z-5\end{bmatrix}&=0\end{aligned}$$

Using Definition 4.8.3, we can determine the scalar equation of the plane.

$$-2x+4y+1z=-2(3)+4(-2)+1(5)=-9$$

Hence, the vector equation of the plane is

$$\begin{bmatrix}-2\\4\\1\end{bmatrix}\bullet\begin{bmatrix}x-3\\y+2\\z-5\end{bmatrix}=0$$

and the scalar equation is

$$-2x+4y+1z=-9$$

Below is a video on finding the equation of a plane given a point and a line.



Below is a video on finding the equation of a plane given three points on the plane.



Below is a video on finding the equation of the plane that is orthogonal to a given line.



Suppose a point P is not contained in a given plane. We are then interested in the shortest distance from that point P to the given plane. Consider the following example.

Example 4.8.3: Shortest Distance From a Point to a Plane

Find the shortest distance from the point $P = (3, 2, 3)$ to the plane given by $2x + y + 2z = 2$, and find the point Q on the plane that is closest to P.

Solution

Pick an arbitrary point P_0 on the plane. Then, it follows that

$$\overrightarrow{QP} = proj_{\vec{n}}\overrightarrow{P_0P}$$

and $\|\overrightarrow{QP}\|$ is the shortest distance from P to the plane. Further, the vector $\overrightarrow{0Q} = \overrightarrow{0P} - \overrightarrow{QP}$ gives the necessary point Q.

From the above scalar equation, we have that $\vec{n} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$. Now, choose $P_0 = (1,0,0)$ so that $\vec{n} \bullet \overrightarrow{0P} = 2 = d$. Then, $\overrightarrow{P_0P} = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$.

Next, compute $\overrightarrow{QP} = proj_{\vec{n}}\overrightarrow{P_0P}$.

$$\begin{aligned} \overrightarrow{QP} &= proj_{\vec{n}}\overrightarrow{P_0P} \\ &= \left(\frac{\overrightarrow{P_0P} \bullet \vec{n}}{\|\vec{n}\|^2} \right) \vec{n} \\ &= \frac{12}{9} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \\ &= \frac{4}{3} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \end{aligned}$$

Then, $\|\overrightarrow{QP}\| = 4$ so the shortest distance from P to the plane is 4.

Next, to find the point Q on the plane which is closest to P we have

$$\begin{aligned} \overrightarrow{0Q} &= \overrightarrow{0P} - \overrightarrow{QP} \\ &= \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix} - \frac{4}{3} \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} \\ &= \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \end{aligned}$$

Therefore, $Q = (\frac{1}{3}, \frac{2}{3}, \frac{1}{3})$.

Below is a video on determining the distance between a point and a plane.



This page titled 4.8: Planes in R^n is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.8: Planes in R^n** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.9: The Cross Product

Outcomes

A. Compute the cross product and box product of vectors in $\mathbb{R}^3$.

Recall that the dot product is one of two important products for vectors. The second type of product for vectors is called the cross product. It is important to note that the cross product is only defined in $\mathbb{R}^3$. First we discuss the geometric meaning and then a description in terms of coordinates is given, both of which are important. The geometric description is essential in order to understand the applications to physics and geometry while the coordinate description is necessary to compute the cross product.

Consider the following definition.

Definition 4.9.1: Right Hand System of Vectors

Three vectors, $\vec{u}, \vec{v}, \vec{w}$ form a right hand system if when you extend the fingers of your right hand along the direction of vector $\vec{u}$ and close them in the direction of $\vec{v}$, the thumb points roughly in the direction of $\vec{w}$.

For an example of a right handed system of vectors, see the following picture.

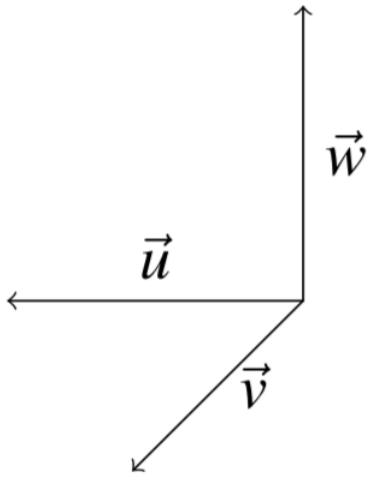


Figure 4.9.1

In this picture the vector $\vec{w}$ points upwards from the plane determined by the other two vectors. Point the fingers of your right hand along $\vec{u}$, and close them in the direction of $\vec{v}$. Notice that if you extend the thumb on your right hand, it points in the direction of $\vec{w}$.

You should consider how a right hand system would differ from a left hand system. Try using your left hand and you will see that the vector $\vec{w}$ would need to point in the opposite direction.

Notice that the special vectors, $\vec{i}, \vec{j}, \vec{k}$ will always form a right handed system. If you extend the fingers of your right hand along $\vec{i}$ and close them in the direction $\vec{j}$, the thumb points in the direction of $\vec{k}$.

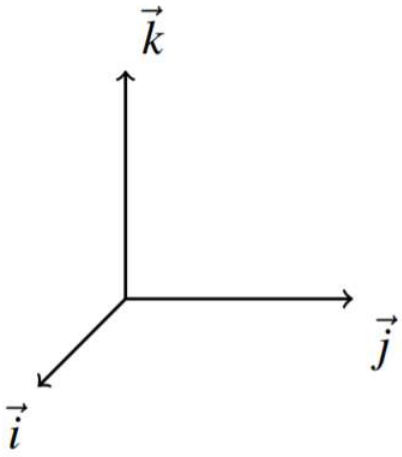


Figure 4.9.2

The following is the geometric description of the cross product. Recall that the dot product of two vectors results in a scalar. In contrast, the cross product results in a vector, as the product gives a direction as well as magnitude.

Definition 4.9.2: Geometric Definition of Cross Product

Let $\vec{u}$ and $\vec{v}$ be two vectors in $\mathbb{R}^3$. Then the **cross product**, written $\vec{u} \times \vec{v}$, is defined by the following two rules.

1. Its length is

$$\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \|\vec{v}\| \sin\theta,$$

   where θ is the included angle between $\vec{u}$ and $\vec{v}$.
2. It is perpendicular to both $\vec{u}$ and $\vec{v}$, that is

$$\begin{aligned} (\vec{u} \times \vec{v}) \cdot \vec{u} &= 0, \\ (\vec{u} \times \vec{v}) \cdot \vec{v} &= 0, \end{aligned} \tag{4.9.1}$$

   and

$$\vec{u}, \vec{v}, \vec{u} \times \vec{v}$$

   form a right hand system.

The cross product of the special vectors $\vec{i}, \vec{j}, \vec{k}$ is as follows.

$$\begin{aligned} \vec{i} \times \vec{j} &= \vec{k} \quad & \vec{j} \times \vec{i} &= -\vec{k} \\ \vec{k} \times \vec{i} &= \vec{j} \quad & \vec{i} \times \vec{k} &= -\vec{j} \\ \vec{j} \times \vec{k} &= \vec{i} \quad & \vec{k} \times \vec{j} &= -\vec{i} \end{aligned}$$

With this information, the following gives the coordinate description of the cross product.

Recall that the vector $\vec{u} = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T$ can be written in terms of $\vec{i}, \vec{j}, \vec{k}$ as $\vec{u} = u_1\vec{i} + u_2\vec{j} + u_3\vec{k}$.

Theorem 4.9.1: Coordinate Description of Cross Product

Let $\vec{u} = u_1\vec{i} + u_2\vec{j} + u_3\vec{k}$ and $\vec{v} = v_1\vec{i} + v_2\vec{j} + v_3\vec{k}$ be two vectors. Then

$$\vec{u} \times \vec{v} = (u_2v_3 - u_3v_2)\vec{i} - (u_1v_3 - u_3v_1)\vec{j} + (u_1v_2 - u_2v_1)\vec{k} \tag{4.9.2}$$

Writing $\vec{u} \times \vec{v}$ in the usual way, it is given by

$$\vec{u}\times\vec{v}=\begin{bmatrix} u_2v_3-u_3v_2 \\ -(u_1v_3-u_3v_1) \\ u_1v_2-u_2v_1 \end{bmatrix}$$

We now prove this proposition.

Proof

From the above table and the properties of the cross product listed,

$$\begin{aligned}\vec{u}\times\vec{v} &= \left(u_1\vec{i}+u_2\vec{j}+u_3\vec{k}\right)\times\left(v_1\vec{i}+v_2\vec{j}+v_3\vec{k}\right)\\ &= u_1v_2\vec{i}\times\vec{j}+u_1v_3\vec{i}\times\vec{k}+u_2v_1\vec{j}\times\vec{i}+u_2v_3\vec{j}\times\vec{k}++u_3v_1\vec{k}\times\vec{i}+u_3v_2\vec{k}\times\vec{j}\\ &= u_1v_2\vec{k}-u_1v_3\vec{j}-u_2v_1\vec{k}+u_2v_3\vec{i}+u_3v_1\vec{j}-u_3v_2\vec{i}\\ &= (u_2v_3-u_3v_2)\vec{i}+(u_3v_1-u_1v_3)\vec{j}+(u_1v_2-u_2v_1)\vec{k}\end{aligned} \tag{4.9.3}$$

There is another version of (4.9.2) which may be easier to remember. We can express the cross product as the determinant of a matrix, as follows.

$$\vec{u}\times\vec{v}=\begin{vmatrix}\vec{i} & \vec{j} & \vec{k}\\ u_1 & u_2 & u_3\\ v_1 & v_2 & v_3\end{vmatrix} \tag{4.9.4}$$

Expanding the determinant along the top row yields

$$\vec{i}(-1)^{1+1}\begin{vmatrix}u_2 & u_3\\ v_2 & v_3\end{vmatrix}+\vec{j}(-1)^{2+1}\begin{vmatrix}u_1 & u_3\\ v_1 & v_3\end{vmatrix}+\vec{k}(-1)^{3+1}\begin{vmatrix}u_1 & u_2\\ v_1 & v_2\end{vmatrix}$$

\

$$=\vec{i}\begin{vmatrix}u_2 & u_3\\ v_2 & v_3\end{vmatrix}-\vec{j}\begin{vmatrix}u_1 & u_3\\ v_1 & v_3\end{vmatrix}+\vec{k}\begin{vmatrix}u_1 & u_2\\ v_1 & v_2\end{vmatrix}$$

Expanding these determinants leads to

$$(u_2v_3-u_3v_2)\vec{i}-(u_1v_3-u_3v_1)\vec{j}+(u_1v_2-u_2v_1)\vec{k}$$

which is the same as (4.9.3).

Below is a video on vector cross products.



The cross product satisfies the following properties.

Proposition 4.9.1: Properties of the Cross Product

Let $\vec{u}, \vec{v}, \vec{w}$ be vectors in $\mathbb{R}^3$, and k a scalar. Then, the following properties of the cross product hold.

1. $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$, and $\vec{u} \times \vec{u} = \vec{0}$
2. $(k\vec{u}) \times \vec{v} = k(\vec{u} \times \vec{v}) = \vec{u} \times (k\vec{v})$
3. $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$
4. $(\vec{v} + \vec{w}) \times \vec{u} = \vec{v} \times \vec{u} + \vec{w} \times \vec{u}$

Proof

Formula 1. follows immediately from the definition. The vectors $\vec{u} \times \vec{v}$ and $\vec{v} \times \vec{u}$ have the same magnitude, $|\vec{u}|\,|\vec{v}| \sin\theta$, and an application of the right hand rule shows they have opposite direction.

Formula 2. is proven as follows. If k is a non-negative scalar, the direction of $(k\vec{u}) \times \vec{v}$ is the same as the direction of $\vec{u} \times \vec{v}, k(\vec{u} \times \vec{v})$ and $\vec{u} \times (k\vec{v})$. The magnitude is k times the magnitude of $\vec{u} \times \vec{v}$ which is the same as the magnitude of $k(\vec{u} \times \vec{v})$ and $\vec{u} \times (k\vec{v})$. Using this yields equality in 2. In the case where $k < 0$, everything works the same way except the vectors are all pointing in the opposite direction and you must multiply by $|k|$ when comparing their magnitudes.

The distributive laws, 3. and 4., are much harder to establish. For now, it suffices to notice that if we know that 3. is true, 4. follows. Thus, assuming 3., and using 1.,

$$\begin{aligned}(\vec{v} + \vec{w}) \times \vec{u} &= -\vec{u} \times (\vec{v} + \vec{w}) \\ &= -(\vec{u} \times \vec{v} + \vec{u} \times \vec{w}) \\ &= \vec{v} \times \vec{u} + \vec{w} \times \vec{u}\end{aligned}$$

Below is a video on properties of cross products.



Below is a video on the cross product of sums of vectors.



We will now look at an example of how to compute a cross product.

Example 4.9.1: Find a Cross Product

Find $\vec{u} \times \vec{v}$ for the following vectors

$$\vec{u} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \vec{v} = \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$$

Solution

Note that we can write $\vec{u}, \vec{v}$ in terms of the special vectors $\vec{i}, \vec{j}, \vec{k}$ as

$$\vec{u} = \vec{i} - \vec{j} + 2\vec{k}$$
$$\vec{v} = 3\vec{i} - 2\vec{j} + \vec{k}$$

We will use the equation given by (4.9.4) to compute the cross product.

$$\vec{u} \times \vec{v} = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & -1 & 2 \\ 3 & -2 & 1 \end{vmatrix} = \begin{vmatrix} -1 & 2 \\ -2 & 1 \end{vmatrix} \vec{i} - \begin{vmatrix} 1 & 2 \\ 3 & 1 \end{vmatrix} \vec{j} + \begin{vmatrix} 1 & -1 \\ 3 & -2 \end{vmatrix} \vec{k} = 3\vec{i} + 5\vec{j} + \vec{k}$$

We can write this result in the usual way, as

$$\vec{u} \times \vec{v} = \begin{bmatrix} 3 \\ 5 \\ 1 \end{bmatrix}$$

Below is a video on finding the cross product of two vectors.



An important geometrical application of the cross product is as follows. The size of the cross product, $\|\vec{u}\times\vec{v}\|$, is the area of the parallelogram determined by $\vec{u}$ and $\vec{v}$, as shown in the following picture.

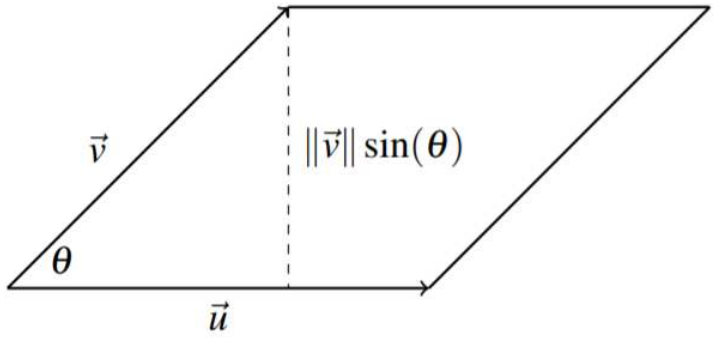


Figure 4.9.3

We examine this concept in the following example.

Example 4.9.2: Area of a Parallelogram

Find the area of the parallelogram determined by the vectors $\vec{u}$ and $\vec{v}$ given by

$$\vec{u}=\begin{bmatrix}1\\-1\\2\end{bmatrix},\vec{v}=\begin{bmatrix}3\\-2\\1\end{bmatrix}$$

Solution

Notice that these vectors are the same as the ones given in Example 4.9.1. Recall from the geometric description of the cross product, that the area of the parallelogram is simply the magnitude of $\vec{u}\times\vec{v}$. From Example 4.9.1, $\vec{u}\times\vec{v}=3\vec{i}+5\vec{j}+\vec{k}$. We can also write this as

$$\vec{u}\times\vec{v}=\begin{bmatrix}3\\5\\1\end{bmatrix}$$

Thus the area of the parallelogram is

$$\|\vec{u}\times\vec{v}\|=\sqrt{(3)(3)+(5)(5)+(1)(1)}=\sqrt{9+25+1}=\sqrt{35}$$

We can also use this concept to find the area of a triangle. Consider the following example.

Example 4.9.3: Area of Triangle

Find the area of the triangle determined by the points $(1,2,3)$, $(0,2,5)$, $(5,1,2)$

Solution

This triangle is obtained by connecting the three points with lines. Picking $(1,2,3)$ as a starting point, there are two displacement vectors, $\begin{bmatrix}-1 & 0 & 2\end{bmatrix}^T$ and $\begin{bmatrix}4 & -1 & -1\end{bmatrix}^T$. Notice that if we add either of these vectors to the position vector of the starting point, the result is the position vectors of the other two points. Now, the area of the triangle is half the area of the parallelogram determined by $\begin{bmatrix}-1 & 0 & 2\end{bmatrix}^T$ and $\begin{bmatrix}4 & -1 & -1\end{bmatrix}^T$. The required cross product is given by

$$\begin{bmatrix}-1\\0\\2\end{bmatrix}\times\begin{bmatrix}4\\-1\\-1\end{bmatrix}=\begin{bmatrix}2 & 7 & 1\end{bmatrix}$$

Taking the size of this vector gives the area of the parallelogram, given by

$$\sqrt{(2)(2)+(7)(7)+(1)(1)}=\sqrt{4+49+1}=\sqrt{54}$$

Hence the area of the triangle is $\frac{1}{2}\sqrt{54}=\frac{3}{2}\sqrt{6}$.

In general, if you have three points in $\mathbb{R}^3$, P, Q, R, the area of the triangle is given by

$$\frac{1}{2}\|\overrightarrow{PQ}\times\overrightarrow{PR}\|$$

Recall that $\overrightarrow{PQ}$ is the vector running from point P to point Q.

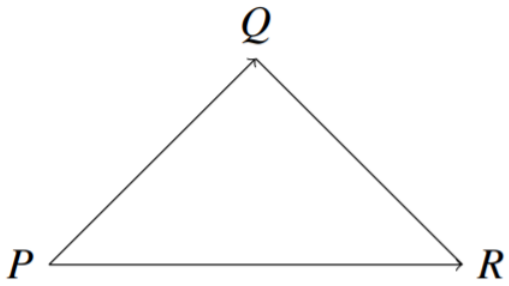


Figure 4.9.4

In the next section, we explore another application of the cross product.

The Box Product

Recall that we can use the cross product to find the the area of a parallelogram. It follows that we can use the cross product together with the dot product to find the volume of a parallelepiped. We begin with a definition.

Definition 4.9.3: Parallelepiped

A **parallelepiped** determined by the three vectors, $\vec{u}$, $\vec{v}$, and $\vec{w}$ consists of

$$\{r\vec{u}+s\vec{v}+t\vec{w} : r,s,t\in[0,1]\}$$

That is, if you pick three numbers, r, s, and t each in $[0,1]$ and form $r\vec{u}+s\vec{v}+t\vec{w}$ then the collection of all such points makes up the parallelepiped determined by these three vectors.

The following is an example of a parallelepiped.

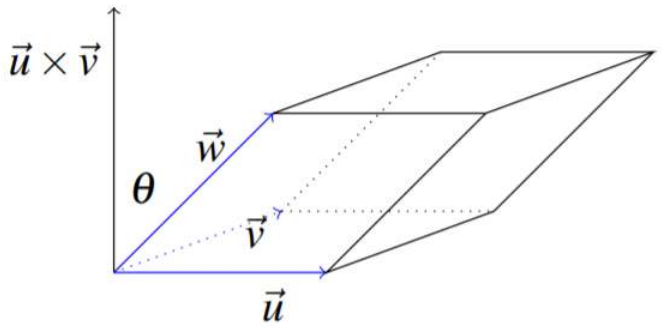


Figure 4.9.5

Notice that the base of the parallelepiped is the parallelogram determined by the vectors $\vec{u}$ and $\vec{v}$. Therefore, its area is equal to $\|\vec{u}\times\vec{v}\|$. The height of the parallelepiped is $\|\vec{w}\|\cos\theta$ where θ is the angle shown in the picture between $\vec{w}$ and $\vec{u}\times\vec{v}$. The volume of this parallelepiped is the area of the base times the height which is just

$$\|\vec{u}\times\vec{v}\|\|\vec{w}\|\cos\theta = (\vec{u}\times\vec{v})\cdot\vec{w}$$

This expression is known as the box product and is sometimes written as $[\vec{u},\vec{v},\vec{w}]$. You should consider what happens if you interchange the $\vec{v}$ with the $\vec{w}$ or the $\vec{u}$ with the $\vec{w}$. You can see geometrically from drawing pictures that this merely introduces a minus sign. In any case the box product of three vectors always equals either the volume of the parallelepiped determined by the three vectors or else -1 times this volume.

Proposition 4.9.2: The Box Product

Let $\vec{u},\vec{v},\vec{w}$ be three vectors in $\mathbb{R}^n$ that define a parallelepiped. Then the volume of the parallelepiped is the absolute value of the box product, given by

$$|(\vec{u}\times\vec{v})\cdot\vec{w}|$$

Consider an example of this concept.

Example 4.9.4: Volume of a Parallelepiped

Find the volume of the parallelepiped determined by the vectors

$$\vec{u} = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}, \vec{v} = \begin{bmatrix} 1 \\ 3 \\ -6 \end{bmatrix}, \vec{w} = \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}$$

Solution

According to the above discussion, pick any two of these vectors, take the cross product and then take the dot product of this with the third of these vectors. The result will be either the desired volume or -1 times the desired volume. Therefore by taking the absolute value of the result, we obtain the volume.

We will take the cross product of $\vec{u}$ and $\vec{v}$. This is given by

$$\vec{u}\times\vec{v} = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix} \times \begin{bmatrix} 1 \\ 3 \\ -6 \end{bmatrix}$$

$$= \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 2 & -5 \\ 1 & 3 & -6 \end{vmatrix} = 3\vec{i}+\vec{j}+\vec{k} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

Now take the dot product of this vector with $\vec{w}$ which yields

$$\begin{aligned}(\vec{u}\times\vec{v})\cdot\vec{w} &= \begin{bmatrix}3\\1\\1\end{bmatrix}\cdot\begin{bmatrix}3\\2\\3\end{bmatrix}\\ &= \left(3\vec{i}+\vec{j}+\vec{k}\right)\cdot\left(3\vec{i}+2\vec{j}+3\vec{k}\right)\\ &= 9+2+3\\ &= 14\end{aligned}$$

This shows the volume of this parallelepiped is 14 cubic units.

Below is a video on the triple scalar product.



There is a fundamental observation which comes directly from the geometric definitions of the cross product and the dot product.

Proposition 4.9.3: Order of the Product

Let $\vec{u},\vec{v}$, and $\vec{w}$ be vectors. Then $(\vec{u}\times\vec{v})\cdot\vec{w}=\vec{u}\cdot(\vec{v}\times\vec{w})$.

Proof

This follows from observing that either $(\vec{u}\times\vec{v})\cdot\vec{w}$ and $\vec{u}\cdot(\vec{v}\times\vec{w})$ both give the volume of the parallelepiped or they both give -1 times the volume.

Recall that we can express the cross product as the determinant of a particular matrix. It turns out that the same can be done for the box product. Suppose you have three vectors, $\vec{u}=[\,a\quad b\quad c\,]^T,\vec{v}=[\,d\quad e\quad f\,]^T$, and $\vec{w}=[\,g\quad h\quad i\,]^T$. Then the box product $\vec{u}\cdot(\vec{v}\times\vec{w})$ is given by the following.

$$\begin{aligned}\vec{u}\cdot(\vec{v}\times\vec{w}) &= \begin{bmatrix}a\\b\\c\end{bmatrix}\cdot\begin{vmatrix}\vec{i}&\vec{j}&\vec{k}\\d&e&f\\g&h&i\end{vmatrix}\\ &= a\begin{vmatrix}e&f\\h&i\end{vmatrix}-b\begin{vmatrix}d&f\\g&i\end{vmatrix}+c\begin{vmatrix}d&e\\g&h\end{vmatrix}\\ &= \det\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}\end{aligned}$$

To take the box product, you can simply take the determinant of the matrix which results by letting the rows be the components of the given vectors in the order in which they occur in the box product.

This follows directly from the definition of the cross product given above and the way we expand determinants. Thus the volume of a parallelepiped determined by the vectors $\vec{u},\vec{v},\vec{w}$ is just the absolute value of the above determinant.

This page titled 4.9: The Cross Product is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.9: The Cross Product** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.10: Spanning, Linear Independence and Basis in R^n

Outcomes

A. Determine the span of a set of vectors, and determine if a vector is contained in a specified span.
B. Determine if a set of vectors is linearly independent.
C. Understand the concepts of subspace, basis, and dimension.
D. Find the row space, column space, and null space of a matrix.

By generating all linear combinations of a set of vectors one can obtain various subsets of $\mathbb{R}^n$ which we call subspaces. For example what set of vectors in $\mathbb{R}^3$ generate the XY-plane? What is the smallest such set of vectors can you find? The tools of spanning, linear independence and basis are exactly what is needed to answer these and similar questions and are the focus of this section. The following definition is essential.

Definition 4.10.1: Subset

Let U and W be sets of vectors in $\mathbb{R}^n$. If all vectors in U are also in W, we say that U is a **subset** of W, denoted

$$U \subseteq W$$

Spanning Set of Vectors

We begin this section with a definition.

Definition 4.10.2: Span of a Set of Vectors

The collection of all linear combinations of a set of vectors $\{\vec{u}_1, \cdots, \vec{u}_k\}$ in $\mathbb{R}^n$ is known as the span of these vectors and is written as $\text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$.

Consider the following example.

Example 4.10.1: Span of Vectors

Describe the span of the vectors $\vec{u} = [\,1 \quad 1 \quad 0\,]^T$ and $\vec{v} = [\,3 \quad 2 \quad 0\,]^T \in \mathbb{R}^3$.

Solution

You can see that any linear combination of the vectors $\vec{u}$ and $\vec{v}$ yields a vector of the form $[\,x \quad y \quad 0\,]^T$ in the XY-plane.

Moreover every vector in the XY-plane is in fact such a linear combination of the vectors $\vec{u}$ and $\vec{v}$. That's because

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = (-2x+3y)\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + (x-y)\begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}$$

Thus $\text{span}\{\vec{u}, \vec{v}\}$ is precisely the XY-plane.

Below is a video on linear combinations and span of vectors.



You can convince yourself that no single vector can span the XY-plane. In fact, take a moment to consider what is meant by the span of a single vector.

However you can make the set larger if you wish. For example consider the larger set of vectors $\{\vec{u},\vec{v},\vec{w}\}$ where $\vec{w}=\begin{bmatrix}4 & 5 & 0\end{bmatrix}^T$. Since the first two vectors already span the entire XY-plane, the span is once again precisely the XY-plane and nothing has been gained. Of course if you add a new vector such as $\vec{w}=\begin{bmatrix}0 & 0 & 1\end{bmatrix}^T$ then it does span a different space. What is the span of $\vec{u},\vec{v},\vec{w}$ in this case?

Below is a video on determining if a vector is a linear combinations of a pair of vectors (Yes).



Below is a video on determining if a vector is a linear combinations of a pair of vectors (No).



The distinction between the sets $\{\vec{u},\vec{v}\}$ and $\{\vec{u},\vec{v},\vec{w}\}$ will be made using the concept of linear independence.

Consider the vectors $\vec{u}, \vec{v}$, and $\vec{w}$ discussed above. In the next example, we will show how to formally demonstrate that $\vec{w}$ is in the span of $\vec{u}$ and $\vec{v}$.

Example 4.10.2: Vector in a Span

Let $\vec{u} = [\,1 \quad 1 \quad 0\,]^T$ and $\vec{v} = [\,3 \quad 2 \quad 0\,]^T \in \mathbb{R}^3$. Show that $\vec{w} = [\,4 \quad 5 \quad 0\,]^T$ is in $\text{span}\{\vec{u}, \vec{v}\}$.

Solution

For a vector to be in $\text{span}\{\vec{u}, \vec{v}\}$, it must be a linear combination of these vectors. If $\vec{w} \in \text{span}\{\vec{u}, \vec{v}\}$, we must be able to find scalars a, b such that

$$\vec{w} = a\vec{u} + b\vec{v}$$

We proceed as follows.

$$\begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix} = a \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}$$

This is equivalent to the following system of equations

$$\begin{aligned} a + 3b &= 4 \\ a + 2b &= 5 \end{aligned}$$

We solving this system the usual way, constructing the augmented matrix and row reducing to find the reduced row-echelon form.

$$\left[\begin{array}{cc|c} 1 & 3 & 4 \\ 1 & 2 & 5 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 7 \\ 0 & 1 & -1 \end{array}\right]$$

The solution is $a = 7, b = -1$. This means that

$$\vec{w} = 7\vec{u} - \vec{v}$$

Therefore we can say that $\vec{w}$ is in $\text{span}\{\vec{u}, \vec{v}\}$.

Below is a video on determining which vectors from a list are in the span of two vectors.



Linearly Independent Set of Vectors

We now turn our attention to the following question: what linear combinations of a given set of vectors $\{\vec{u}_1, \cdots, \vec{u}_k\}$ in $\mathbb{R}^n$ yields the zero vector? Clearly $0\vec{u}_1 + 0\vec{u}_2 + \cdots + 0\vec{u}_k = \vec{0}$, but is it possible to have $\sum_{i=1}^{k} a_i\vec{u}_i = \vec{0}$ without all coefficients being zero?

You can create examples where this easily happens. For example if $\vec{u}_1 = \vec{u}_2$, then $1\vec{u}_1 - \vec{u}_2 + 0\vec{u}_3 + \cdots + 0\vec{u}_k = \vec{0}$, no matter the vectors $\{\vec{u}_3, \cdots, \vec{u}_k\}$. 0But sometimes it can be more subtle.

Example 4.10.3: Linearly Dependent Set of Vectors

Consider the vectors

$$\vec{u}_1 = \begin{bmatrix} 0 & 1 & -2 \end{bmatrix}^T, \vec{u}_2 = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T, \vec{u}_3 = \begin{bmatrix} -2 & 3 & 2 \end{bmatrix}^T, \text{ and } \vec{u}_4 = \begin{bmatrix} 1 & -2 & 0 \end{bmatrix}^T$$

in $\mathbb{R}^3$.

Then verify that

$$1\vec{u}_1 + 0\vec{u}_2 + -\vec{u}_3 - 2\vec{u}_4 = \vec{0}$$

You can see that the linear combination does yield the zero vector but has some non-zero coefficients. Thus we define a set of vectors to be *linearly dependent* if this happens.

Definition 4.10.3: Linearly Dependent Set of Vectors

A set of non-zero vectors $\{\vec{u}_1, \cdots, \vec{u}_k\}$ in $\mathbb{R}^n$ is said to be **linearly dependent** if a linear combination of these vectors without all coefficients being zero does yield the zero vector.

Note that if $\sum_{i=1}^k a_i \vec{u}_i = \vec{0}$ and some coefficient is non-zero, say $a_1 \neq 0$, then

$$\vec{u}_1 = \frac{-1}{a_1} \sum_{i=2}^k a_i \vec{u}_i$$

and thus $\vec{u}_1$ is in the span of the other vectors. And the converse clearly works as well, so we get that a set of vectors is linearly dependent precisely when one of its vector is in the span of the other vectors of that set.

In particular, you can show that the vector $\vec{u}_1$ in the above example is in the span of the vectors $\{\vec{u}_2, \vec{u}_3, \vec{u}_4\}$.

If a set of vectors is NOT linearly dependent, then it must be that any linear combination of these vectors which yields the zero vector must use all zero coefficients. This is a very important notion, and we give it its own name of *linear independence*.

Definition 4.10.4: Linearly Independent Set of Vectors

A set of non-zero vectors $\{\vec{u}_1, \cdots, \vec{u}_k\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if whenever

$$\sum_{i=1}^k a_i \vec{u}_i = \vec{0}$$

it follows that each $a_i = 0$.

Note also that we require all vectors to be non-zero to form a linearly independent set.

Below is a video on linearly independent and dependent sets of vectors.



To view this in a more familiar setting, form the $n \times k$ matrix A having these vectors as columns. Then all we are saying is that the set $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is linearly independent precisely when $AX = 0$ has only the trivial solution.

Here is an example.

Example 4.10.4: Linearly Independent Vectors

Consider the vectors $\vec{u} = [1 \quad 1 \quad 0]^T$, $\vec{v} = [1 \quad 0 \quad 1]^T$, and $\vec{w} = [0 \quad 1 \quad 1]^T$ in $\mathbb{R}^3$. Verify whether the set $\{\vec{u}, \vec{v}, \vec{w}\}$ is linearly independent.

Solution

So suppose that we have a linear combinations $a\vec{u} + b\vec{v} + c\vec{w} = \vec{0}$. Then you can see that this can only happen with $a = b = c = 0$.

As mentioned above, you can equivalently form the 3×3 matrix $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$, and show that $AX = 0$ has only the trivial solution.

Thus this means the set $\{\vec{u}, \vec{v}, \vec{w}\}$ is linearly independent.

Below is a video on determining if vectors are linearly independent.



Below is another video on determining if vectors are linearly independent.



In terms of spanning, a set of vectors is linearly independent if it does not contain unnecessary vectors, that is not vector is in the span of the others.

Thus we put all this together in the following important theorem.

Theorem 4.10.1: Linear Independence as a Linear Combination

Let $\{\vec{u}_1, \cdots, \vec{u}_k\}$ be a collection of vectors in $\mathbb{R}^n$. Then the following are equivalent:

1. It is linearly independent, that is whenever
$$\sum_{i=1}^{k} a_i \vec{u}_i = \vec{0}$$
it follows that each coefficient $a_i = 0$.
2. No vector is in the span of the others.
3. The system of linear equations $AX = 0$ has only the trivial solution, where A is the $n \times k$ matrix having these vectors as columns.

The last sentence of this theorem is useful as it allows us to use the reduced row-echelon form of a matrix to determine if a set of vectors is linearly independent. Let the vectors be columns of a matrix A. Find the reduced row-echelon form of A. If each column has a leading one, then it follows that the vectors are linearly independent.

Sometimes we refer to the condition regarding sums as follows: The set of vectors, $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is linearly independent if and only if there is no nontrivial linear combination which equals the zero vector. A nontrivial linear combination is one in which not all the scalars equal zero. Similarly, a trivial linear combination is one in which all scalars equal zero.

Below is a video on determining if the columns of a matrix are linearly independent.



Here is a detailed example in $\mathbb{R}^4$.

Example 4.10.5: Linear Independence

Determine whether the set of vectors given by

$$\left\{\begin{bmatrix}1\\2\\3\\0\end{bmatrix},\begin{bmatrix}2\\1\\0\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\\2\end{bmatrix},\begin{bmatrix}3\\2\\2\\0\end{bmatrix}\right\}$$

is linearly independent. If it is linearly dependent, express one of the vectors as a linear combination of the others.

Solution

In this case the matrix of the corresponding homogeneous system of linear equations is

$$\left[\begin{array}{cccc|c}1&2&0&3&0\\2&1&1&2&0\\3&0&1&2&0\\0&1&2&0&0\end{array}\right]$$

The reduced row-echelon form is

$$\left[\begin{array}{cccc|c}1&0&0&0&0\\0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\end{array}\right]$$

and so every column is a pivot column and the corresponding system $AX=0$ only has the trivial solution. Therefore, these vectors are linearly independent and there is no way to obtain one of the vectors as a linear combination of the others.

Below is a video on determining a third vector that makes a set dependent (independent).



Consider another example.

Example 4.10.6: Linear Independence

Determine whether the set of vectors given by

$$\left\{\begin{bmatrix}1\\2\\3\\0\end{bmatrix},\begin{bmatrix}2\\1\\0\\1\end{bmatrix},\begin{bmatrix}0\\1\\1\\2\end{bmatrix},\begin{bmatrix}3\\2\\2\\-1\end{bmatrix}\right\}$$

is linearly independent. If it is linearly dependent, express one of the vectors as a linear combination of the others.

Solution

Form the 4×4 matrix A having these vectors as columns:

$$A = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 1 & 1 & 2 \\ 3 & 0 & 1 & 2 \\ 0 & 1 & 2 & -1 \end{bmatrix}$$

Then by Theorem 4.10.1, the given set of vectors is linearly independent exactly if the system $AX = 0$ has only the trivial solution.

The augmented matrix for this system and corresponding reduced row-echelon form are given by

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 3 & 0 \\ 2 & 1 & 1 & 2 & 0 \\ 3 & 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & -1 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Not all the columns of the coefficient matrix are pivot columns and so the vectors are not linearly independent. In this case, we say the vectors are linearly dependent.

It follows that there are infinitely many solutions to $AX = 0$, one of which is

$$\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

Therefore we can write

$$1\begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix} - 1\begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} - 1\begin{bmatrix} 3 \\ 2 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

This can be rearranged as follows

$$1\begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix} - 1\begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 2 \\ -1 \end{bmatrix}$$

This gives the last vector as a linear combination of the first three vectors.

Notice that we could rearrange this equation to write any of the four vectors as a linear combination of the other three.

When given a linearly independent set of vectors, we can determine if related sets are linearly independent.

Example 4.10.7: Related Sets of Vectors

Let $\{\vec{u}, \vec{v}, \vec{w}\}$ be an independent set of $\mathbb{R}^n$. Is $\{\vec{u}+\vec{v}, 2\vec{u}+\vec{w}, \vec{v}-5\vec{w}\}$ linearly independent?

Solution

Suppose $a(\vec{u}+\vec{v}) + b(2\vec{u}+\vec{w}) + c(\vec{v}-5\vec{w}) = \vec{0}_n$ for some $a, b, c \in \mathbb{R}$. Then

$$(a+2b)\vec{u} + (a+c)\vec{v} + (b-5c)\vec{w} = \vec{0}_n.$$

Since $\{\vec{u}, \vec{v}, \vec{w}\}$ is independent,

$$\begin{aligned} a+2b &= 0 \\ a+c &= 0 \\ b-5c &= 0 \end{aligned}$$

This system of three equations in three variables has the unique solution $a=b=c=0$. Therefore, $\{\vec{u}+\vec{v}, 2\vec{u}+\vec{w}, \vec{v}-5\vec{w}\}$ is independent.

The following corollary follows from the fact that if the augmented matrix of a homogeneous system of linear equations has more columns than rows, the system has infinitely many solutions.

Corollary 4.10.1: Linear Dependence in $\mathbb{R}''$

Let $\{\vec{u}_1, \cdots, \vec{u}_k\}$ be a set of vectors in $\mathbb{R}^n$. If $k > n$, then the set is linearly dependent (i.e. NOT linearly independent).

Proof

Form the $n \times k$ matrix A having the vectors $\{\vec{u}_1, \cdots, \vec{u}_k\}$ as its columns and suppose $k > n$. Then A has rank $r \leq n < k$, so the system $AX = 0$ has a nontrivial solution and thus not linearly independent by Theorem 4.10.1.

Example 4.10.8: Linear Dependence

Consider the vectors

$$\left\{ \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\}$$

Are these vectors linearly independent?

Solution

This set contains three vectors in $\mathbb{R}^2$. By Corollary 4.10.1 these vectors are linearly dependent. In fact, we can write

$$(-1) \begin{bmatrix} 1 \\ 4 \end{bmatrix} + (2) \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

showing that this set is linearly dependent.

The third vector in the previous example is in the span of the first two vectors. We could find a way to write this vector as a linear combination of the other two vectors. It turns out that the linear combination which we found is the **only** one, provided that the set is linearly independent.

Theorem 4.10.2: Unique Linear Combination

Let $U \subseteq \mathbb{R}^n$ be an independent set. Then any vector $\vec{x} \in \text{span}(U)$ can be written uniquely as a linear combination of vectors of U.

Proof

To prove this theorem, we will show that two linear combinations of vectors in U that equal $\vec{x}$ must be the same. Let $U = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$. Suppose that there is a vector $\vec{x} \in \text{span}(U)$ such that

$$\begin{aligned} \vec{x} &= s_1\vec{u}_1 + s_2\vec{u}_2 + \cdots + s_k\vec{u}_k, \text{ for some } s_1, s_2, \ldots, s_k \in \mathbb{R}, \text{ and} \\ \vec{x} &= t_1\vec{u}_1 + t_2\vec{u}_2 + \cdots + t_k\vec{u}_k, \text{ for some } t_1, t_2, \ldots, t_k \in \mathbb{R}. \end{aligned}$$

Then $\vec{0}_n = \vec{x} - \vec{x} = (s_1 - t_1)\vec{u}_1 + (s_2 - t_2)\vec{u}_2 + \cdots + (s_k - t_k)\vec{u}_k$.

Since U is independent, the only linear combination that vanishes is the trivial one, so $s_i - t_i = 0$ for all i, $1 \leq i \leq k$.

Therefore, $s_i = t_i$ for all i, $1 \leq i \leq k$, and the representation is unique.Let $U \subseteq \mathbb{R}^n$ be an independent set. Then any vector $\vec{x} \in \text{span}(U)$ can be written uniquely as a linear combination of vectors of U.

Suppose that $\vec{u}, \vec{v}$ and $\vec{w}$ are nonzero vectors in $\mathbb{R}^3$, and that $\{\vec{v}, \vec{w}\}$ is independent. Consider the set $\{\vec{u}, \vec{v}, \vec{w}\}$. When can we know that this set is independent? It turns out that this follows exactly when $\vec{u} \notin \text{span}\{\vec{v}, \vec{w}\}$.

Example 4.10.9

Suppose that $\vec{u}, \vec{v}$ and $\vec{w}$ are nonzero vectors in $\mathbb{R}^3$, and that $\{\vec{v}, \vec{w}\}$ is independent. Prove that $\{\vec{u}, \vec{v}, \vec{w}\}$ is independent if and only if $\vec{u} \notin \text{span}\{\vec{v}, \vec{w}\}$.

Solution

If $\vec{u} \in \text{span}\{\vec{v}, \vec{w}\}$, then there exist $a, b \in \mathbb{R}$ so that $\vec{u} = a\vec{v} + b\vec{w}$. This implies that $\vec{u} - a\vec{v} - b\vec{w} = \vec{0}_3$, so $\vec{u} - a\vec{v} - b\vec{w}$ is a nontrivial linear combination of $\{\vec{u}, \vec{v}, \vec{w}\}$ that vanishes, and thus $\{\vec{u}, \vec{v}, \vec{w}\}$ is dependent.

Now suppose that $\vec{u} \notin \text{span}\{\vec{v}, \vec{w}\}$, and suppose that there exist $a, b, c \in \mathbb{R}$ such that $a\vec{u} + b\vec{v} + c\vec{w} = \vec{0}_3$. If $a \neq 0$, then $\vec{u} = -\frac{b}{a}\vec{v} - \frac{c}{a}\vec{w}$, and $\vec{u} \in \text{span}\{\vec{v}, \vec{w}\}$, a contradiction. Therefore, $a = 0$, implying that $b\vec{v} + c\vec{w} = \vec{0}_3$. Since $\{\vec{v}, \vec{w}\}$ is independent, $b = c = 0$, and thus $a = b = c = 0$, i.e., the only linear combination of $\vec{u}, \vec{v}$ and $\vec{w}$ that vanishes is the trivial one.

Therefore, $\{\vec{u}, \vec{v}, \vec{w}\}$ is independent.

Consider the following useful theorem.

Theorem 4.10.3: Invertible Matrices

Let A be an invertible $n \times n$ matrix. Then the columns of A are independent and span $\mathbb{R}^n$. Similarly, the rows of A are independent and span the set of all $1 \times n$ vectors.

This theorem also allows us to determine if a matrix is invertible. If an $n \times n$ matrix A has columns which are independent, or span $\mathbb{R}^n$, then it follows that A is invertible. If it has rows that are independent, or span the set of all $1 \times n$ vectors, then A is invertible.

A Short Application to Chemistry

The following section applies the concepts of spanning and linear independence to the subject of chemistry.

When working with chemical reactions, there are sometimes a large number of reactions and some are in a sense redundant. Suppose you have the following chemical reactions.

$$CO + \tfrac{1}{2}O_2 \to CO_2$$
$$H_2 + \tfrac{1}{2}O_2 \to H_2O$$
$$CH_4 + \tfrac{3}{2}O_2 \to CO + 2H_2O$$
$$CH_4 + 2O_2 \to CO_2 + 2H_2O$$

There are four chemical reactions here but they are not independent reactions. There is some redundancy. What are the independent reactions? Is there a way to consider a shorter list of reactions? To analyze this situation, we can write the reactions in a matrix as follows

$$\begin{bmatrix} CO & O_2 & CO_2 & H_2 & H_2O & CH_4 \\ 1 & 1/2 & -1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1 & -1 & 0 \\ -1 & 3/2 & 0 & 0 & -2 & 1 \\ 0 & 2 & -1 & 0 & -2 & 1 \end{bmatrix}$$

Each row contains the coefficients of the respective elements in each reaction. For example, the top row of numbers comes from $CO + \frac{1}{2}O_2 - CO_2 = 0$ which represents the first of the chemical reactions.

We can write these coefficients in the following matrix

$$\begin{bmatrix} 1 & 1/2 & -1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1 & -1 & 0 \\ -1 & 3/2 & 0 & 0 & -2 & 1 \\ 0 & 2 & -1 & 0 & -2 & 1 \end{bmatrix}$$

Rather than listing all of the reactions as above, it would be more efficient to only list those which are independent by throwing out that which is redundant. We can use the concepts of the previous section to accomplish this.

First, take the reduced row-echelon form of the above matrix.

$$\begin{bmatrix} 1 & 0 & 0 & 3 & -1 & -1 \\ 0 & 1 & 0 & 2 & -2 & 0 \\ 0 & 0 & 1 & 4 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The top three rows represent “independent" reactions which come from the original four reactions. One can obtain each of the original four rows of the matrix given above by taking a suitable linear combination of rows of this reduced row-echelon matrix.

With the redundant reaction removed, we can consider the simplified reactions as the following equations

$$CO + 3H_2 - 1H_2O - 1CH_4 = 0$$
$$O_2 + 2H_2 - 2H_2O = 0$$
$$CO_2 + 4H_2 - 2H_2O - 1CH_4 = 0$$

In terms of the original notation, these are the reactions

$$CO + 3H_2 \rightarrow H_2O + CH_4$$
$$O_2 + 2H_2 \rightarrow 2H_2O$$
$$CO_2 + 4H_2 \rightarrow 2H_2O + CH_4$$

These three reactions provide an equivalent system to the original four equations. The idea is that, in terms of what happens chemically, you obtain the same information with the shorter list of reactions. Such a simplification is especially useful when dealing with very large lists of reactions which may result from experimental evidence.

Subspaces and Basis

The goal of this section is to develop an understanding of a subspace of $\mathbb{R}^n$. Before a precise definition is considered, we first examine the subspace test given below.

Theorem 4.10.4: Subspace Test

A subset V of $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ if

1. the zero vector of $\mathbb{R}^n$, $\vec{0}_n$, is in V;
2. V is closed under addition, i.e., for all $\vec{u}, \vec{w} \in V$, $\vec{u} + \vec{w} \in V$;
3. V is closed under scalar multiplication, i.e., for all $\vec{u} \in V$ and $k \in \mathbb{R}$, $k\vec{u} \in V$.

This test allows us to determine if a given set is a subspace of $\mathbb{R}^n$. Notice that the subset $V = \left\{\vec{0}\right\}$ is a subspace of $\mathbb{R}^n$ (called the zero subspace), as is $\mathbb{R}^n$ itself. A subspace which is not the zero subspace of $\mathbb{R}^n$ is referred to as a proper subspace.

A subspace is simply a set of vectors with the property that linear combinations of these vectors remain in the set. Geometrically in $\mathbb{R}^3$, it turns out that a subspace can be represented by either the origin as a single point, lines and planes which contain the origin, or the entire space $\mathbb{R}^3$.

Consider the following example of a line in $\mathbb{R}^3$.

Example 4.10.10: Subspace of $\mathbb{R}^3$

In $\mathbb{R}^3$, the line L through the origin that is parallel to the vector $\vec{d} = \begin{bmatrix} -5 \\ 1 \\ -4 \end{bmatrix}$ has (vector) equation $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -5 \\ 1 \\ -4 \end{bmatrix}, t \in \mathbb{R}$, so

$$L = \left\{ t\vec{d} \mid t \in \mathbb{R} \right\}.$$

Then L is a subspace of $\mathbb{R}^3$.

Solution

Using the subspace test given above we can verify that L is a subspace of $\mathbb{R}^3$.

- First: $\vec{0}_3 \in L$ since $0\vec{d} = \vec{0}_3$.
- Suppose $\vec{u}, \vec{v} \in L$. Then by definition, $\vec{u} = s\vec{d}$ and $\vec{v} = t\vec{d}$, for some $s, t \in \mathbb{R}$. Thus

  $$\vec{u} + \vec{v} = s\vec{d} + t\vec{d} = (s+t)\vec{d}.$$

  Since $s + t \in \mathbb{R}$, $\vec{u} + \vec{v} \in L$; i.e., L is closed under addition.
- Suppose $\vec{u} \in L$ and $k \in \mathbb{R}$ (k is a scalar). Then $\vec{u} = t\vec{d}$, for some $t \in \mathbb{R}$, so

  $$k\vec{u} = k(t\vec{d}) = (kt)\vec{d}.$$

  Since $kt \in \mathbb{R}$, $k\vec{u} \in L$; i.e., L is closed under scalar multiplication.

Since L satisfies all conditions of the subspace test, it follows that L is a subspace.

Note that there is nothing special about the vector $\vec{d}$ used in this example; the same proof works for any nonzero vector $\vec{d} \in \mathbb{R}^3$, so any line through the origin is a subspace of $\mathbb{R}^3$.

We are now prepared to examine the precise definition of a subspace as follows.

Definition 4.10.5: Subspace

Let V be a nonempty collection of vectors in $\mathbb{R}^n$. Then V is called a subspace if whenever a and b are scalars and $\vec{u}$ and $\vec{v}$ are vectors in V, the linear combination $a\vec{u} + b\vec{v}$ is also in V.

More generally this means that a subspace contains the span of any finite collection vectors in that subspace. It turns out that in $\mathbb{R}^n$, a subspace is exactly the span of finitely many of its vectors.

Theorem 4.10.5: Subspaces are Spans

Let V be a nonempty collection of vectors in $\mathbb{R}^n$. Then V is a subspace of $\mathbb{R}^n$ if and only if there exist vectors $\{\vec{u}_1, \cdots, \vec{u}_k\}$ in V such that

$$V = \text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$$

Furthermore, let W be another subspace of $\mathbb{R}^n$ and suppose $\{\vec{u}_1, \cdots, \vec{u}_k\} \in W$. Then it follows that V is a subset of W.

Note that since W is arbitrary, the statement that $V \subseteq W$ means that any other subspace of $\mathbb{R}^n$ that contains these vectors will also contain V.

Proof

We first show that if V is a subspace, then it can be written as $V = \text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$. Pick a vector $\vec{u}_1$ in V. If $V = \text{span}\{\vec{u}_1\}$, then you have found your list of vectors and are done. If $V \neq \text{span}\{\vec{u}_1\}$, then there exists $\vec{u}_2$ a vector of V which is not in $\text{span}\{\vec{u}_1\}$. Consider $\text{span}\{\vec{u}_1, \vec{u}_2\}$. If $V = \text{span}\{\vec{u}_1, \vec{u}_2\}$, we are done. Otherwise, pick $\vec{u}_3$ not in $\text{span}\{\vec{u}_1, \vec{u}_2\}$. Continue this way. Note that since V is a subspace, these spans are each contained in V. The process must stop with $\vec{u}_k$ for some $k \leq n$ by Corollary 4.10.1, and thus $V = \text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$.

Now suppose $V = \text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$, we must show this is a subspace. So let $\sum_{i=1}^{k} c_i \vec{u}_i$ and $\sum_{i=1}^{k} d_i \vec{u}_i$ be two vectors in V, and let a and b be two scalars. Then

$$a \sum_{i=1}^{k} c_i \vec{u}_i + b \sum_{i=1}^{k} d_i \vec{u}_i = \sum_{i=1}^{k} (ac_i + bd_i) \vec{u}_i$$

which is one of the vectors in $\text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$ and is therefore contained in V. This shows that $\text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$ has the properties of a subspace.

To prove that $V \subseteq W$, we prove that if $\vec{u}_i \in V$, then $\vec{u}_i \in W$.

Suppose $\vec{u} \in V$. Then $\vec{u} = a_1 \vec{u}_1 + a_2 \vec{u}_2 + \cdots + a_k \vec{u}_k$ for some $a_i \in \mathbb{R}$, $1 \leq i \leq k$. Since W contain each $\vec{u}_i$ and W is a vector space, it follows that $a_1 \vec{u}_1 + a_2 \vec{u}_2 + \cdots + a_k \vec{u}_k \in W$.

Since the vectors $\vec{u}_i$ we constructed in the proof above are not in the span of the previous vectors (by definition), they must be linearly independent and thus we obtain the following corollary.

Corollary 4.10.2: Subspaces are Spans of Independent Vectors

If V is a subspace of $\mathbb{R}^n$, then there exist linearly independent vectors $\{\vec{u}_1, \cdots, \vec{u}_k\}$ in V such that $V = \text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$.

In summary, subspaces of $\mathbb{R}^n$ consist of spans of finite, linearly independent collections of vectors of $\mathbb{R}^n$. Such a collection of vectors is called a basis.

Definition 4.10.6: Basis of a Subspace

Let V be a subspace of $\mathbb{R}^n$. Then $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is a **basis** for V if the following two conditions hold.

1. $\text{span}\{\vec{u}_1, \cdots, \vec{u}_k\} = V$
2. $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is linearly independent

Note the plural of basis is **bases**.

The following is a simple but very useful example of a basis, called the standard basis.

Definition 4.10.7: Standard Basis of $\mathbb{R}^n$

Let $\vec{e}_i$ be the vector in $\mathbb{R}^n$ which has a 1 in the ith entry and zeros elsewhere, that is the ith column of the identity matrix. Then the collection $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$ is a basis for $\mathbb{R}^n$ and is called the standard basis of $\mathbb{R}^n$.

Below is a video on determining if vectors form a basis for R^3.



The main theorem about bases is not only they exist, but that they must be of the same size. To show this, we will need the the following fundamental result, called the Exchange Theorem.

Theorem 4.10.6: Exchange Theorem

Suppose $\{\vec{u}_1, \cdots, \vec{u}_r\}$ is a linearly independent set of vectors in $\mathbb{R}^n$, and each $\vec{u}_k$ is contained in $\text{span}\{\vec{v}_1, \cdots, \vec{v}_s\}$ Then $s \geq r$.
In words, spanning sets have at least as many vectors as linearly independent sets.

Proof

Since each $\vec{u}_j$ is in $\text{span}\{\vec{v}_1, \cdots, \vec{v}_s\}$, there exist scalars a_{ij} such that

$$\vec{u}_j = \sum_{i=1}^{s} a_{ij}\vec{v}_i$$

Suppose for a contradiction that $s < r$. Then the matrix $A = [a_{ij}]$ has fewer rows, s than columns, r. Then the system $AX = 0$ has a non trivial solution $\vec{d}$, that is there is a $\vec{d} \neq \vec{0}$ such that $A\vec{d} = \vec{0}$. In other words,

$$\sum_{j=1}^{r} a_{ij}d_j = 0, \ i = 1, 2, \cdots, s$$

Therefore,

$$\begin{aligned} \sum_{j=1}^{r} d_j\vec{u}_j &= \sum_{j=1}^{r} d_j \sum_{i=1}^{s} a_{ij}\vec{v}_i \\ &= \sum_{i=1}^{s} \left(\sum_{j=1}^{r} a_{ij}d_j \right) \vec{v}_i = \sum_{i=1}^{s} 0\vec{v}_i = 0 \end{aligned}$$

which contradicts the assumption that $\{\vec{u}_1, \cdots, \vec{u}_r\}$ is linearly independent, because not all the d_j are zero. Thus this contradiction indicates that $s \geq r$.

We are now ready to show that any two bases are of the same size.

Theorem 4.10.7: Bases of $\mathbb{R}^n$ are of the Same Size

Let V be a subspace of $\mathbb{R}^n$ with two bases B_1 and B_2. Suppose B_1 contains s vectors and B_2 contains r vectors. Then $s = r$.

Proof

This follows right away from Theorem 9.4.4. Indeed observe that $B_1 = \{\vec{u}_1, \cdots, \vec{u}_s\}$ is a spanning set for V while $B_2 = \{\vec{v}_1, \cdots, \vec{v}_r\}$ is linearly independent, so $s \geq r$. Similarly $B_2 = \{\vec{v}_1, \cdots, \vec{v}_r\}$ is a spanning set for V while $B_1 = \{\vec{u}_1, \cdots, \vec{u}_s\}$ is linearly independent, so $r \geq s$.

Below is another video on determining if vectors form a basis for R^3.



The following definition can now be stated.

Definition 4.10.8: Dimension of a Subspace

Let V be a subspace of $\mathbb{R}^n$. Then the **dimension** of V, written $\dim(V)$ is defined to be the number of vectors in a basis.

The next result follows.

Corollary 4.10.3: Dimension of $\mathbb{R}^n$

The dimension of $\mathbb{R}^n$ is n.

Proof

You only need to exhibit a basis for $\mathbb{R}^n$ which has n vectors. Such a basis is the standard basis $\{\vec{e}_1, \cdots, \vec{e}_n\}$.

Consider the following example.

Example 4.10.11: Basis of Subspace

Let

$$V = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 \ : \ a - b = d - c \right\}.$$

Show that V is a subspace of $\mathbb{R}^4$, find a basis of V, and find $\dim(V)$.

Solution

The condition $a - b = d - c$ is equivalent to the condition $a = b - c + d$, so we may write

$$V = \left\{ \begin{bmatrix} b - c + d \\ b \\ c \\ d \end{bmatrix} \ : \ b, c, d \in \mathbb{R} \right\} = \left\{ b \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + d \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \ : \ b, c, d \in \mathbb{R} \right\}$$

This shows that V is a subspace of $\mathbb{R}^4$, since $V = \text{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ where

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Furthermore,

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is linearly independent, as can be seen by taking the reduced row-echelon form of the matrix whose columns are $\vec{u}_1, \vec{u}_2$ and $\vec{u}_3$.

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Since every column of the reduced row-echelon form matrix has a leading one, the columns are linearly independent.

Therefore $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is linearly independent and spans V, so is a basis of V. Hence V has dimension three.

We continue by stating further properties of a set of vectors in $\mathbb{R}^n$.

Corollary 4.10.4: Linearly Independent and Spanning Sets in $\mathbb{R}^n$

The following properties hold in $\mathbb{R}^n$:

- Suppose $\{\vec{u}_1, \cdots, \vec{u}_n\}$ is linearly independent. Then $\{\vec{u}_1, \cdots, \vec{u}_n\}$ is a basis for $\mathbb{R}^n$.
- Suppose $\{\vec{u}_1, \cdots, \vec{u}_m\}$ spans $\mathbb{R}^n$. Then $m \geq n$.
- If $\{\vec{u}_1, \cdots, \vec{u}_n\}$ spans $\mathbb{R}^n$, then $\{\vec{u}_1, \cdots, \vec{u}_n\}$ is linearly independent.

Proof

Assume first that $\{\vec{u}_1, \cdots, \vec{u}_n\}$ is linearly independent, and we need to show that this set spans $\mathbb{R}^n$. To do so, let $\vec{v}$ be a vector of $\mathbb{R}^n$, and we need to write $\vec{v}$ as a linear combination of $\vec{u}_i$'s. Consider the matrix A having the vectors $\vec{u}_i$ as columns:

$$A = \begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix}$$

By linear independence of the $\vec{u}_i$'s, the reduced row-echelon form of A is the identity matrix. Therefore the system $A\vec{x} = \vec{v}$ has a (unique) solution, so $\vec{v}$ is a linear combination of the $\vec{u}_i$'s.

To establish the second claim, suppose that $m < n$. Then letting $\vec{u}_{i_1}, \cdots, \vec{u}_{i_k}$ be the pivot columns of the matrix

$$\begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{bmatrix}$$

it follows $k \leq m < n$ and these k pivot columns would be a basis for $\mathbb{R}^n$ having fewer than n vectors, contrary to Corollary 4.10.3

Finally consider the third claim. If $\{\vec{u}_1, \cdots, \vec{u}_n\}$ is not linearly independent, then replace this list with $\{\vec{u}_{i_1}, \cdots, \vec{u}_{i_k}\}$ where these are the pivot columns of the matrix

$$\begin{bmatrix} \vec{u}_1 & \cdots & \vec{u}_n \end{bmatrix}$$

Then $\{\vec{u}_{i_1}, \cdots, \vec{u}_{i_k}\}$ spans $\mathbb{R}^n$ and is linearly independent, so it is a basis having less than n vectors again contrary to Corollary 4.10.3

The next theorem follows from the above claim.

Theorem 4.10.8: Existence of Basis

Let V be a subspace of $\mathbb{R}^n$. Then there exists a basis of V with $\dim(V) \leq n$.

Consider Corollary 4.10.4 together with Theorem 4.10.8. Let $\dim(V) = r$. Suppose there exists an independent set of vectors in V. If this set contains r vectors, then it is a basis for V. If it contains less than r vectors, then vectors can be added to the set to create a basis of V. Similarly, any spanning set of V which contains more than r vectors can have vectors removed to create a basis of V.

We illustrate this concept in the next example.

Example 4.10.12: Extending an Independent Set

Consider the set U given by

$$U = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; a - b = d - c \right\}$$

Then U is a subspace of $\mathbb{R}^4$ and $\dim(U) = 3$.

Then

$$S = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 3 \\ 2 \end{bmatrix} \right\},$$

is an independent subset of U. Therefore S can be extended to a basis of U.

Solution

To extend S to a basis of U, find a vector in U that is **not** in $\text{span}(S)$.

$$\begin{bmatrix} 1 & 2 & ? \\ 1 & 3 & ? \\ 1 & 3 & ? \\ 1 & 2 & ? \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 1 \\ 1 & 3 & 0 \\ 1 & 3 & -1 \\ 1 & 2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore, S can be extended to the following basis of U:

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \right\},$$

Next we consider the case of removing vectors from a spanning set to result in a basis.

Theorem 4.10.9: Finding a Basis from a Span

Let W be a subspace. Also suppose that $W = span\{\vec{w}_1, \cdots, \vec{w}_m\}$. Then there exists a subset of $\{\vec{w}_1, \cdots, \vec{w}_m\}$ which is a basis for W.

Proof

Let S denote the set of positive integers such that for $k \in S$, there exists a subset of $\{\vec{w}_1, \cdots, \vec{w}_m\}$ consisting of exactly k vectors which is a spanning set for W. Thus $m \in S$. Pick the smallest positive integer in S. Call it k. Then there exists $\{\vec{u}_1, \cdots, \vec{u}_k\} \subseteq \{\vec{w}_1, \cdots, \vec{w}_m\}$ such that $\text{span}\{\vec{u}_1, \cdots, \vec{u}_k\} = W$. If

$$\sum_{i=1}^{k} c_i \vec{w}_i = \vec{0}$$

and not all of the $c_i = 0$, then you could pick $c_j \neq 0$, divide by it and solve for $\vec{u}_j$ in terms of the others,

$$\vec{w}_j = \sum_{i \neq j} \left(-\frac{c_i}{c_j} \right) \vec{w}_i$$

Then you could delete $\vec{w}_j$ from the list and have the same span. Any linear combination involving $\vec{w}_j$ would equal one in which $\vec{w}_j$ is replaced with the above sum, showing that it could have been obtained as a linear combination of $\vec{w}_i$ for $i \neq j$. Thus $k - 1 \in S$ contrary to the choice of k. Hence each $c_i = 0$ and so $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is a basis for W consisting of vectors of $\{\vec{w}_1, \cdots, \vec{w}_m\}$.

Below is a video on determining a basis for a set of 4 vectors.



The following example illustrates how to carry out this shrinking process which will obtain a subset of a span of vectors which is linearly independent.

Example 4.10.13: Subset of a Span

Let W be the subspace

$$span \left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ 19 \\ -8 \\ 8 \end{bmatrix}, \begin{bmatrix} -6 \\ -15 \\ 6 \\ -6 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Find a basis for W which consists of a subset of the given vectors.

Solution

You can use the reduced row-echelon form to accomplish this reduction. Form the matrix which has the given vectors as columns.

$$\begin{bmatrix} 1 & 1 & 8 & -6 & 1 & 1 \\ 2 & 3 & 19 & -15 & 3 & 5 \\ -1 & -1 & -8 & 6 & 0 & 0 \\ 1 & 1 & 8 & -6 & 1 & 1 \end{bmatrix}$$

Then take the reduced row-echelon form

$$\begin{bmatrix} 1 & 0 & 5 & -3 & 0 & -2 \\ 0 & 1 & 3 & -3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

It follows that a basis for W is

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Since the first, second, and fifth columns are obviously a basis for the column space of the , the same is true for the matrix having the given vectors as columns.

Consider the following theorems regarding a subspace contained in another subspace.

Theorem 4.10.10: Subset of a Subspace

Let V and W be subspaces of $\mathbb{R}^n$, and suppose that $W \subseteq V$. Then $\dim(W) \leq \dim(V)$ with equality when $W = V$.

Theorem 4.10.11: Extending a Basis

Let W be any non-zero subspace $\mathbb{R}^n$ and let $W \subseteq V$ where V is also a subspace of $\mathbb{R}^n$. Then every basis of W can be extended to a basis for V.

The proof is left as an exercise but proceeds as follows. Begin with a basis for $W, \{\vec{w}_1, \cdots, \vec{w}_s\}$ and add in vectors from V until you obtain a basis for V. Not that the process will stop because the dimension of V is no more than n.

Consider the following example.

Example 4.10.14: Extending a Basis

Let $V = \mathbb{R}^4$ and let

$$W = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Extend this basis of W to a basis of $\mathbb{R}^n$.

Solution

An easy way to do this is to take the reduced row-echelon form of the matrix

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.10.1}$$

Note how the given vectors were placed as the first two columns and then the matrix was extended in such a way that it is clear that the span of the columns of this matrix yield all of $\mathbb{R}^4$. Now determine the pivot columns. The reduced row-echelon form is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix} \tag{4.10.2}$$

Therefore the pivot columns are

$$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

and now this is an extension of the given basis for W to a basis for $\mathbb{R}^4$.

Why does this work? The columns of (4.10.1) obviously span $\mathbb{R}^4$. In fact the span of the first four is the same as the span of all six.

Consider another example.

Example 4.10.15: Extending a Basis

Let W be the span of $\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ in $\mathbb{R}^4$. Let V consist of the span of the vectors

$$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -6 \\ 1 \\ -6 \end{bmatrix}, \begin{bmatrix} -5 \\ 7 \\ 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Find a basis for V which extends the basis for W.

Solution

Note that the above vectors are not linearly independent, but their span, denoted as V is a subspace which does include the subspace W.

Using the process outlined in the previous example, form the following matrix

$$\begin{bmatrix} 1 & 0 & 7 & -5 & 0 \\ 0 & 1 & -6 & 7 & 0 \\ 1 & 1 & 1 & 2 & 0 \\ 0 & 1 & -6 & 7 & 1 \end{bmatrix}$$

Next find its reduced row-echelon form

$$\begin{bmatrix} 1 & 0 & 7 & -5 & 0 \\ 0 & 1 & -6 & 7 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

It follows that a basis for V consists of the first two vectors and the last.

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Thus V is of dimension 3 and it has a basis which extends the basis for W.

Row Space, Column Space, and Null Space of a Matrix

We begin this section with a new definition.

Definition 4.10.9: Row and Column Space

Let A be an $m \times n$ matrix. The **column space** of A, written $\text{col}(A)$, is the span of the columns. The **row space** of A, written $\text{row}(A)$, is the span of the rows.

Below is a video on the column space of a matrix.



Below is a video on the null space of a matrix.



Using the reduced row-echelon form, we can obtain an efficient description of the row and column space of a matrix. Consider the following lemma.

Lemma 4.10.1: Effect of Row Operations on Row Space

Let A and B be $m \times n$ matrices such that A can be carried to B by elementary row [column] operations. Then $\text{row}(A) = \text{row}(B)$ $[\text{col}(A) = \text{col}(B)]$.

Proof

We will prove that the above is true for row operations, which can be easily applied to column operations.

Let $\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m$ denote the rows of A.

- If B is obtained from A by a interchanging two rows of A, then A and B have exactly the same rows, so $\text{row}(B) = \text{row}(A)$.
- Suppose $p \neq 0$, and suppose that for some j, $1 \leq j \leq m$, B is obtained from A by multiplying row j by p. Then

$$\text{row}(B) = \text{span}\{\vec{r}_1, \ldots, p\vec{r}_j, \ldots, \vec{r}_m\}.$$

Since

$$\{\vec{r}_1, \ldots, p\vec{r}_j, \ldots, \vec{r}_m\} \subseteq \text{row}(A),$$

it follows that $\text{row}(B) \subseteq \text{row}(A)$. Conversely, since

$$\{\vec{r}_1, \ldots, \vec{r}_m\} \subseteq \text{row}(B),$$

it follows that $\text{row}(A) \subseteq \text{row}(B)$. Therefore, $\text{row}(B) = \text{row}(A)$.

Suppose $p \neq 0$, and suppose that for some i and j, $1 \leq i, j \leq m$, B is obtained from A by adding p time row j to row i. Without loss of generality, we may assume $i < j$.

Then

$$\text{row}(B) = \text{span}\{\vec{r}_1, \ldots, \vec{r}_{i-1}, \vec{r}_i + p\vec{r}_j, \ldots, \vec{r}_j, \ldots, \vec{r}_m\}.$$

Since

$$\{\vec{r}_1, \ldots, \vec{r}_{i-1}, \vec{r}_i + p\vec{r}_j, \ldots, \vec{r}_m\} \subseteq \text{row}(A),$$

it follows that $\text{row}(B) \subseteq \text{row}(A)$.

Conversely, since

$$\{\vec{r}_1, \ldots, \vec{r}_m\} \subseteq \text{row}(B),$$

it follows that $\text{row}(A) \subseteq \text{row}(B)$. Therefore, $\text{row}(B) = \text{row}(A)$.

Consider the following lemma.

Lemma 4.10.2: Row Space of a reduced row-echelon form Matrix

Let A be an $m \times n$ matrix and let R be its reduced row-echelon form. Then the nonzero rows of R form a basis of $\text{row}(R)$, and consequently of $\text{row}(A)$.

This lemma suggests that we can examine the reduced row-echelon form of a matrix in order to obtain the row space. Consider now the column space. The column space can be obtained by simply saying that it equals the span of all the columns. However, you can often get the column space as the span of fewer columns than this. A variation of the previous lemma provides a solution. Suppose A is row reduced to its reduced row-echelon form R. Identify the pivot columns of R (columns which have leading ones), and take the corresponding columns of A. It turns out that this forms a basis of $\text{col}(A)$.

Below is a video on finding a basis for the column space of a matrix.



Below is a video on finding a basis for the null space of a matrix.



Below is a video on the determining if a vector is in the column space of a matrix and then finding a basis for the column space.



Before proceeding to an example of this concept, we revisit the definition of rank.

Definition 4.10.10: Rank of a Matrix

Previously, we defined $\text{rank}(A)$ to be the number of leading entries in the row-echelon form of A. Using an understanding of dimension and row space, we can now define rank as follows:

$$\text{rank}(A) = \dim(\text{row}(A))$$

Consider the following example.

Example 4.10.16: Rank, Column and Row Space

Find the rank of the following matrix and describe the column and row spaces.

$$A = \begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 1 & 3 & 6 & 0 & 2 \\ 3 & 7 & 8 & 6 & 6 \end{bmatrix}$$

Solution

The reduced row-echelon form of A is

$$\begin{bmatrix} 1 & 0 & -9 & 9 & 2 \\ 0 & 1 & 5 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore, the rank is 2.

Notice that the first two columns of R are pivot columns. By the discussion following Lemma 4.10.2, we find the corresponding columns of A, in this case the first two columns. Therefore a basis for $\text{col}(A)$ is given by

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} \right\}$$

For example, consider the third column of the original matrix. It can be written as a linear combination of the first two columns of the original matrix as follows.

$$\begin{bmatrix} 1 \\ 6 \\ 8 \end{bmatrix} = -9 \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix} + 5 \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix}$$

What about an efficient description of the row space? By Lemma 4.10.2 we know that the nonzero rows of R create a basis of $\text{row}(A)$. For the above matrix, the row space equals

$$\text{row}(A) = \text{span}\{[\,1 \quad 0 \quad -9 \quad 9 \quad 2\,], [\,0 \quad 1 \quad 5 \quad -3 \quad 0\,]\}$$

Below is a video on determining the number of vectors in a basis for the column space and the null space of a matrix.



Notice that the column space of A is given as the span of columns of the original matrix, while the row space of A is the span of rows of the reduced row-echelon form of A.

Consider another example.

Example 4.10.17: Rank, Column and Row Space

Find the rank of the following matrix and describe the column and row spaces.

$$\begin{bmatrix} 1 & 2 & 1 & 3 & 2 \\ 1 & 3 & 6 & 0 & 2 \\ 1 & 2 & 1 & 3 & 2 \\ 1 & 3 & 2 & 4 & 0 \end{bmatrix}$$

Solution

The reduced row-echelon form is

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \frac{13}{2} \\ 0 & 1 & 0 & 2 & -\frac{5}{2} \\ 0 & 0 & 1 & -1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and so the rank is 3. The row space is given by

$$\text{row}(A) = \text{span}\left\{ \begin{bmatrix} 1 & 0 & 0 & 0 & \frac{13}{2} \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 2 & -\frac{5}{2} \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & -1 & \frac{1}{2} \end{bmatrix} \right\}$$

Notice that the first three columns of the reduced row-echelon form are pivot columns. The column space is the span of the first three columns in the **original matrix**,

$$\text{col}(A) = \text{span}\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \\ 1 \\ 2 \end{bmatrix} \right\}$$

Consider the solution given above for Example 4.10.17, where the rank of A equals 3. Notice that the row space and the column space each had dimension equal to 3. It turns out that this is not a coincidence, and this essential result is referred to as the Rank Theorem and is given now. Recall that we defined $\text{rank}(A) = \dim(\text{row}(A))$.

Theorem 4.10.12: Rank Theorem

Let A be an $m \times n$ matrix. Then $\dim(\text{col}(A))$, the dimension of the column space, is equal to the dimension of the row space, $\dim(\text{row}(A))$.

Below is a video on the null space, rank, and nullity of a matrix.



The following statements all follow from the Rank Theorem.

Corollary 4.10.5: Results of the Rank Theorem

Let A be a matrix. Then the following are true:

1. $\text{rank}(A) = \text{rank}(A^T)$.
2. For A of size $m \times n$, $\text{rank}(A) \leq m$ and $\text{rank}(A) \leq n$.
3. For A of size $n \times n$, A is invertible if and only if $\text{rank}(A) = n$.
4. For invertible matrices B and C of appropriate size, $\text{rank}(A) = \text{rank}(BA) = \text{rank}(AC)$.

Consider the following example.

Example 4.10.18: Rank of the Transpose

Let

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$$

Find $\text{rank}(A)$ and $\text{rank}(A^T)$.

Solution

To find $\text{rank}(A)$ we first row reduce to find the reduced row-echelon form.

$$A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Therefore the rank of A is 2. Now consider A^T given by

$$A^T = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$$

Again we row reduce to find the reduced row-echelon form.

$$\begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

You can see that $\text{rank}(A^T) = 2$, the same as $\text{rank}(A)$.

We now define what is meant by the null space of a general $m \times n$ matrix.

Definition 4.10.11: Null Space, or Kernel, of A

The null space of a matrix A, also referred to as the kernel of A, is defined as follows.

$$\text{null}\,(A) = \left\{ \vec{x} : A\vec{x} = \vec{0} \right\}$$

It can also be referred to using the notation $\ker(A)$. Similarly, we can discuss the image of A, denoted by $\text{im}\,(A)$. The image of A consists of the vectors of $\mathbb{R}^m$ which "get hit" by A. The formal definition is as follows.

Definition 4.10.12: Image of A

The image of A, written $\text{im}\,(A)$ is given by

$$\text{im}\,(A) = \{ A\vec{x} : \vec{x} \in \mathbb{R}^n \}$$

Consider A as a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ whose action is given by multiplication. The following diagram displays this scenario.

$$\begin{matrix} \text{null}(A) & & \text{im}(A) \\ \mathbb{R}^n & \overset{A}{\to} & \mathbb{R}^m \end{matrix}$$

As indicated, $\text{im}\,(A)$ is a subset of $\mathbb{R}^m$ while $\text{null}\,(A)$ is a subset of $\mathbb{R}^n$.

It turns out that the null space and image of A are both subspaces. Consider the following example.

Example 4.10.19: Null Space

Let A be an $m \times n$ matrix. Then the null space of A, $\text{null}(A)$ is a subspace of $\mathbb{R}^n$.

Solution

- Since $A\vec{0}_n = \vec{0}_m$, $\vec{0}_n \in \text{null}(A)$.
- Let $\vec{x}, \vec{y} \in \text{null}(A)$. Then $A\vec{x} = \vec{0}_m$ and $A\vec{y} = \vec{0}_m$, so

$$A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} = \vec{0}_m + \vec{0}_m = \vec{0}_m,$$

  and thus $\vec{x} + \vec{y} \in \text{null}(A)$.
- Let $\vec{x} \in \text{null}(A)$ and $k \in \mathbb{R}$. Then $A\vec{x} = \vec{0}_m$, so

$$A(k\vec{x}) = k(A\vec{x}) = k\vec{0}_m = \vec{0}_m,$$

  and thus $k\vec{x} \in \text{null}(A)$.

Therefore by the subspace test, $\text{null}(A)$ is a subspace of $\mathbb{R}^n$.

The proof that $\text{im}(A)$ is a subspace of $\mathbb{R}^m$ is similar and is left as an exercise to the reader.

We now wish to find a way to describe $\text{null}(A)$ for a matrix A. However, finding $\text{null}\,(A)$ is not new! There is just some new terminology being used, as $\text{null}\,(A)$ is simply the solution to the system $A\vec{x} = \vec{0}$.

Theorem 4.10.13: Basis of null(A)

Let A be an $m \times n$ matrix such that $\text{rank}(A) = r$. Then the system $A\vec{x} = \vec{0}_m$ has $n - r$ basic solutions, providing a basis of $\text{null}(A)$ with $\dim(\text{null}(A)) = n - r$.

Consider the following example.

Example 4.10.20: Null Space of A

Let

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 2 & 3 & 3 \end{bmatrix}$$

Find $\text{null}\,(A)$ and $\text{im}\,(A)$.

Solution

In order to find $\text{null}\,(A)$, we simply need to solve the equation $A\vec{x} = \vec{0}$. This is the usual procedure of writing the augmented matrix, finding the reduced row-echelon form and then the solution. The augmented matrix and corresponding reduced row-echelon form are

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 2 & 3 & 3 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

The third column is not a pivot column, and therefore the solution will contain a parameter. The solution to the system $A\vec{x} = \vec{0}$ is given by

$$\begin{bmatrix} -3t \\ t \\ t \end{bmatrix} : t \in \mathbb{R}$$

which can be written as

$$t \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} : t \in \mathbb{R}$$

Therefore, the null space of A is all multiples of this vector, which we can write as

$$\text{null}(A) = \text{span} \left\{ \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Finally $\text{im}\,(A)$ is just $\{A\vec{x} : \vec{x} \in \mathbb{R}^n\}$ and hence consists of the span of all columns of A, that is $\text{im}\,(A) = \text{col}(A)$.

Notice from the above calculation that that the first two columns of the reduced row-echelon form are pivot columns. Thus the column space is the span of the first two columns in the **original matrix**, and we get

$$\text{im}\,(A) = \text{col}(A) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \right\}$$

Here is a larger example, but the method is entirely similar.

Example 4.10.21: Null Space of A

Let

$$A = \begin{bmatrix} 1 & 2 & 1 & 0 & 1 \\ 2 & -1 & 1 & 3 & 0 \\ 3 & 1 & 2 & 3 & 1 \\ 4 & -2 & 2 & 6 & 0 \end{bmatrix}$$

Find the null space of A.

Solution

To find the null space, we need to solve the equation $AX = 0$. The augmented matrix and corresponding reduced row-echelon form are given by

$$\left[\begin{array}{ccccc|c} 1 & 2 & 1 & 0 & 1 & 0 \\ 2 & -1 & 1 & 3 & 0 & 0 \\ 3 & 1 & 2 & 3 & 1 & 0 \\ 4 & -2 & 2 & 6 & 0 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccccc|c} 1 & 0 & \frac{3}{5} & \frac{6}{5} & \frac{1}{5} & 0 \\ 0 & 1 & \frac{1}{5} & -\frac{3}{5} & \frac{2}{5} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

It follows that the first two columns are pivot columns, and the next three correspond to parameters. Therefore, $\text{null}\,(A)$ is given by

$$\begin{bmatrix} \left(-\frac{3}{5}\right) s + \left(-\frac{6}{5}\right) t + \left(\frac{1}{5}\right) r \\ \left(-\frac{1}{5}\right) s + \left(\frac{3}{5}\right) t + \left(-\frac{2}{5}\right) r \\ s \\ t \\ r \end{bmatrix} : s, t, r \in \mathbb{R}.$$

We write this in the form

$$s\begin{bmatrix} -\frac{3}{5} \\ -\frac{1}{5} \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -\frac{6}{5} \\ \frac{3}{5} \\ 0 \\ 1 \\ 0 \end{bmatrix} + r\begin{bmatrix} \frac{1}{5} \\ -\frac{2}{5} \\ 0 \\ 0 \\ 1 \end{bmatrix} : s,t,r \in \mathbb{R}.$$

In other words, the null space of this matrix equals the span of the three vectors above. Thus

$$\text{null}\,(A) = \text{span}\left\{ \begin{bmatrix} -\frac{3}{5} \\ -\frac{1}{5} \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{6}{5} \\ \frac{3}{5} \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{5} \\ -\frac{2}{5} \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Notice also that the three vectors above are linearly independent and so the dimension of null (A) is 3. The following is true in general, the number of parameters in the solution of $AX = 0$ equals the dimension of the null space. Recall also that the number of leading ones in the reduced row-echelon form equals the number of pivot columns, which is the rank of the matrix, which is the same as the dimension of either the column or row space.

Before we proceed to an important theorem, we first define what is meant by the nullity of a matrix.

Definition 4.10.13: Nullity

The dimension of the null space of a matrix is called the **nullity**, denoted $\dim(\text{null}\,(A))$.

From our observation above we can now state an important theorem.

Theorem 4.10.14: Rank and Nullity

Let A be an $m \times n$ matrix. Then $\text{rank}\,(A) + \dim(\text{null}\,(A)) = n$.

Example 4.10.22: Rank and Nullity

Let

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 1 \\ 2 & 3 & 3 \end{bmatrix}$$

Find $\text{rank}\,(A)$ and $\dim(\text{null}\,(A))$.

Solution

In the above Example 4.10.20we determined that the reduced row-echelon form of A is given by

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

Therefore the rank of A is 2. We also determined that the null space of A is given by

$$\text{null}(A) = \text{span}\left\{ \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Therefore the nullity of A is 1. It follows from Theorem 4.10.14that $\text{rank}\,(A) + \dim(\text{null}\,(A)) = 2 + 1 = 3$, which is the number of columns of A.

Below is a video on the null space, rank, and nullity of a matrix.



We conclude this section with two similar, and important, theorems.

Theorem 4.10.15

Let A be an $m \times n$ matrix. The following are equivalent.

1. $\text{rank}(A) = n$.
2. $\text{row}(A) = \mathbb{R}^n$, i.e., the rows of A span $\mathbb{R}^n$.
3. The columns of A are independent in $\mathbb{R}^m$.
4. The $n \times n$ matrix $A^T A$ is invertible.
5. There exists an $n \times m$ matrix C so that $CA = I_n$.
6. If $A\vec{x} = \vec{0}_m$ for some $\vec{x} \in \mathbb{R}^n$, then $\vec{x} = \vec{0}_n$.

Theorem 4.10.16

Let A be an $m \times n$ matrix. The following are equivalent.

1. $\text{rank}(A) = m$.
2. $\text{col}(A) = \mathbb{R}^m$, i.e., the columns of A span $\mathbb{R}^m$.
3. The rows of A are independent in $\mathbb{R}^n$.
4. The $m \times m$ matrix AA^T is invertible.
5. There exists an $n \times m$ matrix C so that $AC = I_m$.
6. The system $A\vec{x} = \vec{b}$ is consistent for every $\vec{b} \in \mathbb{R}^m$.

This page titled 4.10: Spanning, Linear Independence and Basis in Rⁿ is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.10: Spanning, Linear Independence and Basis in Rⁿ** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.11: Orthogonality

Outcomes

A. Determine if a given set is orthogonal or orthonormal.
B. Determine if a given matrix is orthogonal.
C. Given a linearly independent set, use the Gram-Schmidt Process to find corresponding orthogonal and orthonormal sets.
D. Find the orthogonal projection of a vector onto a subspace.
E. Find the least squares approximation for a collection of points.

In this section, we examine what it means for vectors (and sets of vectors) to be **orthogonal** and **orthonormal**. First, it is necessary to review some important concepts. You may recall the definitions for the span of a set of vectors and a linear independent set of vectors. We include the definitions and examples here for convenience.

Definition 4.11.1: Span of a Set of Vectors and Subspace

The collection of all linear combinations of a set of vectors $\{\vec{u}_1, \cdots, \vec{u}_k\}$ in $\mathbb{R}^n$ is known as the span of these vectors and is written as $\text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$.
We call a collection of the form $\text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$ a subspace of $\mathbb{R}^n$.

Consider the following example.

Example 4.11.1: Spanning Vectors

Describe the span of the vectors $\vec{u} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T$ and $\vec{v} = \begin{bmatrix} 3 & 2 & 0 \end{bmatrix}^T \in \mathbb{R}^3$.

Solution

You can see that any linear combination of the vectors $\vec{u}$ and $\vec{v}$ yields a vector $\begin{bmatrix} x & y & 0 \end{bmatrix}^T$ in the XY-plane.

Moreover every vector in the XY-plane is in fact such a linear combination of the vectors $\vec{u}$ and $\vec{v}$. That's because

$$\begin{bmatrix} x \\ y \\ 0 \end{bmatrix} = (-2x+3y) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + (x-y) \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}$$

Thus $\text{span}\{\vec{u}, \vec{v}\}$ is precisely the XY-plane.

The span of a set of a vectors in $\mathbb{R}^n$ is what we call a **subspace of $\mathbb{R}^n$**. A subspace W is characterized by the feature that any linear combination of vectors of W is again a vector contained in W.

Another important property of sets of vectors is called **linear independence**.

Definition 4.11.2: Linear Independence

A set of non-zero vectors $\{\vec{u}_1, \cdots, \vec{u}_k\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if no vector in that set is in the span of the other vectors of that set.

Here is an example.

Example 4.11.2: Linearly Independent Vectors

Consider vectors $\vec{u} = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^T$, $\vec{v} = \begin{bmatrix} 3 & 2 & 0 \end{bmatrix}^T$, and $\vec{w} = \begin{bmatrix} 4 & 5 & 0 \end{bmatrix}^T \in \mathbb{R}^3$. Verify whether the set $\{\vec{u}, \vec{v}, \vec{w}\}$ is linearly independent.

Solution

We already verified in Example 4.11.1 that $\text{span}\{\vec{u}, \vec{v}\}$ is the XY-plane. Since $\vec{w}$ is clearly also in the XY-plane, then the set $\{\vec{u}, \vec{v}, \vec{w}\}$ is **not** linearly independent.

In terms of spanning, a set of vectors is linearly independent if it does not contain unnecessary vectors. In the previous example you can see that the vector $\vec{w}$ does not help to span any new vector not already in the span of the other two vectors. However you can verify that the set $\{\vec{u}, \vec{v}\}$ is linearly independent, since you will not get the XY-plane as the span of a single vector.

We can also determine if a set of vectors is linearly independent by examining linear combinations. A set of vectors is linearly independent if and only if whenever a linear combination of these vectors equals zero, it follows that all the coefficients equal zero. It is a good exercise to verify this equivalence, and this latter condition is often used as the (equivalent) definition of linear independence.

If a subspace is spanned by a linearly independent set of vectors, then we say that it is a basis for the subspace.

Definition 4.11.3: Basis

Let V be a subspace of $\mathbb{R}^n$. Then $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is a **basis** for V if the following two conditions hold.

1. $\text{span}\{\vec{u}_1, \cdots, \vec{u}_k\} = V$
2. $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is linearly independent

Thus the set of vectors $\{\vec{u}, \vec{v}\}$ from Example 4.11.2 is a basis for XY-plane in $\mathbb{R}^3$ since it is both linearly independent and spans the XY-plane.

Recall from the properties of the dot product of vectors that two vectors $\vec{u}$ and $\vec{v}$ are orthogonal if $\vec{u} \cdot \vec{v} = 0$. Suppose a vector is orthogonal to a spanning set of $\mathbb{R}^n$. What can be said about such a vector? This is the discussion in the following example.

Example 4.11.3: Orthogonal Vector to a Spanning Set

Let $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\} \in \mathbb{R}^n$ and suppose $\mathbb{R}^n = \text{span}\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$. Furthermore, suppose that there exists a vector $\vec{u} \in \mathbb{R}^n$ for which $\vec{u} \cdot \vec{x}_j = 0$ for all j, $1 \leq j \leq k$. What type of vector is $\vec{u}$?

Solution

Write $\vec{u} = t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k$ for some $t_1, t_2, \ldots, t_k \in \mathbb{R}$ (this is possible because $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k$ span $\mathbb{R}^n$).

Then

$$\begin{aligned} \|\vec{u}\|^2 &= \vec{u} \cdot \vec{u} \\ &= \vec{u} \cdot (t_1\vec{x}_1 + t_2\vec{x}_2 + \cdots + t_k\vec{x}_k) \\ &= \vec{u} \cdot (t_1\vec{x}_1) + \vec{u} \cdot (t_2\vec{x}_2) + \cdots + \vec{u} \cdot (t_k\vec{x}_k) \\ &= t_1(\vec{u} \cdot \vec{x}_1) + t_2(\vec{u} \cdot \vec{x}_2) + \cdots + t_k(\vec{u} \cdot \vec{x}_k) \\ &= t_1(0) + t_2(0) + \cdots + t_k(0) = 0. \end{aligned}$$

Since $\|\vec{u}\|^2 = 0$, $\|\vec{u}\| = 0$. We know that $\|\vec{u}\| = 0$ if and only if $\vec{u} = \vec{0}_n$. Therefore, $\vec{u} = \vec{0}_n$. In conclusion, the only vector orthogonal to every vector of a spanning set of $\mathbb{R}^n$ is the zero vector.

We can now discuss what is meant by an orthogonal set of vectors.

Definition 4.11.4: Orthogonal Set of Vectors

Let $\{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_m\}$ be a set of vectors in $\mathbb{R}^n$. Then this set is called an **orthogonal set** if the following conditions hold:

1. $\vec{u}_i \cdot \vec{u}_j = 0$ for all $i \neq j$
2. $\vec{u}_i \neq \vec{0}$ for all i

If we have an orthogonal set of vectors and normalize each vector so they have length 1, the resulting set is called an **orthonormal set** of vectors. They can be described as follows.

Definition 4.11. :5 Orthonormal Set of Vectors

A set of vectors, $\{\vec{w}_1, \cdots, \vec{w}_m\}$ is said to be an **orthonormal** set if

$$\vec{w}_i \cdot \vec{w}_j = \delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

Note that all orthonormal sets are orthogonal, but the reverse is not necessarily true since the vectors may not be normalized. In order to normalize the vectors, we simply need divide each one by its length.

Below is a video on orthogonal and orthonormal sets of vectors.



Definition 4.11.6: Normalizing an Orthogonal Set

Normalizing an orthogonal set is the process of turning an orthogonal (but not orthonormal) set into an orthonormal set. If $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ is an orthogonal subset of $\mathbb{R}^n$, then

$$\left\{ \frac{1}{\|\vec{u}_1\|}\vec{u}_1, \frac{1}{\|\vec{u}_2\|}\vec{u}_2, \ldots, \frac{1}{\|\vec{u}_k\|}\vec{u}_k \right\}$$

is an orthonormal set.

We illustrate this concept in the following example.

Example 4.11.4: Orthonormal Set

Consider the set of vectors given by

$$\{\vec{u}_1, \vec{u}_2\} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

Show that it is an orthogonal set of vectors but not an orthonormal one. Find the corresponding orthonormal set.

Solution

One easily verifies that $\vec{u}_1 \cdot \vec{u}_2 = 0$ and $\{\vec{u}_1, \vec{u}_2\}$ is an orthogonal set of vectors. On the other hand one can compute that $\|\vec{u}_1\| = \|\vec{u}_2\| = \sqrt{2} \neq 1$ and thus it is not an orthonormal set.

Thus to find a corresponding orthonormal set, we simply need to normalize each vector. We will write $\{\vec{w}_1, \vec{w}_2\}$ for the corresponding orthonormal set. Then,

$$\begin{aligned}\vec{w}_1 &= \frac{1}{\|\vec{u}_1\|}\vec{u}_1 \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} \\ &= \begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}\end{aligned}$$

Similarly,

$$\begin{aligned}\vec{w}_2 &= \frac{1}{\|\vec{u}_2\|}\vec{u}_2 \\ &= \frac{1}{\sqrt{2}}\begin{bmatrix}-1\\1\end{bmatrix} \\ &= \begin{bmatrix}-\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}\end{aligned}$$

Therefore the corresponding orthonormal set is

$$\{\vec{w}_1, \vec{w}_2\} = \left\{\begin{bmatrix}\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}, \begin{bmatrix}-\frac{1}{\sqrt{2}}\\ \frac{1}{\sqrt{2}}\end{bmatrix}\right\}$$

You can verify that this set is orthogonal.

Consider an orthogonal set of vectors in $\mathbb{R}^n$, written $\{\vec{w}_1, \cdots, \vec{w}_k\}$ with $k \leq n$. The span of these vectors is a subspace W of $\mathbb{R}^n$. If we could show that this orthogonal set is also linearly independent, we would have a basis of W. We will show this in the next theorem.

Theorem 4.11.1: Orthogonal Basis of a Subspace

Let $\{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_k\}$ be an orthonormal set of vectors in $\mathbb{R}^n$. Then this set is linearly independent and forms a basis for the subspace $W = \text{span}\{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_k\}$.

Proof

To show it is a linearly independent set, suppose a linear combination of these vectors equals $\vec{0}$, such as:

$$a_1\vec{w}_1 + a_2\vec{w}_2 + \cdots + a_k\vec{w}_k = \vec{0}, a_i \in \mathbb{R}$$

We need to show that all $a_i = 0$. To do so, take the dot product of each side of the above equation with the vector $\vec{w}_i$ and obtain the following.

$$\begin{aligned}\vec{w}_i \cdot (a_1\vec{w}_1 + a_2\vec{w}_2 + \cdots + a_k\vec{w}_k) &= \vec{w}_i \cdot \vec{0} \\ a_1(\vec{w}_i \cdot \vec{w}_1) + a_2(\vec{w}_i \cdot \vec{w}_2) + \cdots + a_k(\vec{w}_i \cdot \vec{w}_k) &= 0\end{aligned}$$

Now since the set is orthogonal, $\vec{w}_i \cdot \vec{w}_m = 0$ for all $m \neq i$, so we have:

$$a_1(0) + \cdots + a_i(\vec{w}_i \cdot \vec{w}_i) + \cdots + a_k(0) = 0$$

$$a_i\|\vec{w}_i\|^2 = 0$$

Since the set is orthogonal, we know that $\|\vec{w}_i\|^2 \neq 0$. It follows that $a_i = 0$. Since the a_i was chosen arbitrarily, the set $\{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_k\}$ is linearly independent.

Finally since $W = \text{span}\{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_k\}$, the set of vectors also spans W and therefore forms a basis of W.

If an orthogonal set is a basis for a subspace, we call this an orthogonal basis. Similarly, if an orthonormal set is a basis, we call this an orthonormal basis.

We conclude this section with a discussion of Fourier expansions. Given any orthogonal basis B of $\mathbb{R}^n$ and an arbitrary vector $\vec{x} \in \mathbb{R}^n$, how do we express $\vec{x}$ as a linear combination of vectors in B? The solution is Fourier expansion.

Theorem 4.11.2: Fourier Expansion

Let V be a subspace of $\mathbb{R}^n$ and suppose $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m\}$ is an orthogonal basis of V. Then for any $\vec{x} \in V$,

$$\vec{x} = \left(\frac{\vec{x} \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\right)\vec{u}_1 + \left(\frac{\vec{x} \cdot \vec{u}_2}{\|\vec{u}_2\|^2}\right)\vec{u}_2 + \cdots + \left(\frac{\vec{x} \cdot \vec{u}_m}{\|\vec{u}_m\|^2}\right)\vec{u}_m$$

This expression is called the **Fourier expansion** of $\vec{x}$, and

$$\frac{\vec{x} \cdot \vec{u}_j}{\|\vec{u}_j\|^2},$$

$j = 1, 2, \ldots, m$ are the Fourier coefficients.

Consider the following example.

Example 4.11.5: Fourier Expansion

Let $\vec{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$, $\vec{u}_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, and $\vec{u}_3 = \begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}$, and let $\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Then $B = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthogonal basis of $\mathbb{R}^3$.

Compute the Fourier expansion of $\vec{x}$, thus writing $\vec{x}$ as a linear combination of the vectors of B.

Solution

Since B is a basis (verify!) there is a unique way to express $\vec{x}$ as a linear combination of the vectors of B. Moreover since B is an orthogonal basis (verify!), then this can be done by computing the Fourier expansion of $\vec{x}$.

That is:

$$\vec{x} = \left(\frac{\vec{x} \cdot \vec{u}_1}{\|\vec{u}_1\|^2}\right)\vec{u}_1 + \left(\frac{\vec{x} \cdot \vec{u}_2}{\|\vec{u}_2\|^2}\right)\vec{u}_2 + \left(\frac{\vec{x} \cdot \vec{u}_3}{\|\vec{u}_3\|^2}\right)\vec{u}_3.$$

We readily compute:

$$\frac{\vec{x} \cdot \vec{u}_1}{\|\vec{u}_1\|^2} = \frac{2}{6}, \quad \frac{\vec{x} \cdot \vec{u}_2}{\|\vec{u}_2\|^2} = \frac{3}{5}, \text{ and } \frac{\vec{x} \cdot \vec{u}_3}{\|\vec{u}_3\|^2} = \frac{4}{30}.$$

Therefore,

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + \frac{3}{5}\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} + \frac{2}{15}\begin{bmatrix} 5 \\ 1 \\ -2 \end{bmatrix}.$$

Orthogonal Matrices

Recall that the process to find the inverse of a matrix was often cumbersome. In contrast, it was very easy to take the transpose of a matrix. Luckily for some special matrices, the transpose equals the inverse. When an $n \times n$ matrix has all real entries and its transpose equals its inverse, the matrix is called an **orthogonal matrix**.

The precise definition is as follows.

Definition 4.11.7: Orthogonal Matrices

A real $n \times n$ matrix U is called an **orthogonal** matrix if

$$UU^T = U^T U = I.$$

Note since U is assumed to be a square matrix, it suffices to verify only one of these equalities $UU^T = I$ or $U^T U = I$ holds to guarantee that U^T is the inverse of U.

Consider the following example.

Example 4.11.6: Orthogonal Matrix

Orthogonal Matrix Show the matrix

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

is orthogonal.

Solution

All we need to do is verify (one of the equations from) the requirements of Definition 4.11.7.

$$UU^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Since $UU^T = I$, this matrix is orthogonal.

Here is another example.

Example 4.11.7: Orthogonal Matrix

Orthogonal Matrix Let $U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}$. Is U orthogonal?

Solution

Again the answer is yes and this can be verified simply by showing that $U^T U = I$:

$$\begin{aligned} U^T U &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}^T \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

When we say that U is orthogonal, we are saying that $UU^T = I$, meaning that

$$\sum_j u_{ij} u_{jk}^T = \sum_j u_{ij} u_{kj} = \delta_{ik}$$

where δ_{ij} is the **Kronecker symbol** defined by

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

In words, the product of the ith row of U with the kth row gives 1 if $i = k$ and 0 if $i \neq k$. The same is true of the columns because $U^T U = I$ also. Therefore,

$$\sum_j u_{ij}^T u_{jk} = \sum_j u_{ji} u_{jk} = \delta_{ik}$$

which says that the product of one column with another column gives 1 if the two columns are the same and 0 if the two columns are different.

More succinctly, this states that if $\vec{u}_1, \cdots, \vec{u}_n$ are the columns of U, an orthogonal matrix, then

$$\vec{u}_i \cdot \vec{u}_j = \delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

We will say that the columns form an orthonormal set of vectors, and similarly for the rows. Thus a matrix is **orthogonal** if its rows (or columns) form an **orthonormal** set of vectors. Notice that the convention is to call such a matrix orthogonal rather than orthonormal (although this may make more sense!).

Proposition 4.11.1: Orthonormal Basis

The rows of an $n \times n$ orthogonal matrix form an orthonormal basis of $\mathbb{R}^n$. Further, any orthonormal basis of $\mathbb{R}^n$ can be used to construct an $n \times n$ orthogonal matrix.

Proof

Recall from Theorem 4.11.1that an orthonormal set is linearly independent and forms a basis for its span. Since the rows of an $n \times n$ orthogonal matrix form an orthonormal set, they must be linearly independent. Now we have n linearly independent vectors, and it follows that their span equals $\mathbb{R}^n$. Therefore these vectors form an orthonormal basis for $\mathbb{R}^n$.

Suppose now that we have an orthonormal basis for $\mathbb{R}^n$. Since the basis will contain n vectors, these can be used to construct an $n \times n$ matrix, with each vector becoming a row. Therefore the matrix is composed of orthonormal rows, which by our above discussion, means that the matrix is orthogonal. Note we could also have construct a matrix with each vector becoming a column instead, and this would again be an orthogonal matrix. In fact this is simply the transpose of the previous matrix.

Consider the following proposition.

Proposition 4.11.2: Determinant of Orthogonal Matrices

Det Suppose U is an orthogonal matrix. Then $\det(U) = \pm 1$.

Proof

This result follows from the properties of determinants. Recall that for any matrix A, $\det(A)^T = \det(A)$. Now if U is orthogonal, then:

$$(\det(U))^2 = \det\left(U^T\right) \det(U) = \det\left(U^T U\right) = \det(I) = 1$$

Therefore $(\det(U))^2 = 1$ and it follows that $\det(U) = \pm 1$.

Orthogonal matrices are divided into two classes, proper and improper. The proper orthogonal matrices are those whose determinant equals 1 and the improper ones are those whose determinant equals -1. The reason for the distinction is that the improper orthogonal matrices are sometimes considered to have no physical significance. These matrices cause a change in orientation which would correspond to material passing through itself in a non physical manner. Thus in considering which coordinate systems must be considered in certain applications, you only need to consider those which are related by a proper

orthogonal transformation. Geometrically, the linear transformations determined by the proper orthogonal matrices correspond to the composition of rotations.

We conclude this section with two useful properties of orthogonal matrices.

Example 4.11.8: Product and Inverse of Orthogonal Matrices

Suppose A and B are orthogonal matrices. Then AB and A^{-1} both exist and are orthogonal.

Solution

First we examine the product AB.

$$(AB)(B^TA^T) = A(BB^T)A^T = AA^T = I$$

Since AB is square, $B^TA^T = (AB)^T$ is the inverse of AB, so AB is invertible, and $(AB)^{-1} = (AB)^T$ Therefore, AB is orthogonal.

Next we show that $A^{-1} = A^T$ is also orthogonal.

$$(A^{-1})^{-1} = A = (A^T)^T = (A^{-1})^T$$

Therefore A^{-1} is also orthogonal.

Gram-Schmidt Process

The Gram-Schmidt process is an algorithm to transform a set of vectors into an orthonormal set spanning the same subspace, that is generating the same collection of linear combinations (see Definition 9.2.2).

The goal of the Gram-Schmidt process is to take a linearly independent set of vectors and transform it into an orthonormal set with the same span. The first objective is to construct an orthogonal set of vectors with the same span, since from there an orthonormal set can be obtained by simply dividing each vector by its length.

Algorithm 4.11.1: Gram-Schmidt Process

Let $\{\vec{u}_1, \cdots, \vec{u}_n\}$ be a set of linearly independent vectors in $\mathbb{R}^n$.

I: Construct a new set of vectors $\{\vec{v}_1, \cdots, \vec{v}_n\}$ as follows:

$$\begin{aligned}
\vec{v}_1 &= \vec{u}_1 \\
\vec{v}_2 &= \vec{u}_2 - \left(\frac{\vec{u}_2 \cdot \vec{v}_1}{\|\vec{v}_1\|^2}\right)\vec{v}_1 \\
\vec{v}_3 &= \vec{u}_3 - \left(\frac{\vec{u}_3 \cdot \vec{v}_1}{\|\vec{v}_1\|^2}\right)\vec{v}_1 - \left(\frac{\vec{u}_3 \cdot \vec{v}_2}{\|\vec{v}_2\|^2}\right)\vec{v}_2 \\
&\vdots \\
\vec{v}_n &= \vec{u}_n - \left(\frac{\vec{u}_n \cdot \vec{v}_1}{\|\vec{v}_1\|^2}\right)\vec{v}_1 - \left(\frac{\vec{u}_n \cdot \vec{v}_2}{\|\vec{v}_2\|^2}\right)\vec{v}_2 - \cdots - \left(\frac{\vec{u}_n \cdot \vec{v}_{n-1}}{\|\vec{v}_{n-1}\|^2}\right)\vec{v}_{n-1}
\end{aligned}$$

II: Now let $\vec{w}_i = \dfrac{\vec{v}_i}{\|\vec{v}_i\|}$ for $i = 1, \cdots, n$.

Then

1. $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is an orthogonal set.
2. $\{\vec{w}_1, \cdots, \vec{w}_n\}$ is an orthonormal set.
3. $\text{span}\{\vec{u}_1, \cdots, \vec{u}_n\} = \text{span}\{\vec{v}_1, \cdots, \vec{v}_n\} = \text{span}\{\vec{w}_1, \cdots, \vec{w}_n\}$.

Solution

The full proof of this algorithm is beyond this material, however here is an indication of the arguments.

To show that $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is an orthogonal set, let

$$a_2 = \frac{\vec{u}_2 \cdot \vec{v}_1}{\|\vec{v}_1\|^2}$$

then:

$$\begin{aligned} \vec{v}_1 \cdot \vec{v}_2 &= \vec{v}_1 \cdot (\vec{u}_2 - a_2 \vec{v}_1) \\ &= \vec{v}_1 \cdot \vec{u}_2 - a_2 (\vec{v}_1 \cdot \vec{v}_1 \\ &= \vec{v}_1 \cdot \vec{u}_2 - \frac{\vec{u}_2 \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \|\vec{v}_1\|^2 \\ &= (\vec{v}_1 \cdot \vec{u}_2) - (\vec{u}_2 \cdot \vec{v}_1) = 0 \end{aligned}$$

Now that you have shown that $\{\vec{v}_1, \vec{v}_2\}$ is orthogonal, use the same method as above to show that $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is also orthogonal, and so on.

Then in a similar fashion you show that $\text{span}\{\vec{u}_1, \cdots, \vec{u}_n\} = \text{span}\{\vec{v}_1, \cdots, \vec{v}_n\}$.

Finally defining $\vec{w}_i = \frac{\vec{v}_i}{\|\vec{v}_i\|}$ for $i = 1, \cdots, n$ does not affect orthogonality and yields vectors of length 1, hence an orthonormal set. You can also observe that it does not affect the span either and the proof would be complete.

Below is a video on the Gram Schmidt process.



Consider the following example.

Example 4.11.9: Find Orthonormal Set with Same Span

Consider the set of vectors $\{\vec{u}_1, \vec{u}_2\}$ given as in Example 4.11.1. That is

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} \in \mathbb{R}^3$$

Use the Gram-Schmidt algorithm to find an orthonormal set of vectors $\{\vec{w}_1, \vec{w}_2\}$ having the same span.

Solution

We already remarked that the set of vectors in $\{\vec{u}_1, \vec{u}_2\}$ is linearly independent, so we can proceed with the Gram-Schmidt algorithm:

$$\vec{v}_1 = \vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \vec{v}_2 &= \vec{u}_2 - \left(\frac{\vec{u}_2 \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \right) \vec{v}_1 \\ &= \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} - \frac{5}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 0 \end{bmatrix} \end{aligned}$$

Now to normalize simply let

$$\vec{w}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

You can verify that $\{\vec{w}_1, \vec{w}_2\}$ is an orthonormal set of vectors having the same span as $\{\vec{u}_1, \vec{u}_2\}$, namely the XY-plane.

In this example, we began with a linearly independent set and found an orthonormal set of vectors which had the same span. It turns out that if we start with a basis of a subspace and apply the Gram-Schmidt algorithm, the result will be an orthogonal basis of the same subspace. We examine this in the following example.

Example 4.11.10: Find a Corresponding Orthogonal Basis

Let

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \text{ and } \vec{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix},$$

and let $U = \text{span}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$. Use the Gram-Schmidt Process to construct an orthogonal basis B of U.

Solution

First $\vec{f}_1 = \vec{x}_1$.

Next,

$$\vec{f}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Finally,

$$\vec{f}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{0}{1} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \\ -1/2 \\ 0 \end{bmatrix}.$$

Therefore,

$$\left\{\begin{bmatrix}1\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\end{bmatrix},\begin{bmatrix}1/2\\1\\-1/2\\0\end{bmatrix}\right\}$$

is an orthogonal basis of U. However, it is sometimes more convenient to deal with vectors having integer entries, in which case we take

$$B=\left\{\begin{bmatrix}1\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\end{bmatrix},\begin{bmatrix}1\\2\\-1\\0\end{bmatrix}\right\}.$$

Below is another video on the Gram Schmidt process.



Orthogonal Projections

An important use of the Gram-Schmidt Process is in orthogonal projections, the focus of this section.

You may recall that a subspace of $\mathbb{R}^n$ is a set of vectors which contains the zero vector, and is closed under addition and scalar multiplication. Let's call such a subspace W. In particular, a plane in $\mathbb{R}^n$ which contains the origin, $(0,0,\cdots,0)$, is a subspace of $\mathbb{R}^n$.

Suppose a point Y in $\mathbb{R}^n$ is not contained in W, then what point Z in W is closest to Y? Using the Gram-Schmidt Process, we can find such a point. Let $\vec{y},\vec{z}$ represent the position vectors of the points Y and Z respectively, with $\vec{y}-\vec{z}$ representing the vector connecting the two points Y and Z. It will follow that if Z is the point on W closest to Y, then $\vec{y}-\vec{z}$ will be perpendicular to W (can you see why?); in other words, $\vec{y}-\vec{z}$ is orthogonal to W (and to every vector contained in W) as in the following diagram.

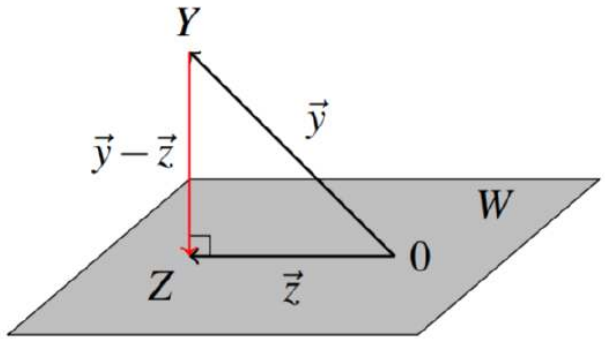


Figure 4.11.1

The vector $\vec{z}$ is called the **orthogonal projection** of $\vec{y}$ on W. The definition is given as follows.

Definition 4.11.8: Orthogonal Projection

Let W be a subspace of $\mathbb{R}^n$, and Y be any point in $\mathbb{R}^n$. Then the orthogonal projection of Y onto W is given by

$$\vec{z} = \text{proj}_W(\vec{y}) = \left(\frac{\vec{y}\cdot\vec{w}_1}{\|\vec{w}_1\|^2}\right)\vec{w}_1 + \left(\frac{\vec{y}\cdot\vec{w}_2}{\|\vec{w}_2\|^2}\right)\vec{w}_2 + \cdots + \left(\frac{\vec{y}\cdot\vec{w}_m}{\|\vec{w}_m\|^2}\right)\vec{w}_m$$

where $\{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_m\}$ is any orthogonal basis of W.

Therefore, in order to find the orthogonal projection, we must first find an orthogonal basis for the subspace. Note that one could use an orthonormal basis, but it is not necessary in this case since as you can see above the normalization of each vector is included in the formula for the projection.

Below is a video on finding an orthogonal projection of a vector onto a line.



Below is a video on finding an orthogonal projection of a vector onto a plane.



Below is a video on finding an orthogonal projection of a vector onto a subspace of R^4.



Before we explore this further through an example, we show that the orthogonal projection does indeed yield a point Z (the point whose position vector is the vector $\vec{z}$ above) which is the point of W closest to Y.

Theorem 4.11.3: Approximation Theorem

Let W be a subspace of $\mathbb{R}^n$ and Y any point in $\mathbb{R}^n$. Let Z be the point whose position vector is the orthogonal projection of Y onto W.

Then, Z is the point in W closest to Y.

Proof

First Z is certainly a point in W since it is in the span of a basis of W.

To show that Z is the point in W closest to Y, we wish to show that $|\vec{y}-\vec{z}_1| > |\vec{y}-\vec{z}|$ for all $\vec{z}_1 \neq \vec{z} \in W$. We begin by writing $\vec{y}-\vec{z}_1 = (\vec{y}-\vec{z})+(\vec{z}-\vec{z}_1)$. Now, the vector $\vec{y}-\vec{z}$ is orthogonal to W, and $\vec{z}-\vec{z}_1$ is contained in W. Therefore these vectors are orthogonal to each other. By the Pythagorean Theorem, we have that

$$\|\vec{y}-\vec{z}_1\|^2 = \|\vec{y}-\vec{z}\|^2 + \|\vec{z}-\vec{z}_1\|^2 > \|\vec{y}-\vec{z}\|^2$$

This follows because $\vec{z} \neq \vec{z}_1$ so $\|\vec{z}-\vec{z}_1\|^2 > 0$.

Hence, $\|\vec{y}-\vec{z}_1\|^2 > \|\vec{y}-\vec{z}\|^2$. Taking the square root of each side, we obtain the desired result.

Consider the following example.

Example 4.11.11: Orthogonal Projection

Let W be the plane through the origin given by the equation $x-2y+z=0$.
Find the point in W closest to the point $Y=(1,0,3)$.

Solution

We must first find an orthogonal basis for W. Notice that W is characterized by all points (a,b,c) where $c=2b-a$. In other words,

$$W = \begin{bmatrix} a \\ b \\ 2b-a \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix},\ a,b \in \mathbb{R}$$

We can thus write W as

$$\begin{aligned} W &= \text{span}\{\vec{u}_1, \vec{u}_2\} \\ &= \text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\right\} \end{aligned}$$

Notice that this span is a basis of W as it is linearly independent. We will use the Gram-Schmidt Process to convert this to an orthogonal basis, $\{\vec{w}_1, \vec{w}_2\}$. In this case, as we remarked it is only necessary to find an orthogonal basis, and it is not required that it be orthonormal.

$$\begin{aligned}
\vec{w}_1 = \vec{u}_1 &= \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \\
\vec{w}_2 &= \vec{u}_2 - \left(\frac{\vec{u}_2 \cdot \vec{w}_1}{\|\vec{w}_1\|^2} \right) \vec{w}_1 \\
&= \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} - \left(\frac{-2}{2} \right) \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}
\end{aligned}$$

Therefore an orthogonal basis of W is

$$\{\vec{w}_1, \vec{w}_2\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

We can now use this basis to find the orthogonal projection of the point $Y = (1, 0, 3)$ on the subspace W. We will write the position vector $\vec{y}$ of Y as $\vec{y} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$. Using Definition 4.11.8, we compute the projection as follows:

$$\begin{aligned}
\vec{z} &= \text{proj}_W(\vec{y}) \\
&= \left(\frac{\vec{y} \cdot \vec{w}_1}{\|\vec{w}_1\|^2} \right) \vec{w}_1 + \left(\frac{\vec{y} \cdot \vec{w}_2}{\|\vec{w}_2\|^2} \right) \vec{w}_2 \\
&= \left(\frac{-2}{2} \right) \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \left(\frac{4}{3} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{3} \\ \frac{4}{3} \\ \frac{7}{3} \end{bmatrix}
\end{aligned}$$

Therefore the point Z on W closest to the point $(1, 0, 3)$ is $\left(\frac{1}{3}, \frac{4}{3}, \frac{7}{3}\right)$.

Recall that the vector $\vec{y} - \vec{z}$ is perpendicular (orthogonal) to all the vectors contained in the plane W. Using a basis for W, we can in fact find all such vectors which are perpendicular to W. We call this set of vectors the **orthogonal complement** of W and denote it $W^\perp$.

Definition 4.11.9: Orthogonal Complement

Let W be a subspace of $\mathbb{R}^n$. Then the orthogonal complement of W, written $W^\perp$, is the set of all vectors $\vec{x}$ such that $\vec{x} \cdot \vec{z} = 0$ for all vectors $\vec{z}$ in W.

$$W^\perp = \{\vec{x} \in \mathbb{R}^n \text{ such that } \vec{x} \cdot \vec{z} = 0 \text{ for all } \vec{z} \in W\}$$

The orthogonal complement is defined as the set of all vectors which are orthogonal to all vectors in the original subspace. It turns out that it is sufficient that the vectors in the orthogonal complement be orthogonal to a spanning set of the original space.

Proposition 4.11.3: Orthogonal to Spanning Set

Let W be a subspace of $\mathbb{R}^n$ such that $W = \text{span}\{\vec{w}_1, \vec{w}_2, \cdots, \vec{w}_m\}$. Then $W^\perp$ is the set of all vectors which are orthogonal to each $\vec{w}_i$ in the spanning set.

The following proposition demonstrates that the orthogonal complement of a subspace is itself a subspace.

Proposition 4.11.4: The Orthogonal Complement

Let W be a subspace of $\mathbb{R}^n$. Then the orthogonal complement $W^\perp$ is also a subspace of $\mathbb{R}^n$.

Consider the following proposition.

Proposition 4.11.5: Orthogonal Complement of $\mathbb{R}^n$

The complement of $\mathbb{R}^n$ is the set containing the zero vector:

$$(\mathbb{R}^n)^\perp = \left\{\vec{0}\right\}$$

Similarly,

$$\left\{\vec{0}\right\}^\perp = (\mathbb{R}^n).$$

Proof

Here, $\vec{0}$ is the zero vector of $\mathbb{R}^n$. Since $\vec{x} \cdot \vec{0} = 0$ for all $\vec{x} \in \mathbb{R}^n$, $\mathbb{R}^n \subseteq \{\vec{0}\}^\perp$. Since $\{\vec{0}\}^\perp \subseteq \mathbb{R}^n$, the equality follows, i.e., $\{\vec{0}\}^\perp = \mathbb{R}^n$.

Again, since $\vec{x} \cdot \vec{0} = 0$ for all $\vec{x} \in \mathbb{R}^n$, $\vec{0} \in (\mathbb{R}^n)^\perp$, so $\{\vec{0}\} \subseteq (\mathbb{R}^n)^\perp$. Suppose $\vec{x} \in \mathbb{R}^n$, $\vec{x} \neq \vec{0}$. Since $\vec{x} \cdot \vec{x} = ||\vec{x}||^2$ and $\vec{x} \neq \vec{0}$, $\vec{x} \cdot \vec{x} \neq 0$, so $\vec{x} \notin (\mathbb{R}^n)^\perp$. Therefore $(\mathbb{R}^n)^\perp \subseteq \{\vec{0}\}$, and thus $(\mathbb{R}^n)^\perp = \{\vec{0}\}$.

In the next example, we will look at how to find $W^\perp$.

Example 4.11.12: Orthogonal Complement

Let W be the plane through the origin given by the equation $x - 2y + z = 0$. Find a basis for the orthogonal complement of W.

Solution

From Example 4.11.11 we know that we can write W as

$$W = \text{span}\{\vec{u}_1, \vec{u}_2\} = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\right\}$$

In order to find $W^\perp$, we need to find all $\vec{x}$ which are orthogonal to every vector in this span.

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. In order to satisfy $\vec{x} \cdot \vec{u}_1 = 0$, the following equation must hold.

$$x_1 - x_3 = 0$$

In order to satisfy $\vec{x} \cdot \vec{u}_2 = 0$, the following equation must hold.

$$x_2 + 2x_3 = 0$$

Both of these equations must be satisfied, so we have the following system of equations.

$$x_1 - x_3 = 0$$
$$x_2 + 2x_3 = 0$$

To solve, set up the augmented matrix.

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right]$$

Using Gaussian Elimination, we find that $W^\perp = \text{span}\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right\}$, and hence $\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right\}$ is a basis for $W^\perp$.

The following results summarize the important properties of the orthogonal projection.

Theorem 4.11.4: Orthogonal Projection

Let W be a subspace of $\mathbb{R}^n$, Y be any point in $\mathbb{R}^n$, and let Z be the point in W closest to Y. Then,

1. The position vector $\vec{z}$ of the point Z is given by $\vec{z} = \text{proj}_W(\vec{y})$
2. $\vec{z} \in W$ and $\vec{y} - \vec{z} \in W^\perp$
3. $|Y - Z| < |Y - Z_1|$ for all $Z_1 \neq Z \in W$

Consider the following example of this concept.

Example 4.11.13: Find a Vector Closest to a Given Vector

Let

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \text{ and } \vec{v} = \begin{bmatrix} 4 \\ 3 \\ -2 \\ 5 \end{bmatrix}.$$

We want to find the vector in $W = \text{span}\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ closest to $\vec{y}$.

Solution

We will first use the Gram-Schmidt Process to construct the orthogonal basis, B, of W:

$$B = \left\{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}\right\}.$$

By Theorem 4.11.4,

$$\text{proj}_U(\vec{v}) = \frac{2}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \frac{5}{1}\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \frac{12}{6}\begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

is the vector in U closest to $\vec{y}$.

Consider the next example.

Example 4.11.14: Vector Written as a Sum of Two Vectors

Let W be a subspace given by $W = \text{span}\left\{\begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\2\end{bmatrix}\right\}$, and $Y = (1,2,3,4)$.

Find the point Z in W closest to Y, and moreover write $\vec{y}$ as the sum of a vector in W and a vector in $W^{\perp}$.

Solution

From Theorem 4.11.3 the point Z in W closest to Y is given by $\vec{z} = \text{proj}_W(\vec{y})$.

Notice that since the above vectors already give an orthogonal basis for W, we have:

$$\begin{aligned}\vec{z} &= \text{proj}_W(\vec{y}) \\ &= \left(\frac{\vec{y}\cdot\vec{w}_1}{\|\vec{w}_1\|^2}\right)\vec{w}_1 + \left(\frac{\vec{y}\cdot\vec{w}_2}{\|\vec{w}_2\|^2}\right)\vec{w}_2 \\ &= \left(\frac{4}{2}\right)\begin{bmatrix}1\\0\\1\\0\end{bmatrix} + \left(\frac{10}{5}\right)\begin{bmatrix}0\\1\\0\\2\end{bmatrix} \\ &= \begin{bmatrix}2\\2\\2\\4\end{bmatrix}\end{aligned}$$

Therefore the point in W closest to Y is $Z = (2,2,2,4)$.
Now, we need to write $\vec{y}$ as the sum of a vector in W and a vector in $W^{\perp}$. This can easily be done as follows:

$$\vec{y} = \vec{z} + (\vec{y} - \vec{z})$$

since $\vec{z}$ is in W and as we have seen $\vec{y} - \vec{z}$ is in $W^{\perp}$.
The vector $\vec{y} - \vec{z}$ is given by

$$\vec{y} - \vec{z} = \begin{bmatrix}1\\2\\3\\4\end{bmatrix} - \begin{bmatrix}2\\2\\2\\4\end{bmatrix} = \begin{bmatrix}-1\\0\\1\\0\end{bmatrix}$$

Therefore, we can write $\vec{y}$ as

$$\begin{bmatrix}1\\2\\3\\4\end{bmatrix} = \begin{bmatrix}2\\2\\2\\4\end{bmatrix} + \begin{bmatrix}-1\\0\\1\\0\end{bmatrix}$$

Example 4.11.15: Point in a Plane Closest to a Given Point

Find the point Z in the plane $3x + y - 2z = 0$ that is closest to the point $Y = (1,1,1)$.

Solution

The solution will proceed as follows.

1. Find a basis X of the subspace W of $\mathbb{R}^3$ defined by the equation $3x + y - 2z = 0$.
2. Orthogonalize the basis X to get an orthogonal basis B of W.
3. Find the projection on W of the position vector of the point Y.

We now begin the solution.

1. $3x+y-2z=0$ is a system of one equation in three variables. Putting the augmented matrix in reduced row-echelon form:

$$\left[\begin{array}{ccc|c} 3 & 1 & -2 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & \frac{1}{3} & -\frac{2}{3} & 0 \end{array}\right]$$

gives general solution $x=\frac{1}{3}s+\frac{2}{3}t$, $y=s$, $z=t$ for any $s,t\in\mathbb{R}$. Then

$$W=\text{span}\left\{\begin{bmatrix} -\frac{1}{3} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ 0 \\ 1 \end{bmatrix}\right\}$$

Let $X=\left\{\begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}\right\}$. Then X is linearly independent and $\text{span}(X)=W$, so X is a basis of W.

2. Use the Gram-Schmidt Process to get an orthogonal basis of W:

$$\vec{f}_1=\begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix} \text{ and } \vec{f}_2=\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}-\frac{-2}{10}\begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}=\frac{1}{5}\begin{bmatrix} 9 \\ 3 \\ 15 \end{bmatrix}.$$

Therefore $B=\left\{\begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}\right\}$ is an orthogonal basis of W.

3. To find the point Z on W closest to $Y=(1,1,1)$, compute

$$\begin{aligned} \text{proj}_W\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} &= \frac{2}{10}\begin{bmatrix} -1 \\ 3 \\ 0 \end{bmatrix}+\frac{9}{35}\begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} \\ &= \frac{1}{7}\begin{bmatrix} 4 \\ 6 \\ 9 \end{bmatrix}. \end{aligned}$$

Therefore, $Z=\left(\frac{4}{7},\frac{6}{7},\frac{9}{7}\right)$.

Least Squares Approximation

It should not be surprising to hear that many problems do not have a perfect solution, and in these cases the objective is always to try to do the best possible. For example what does one do if there are no solutions to a system of linear equations $A\vec{x}=\vec{b}$? It turns out that what we do is find $\vec{x}$ such that $A\vec{x}$ is as close to $\vec{b}$ as possible. A very important technique that follows from orthogonal projections is that of the least square approximation, and allows us to do exactly that.

We begin with a lemma.

Recall that we can form the image of an $m\times n$ matrix A by $\text{im}\,(A)=\{A\vec{x}:\vec{x}\in\mathbb{R}^n\}$. Rephrasing Theorem 4.11.4 using the subspace $W=\text{im}\,(A)$ gives the equivalence of an orthogonality condition with a minimization condition. The following picture illustrates this orthogonality condition and geometric meaning of this theorem.

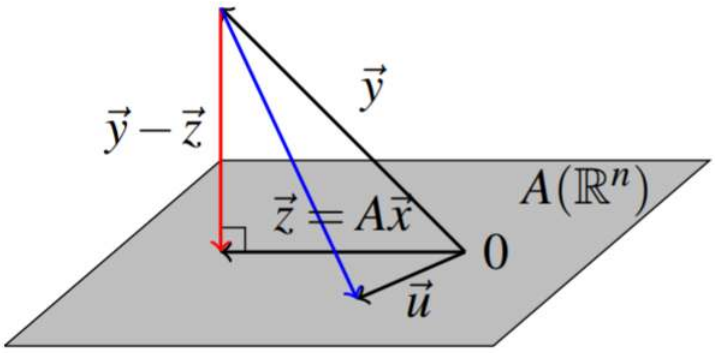


Figure 4.11.2

Theorem 4.11.5:Existence of Minimizers

Let $\vec{y} \in \mathbb{R}^m$ and let A be an $m \times n$ matrix.

Choose $\vec{z} \in W = \text{im}(A)$ given by $\vec{z} = \text{proj}_W(\vec{y})$, and let $\vec{x} \in \mathbb{R}^n$ such that $\vec{z} = A\vec{x}$.

Then

1. $\vec{y} - A\vec{x} \in W^\perp$
2. $\|\vec{y} - A\vec{x}\| < \|\vec{y} - \vec{u}\|$ for all $\vec{u} \neq \vec{z} \in W$

We note a simple but useful observation.

Lemma 4.11.1:Transpose and Dot Product

Let A be an $m \times n$ matrix. Then

$$A\vec{x} \cdot \vec{y} = \vec{x} \cdot A^T\vec{y}$$

Proof

This follows from the definitions:

$$A\vec{x} \cdot \vec{y} = \sum_{i,j} a_{ij}x_j y_i = \sum_{i,j} x_j a_{ji} y_i = \vec{x} \cdot A^T\vec{y}$$

The next corollary gives the technique of least squares.

Corollary 4.11.1:Least Squares and Normal Equation

A specific value of $\vec{x}$ which solves the problem of Theorem 4.11.5 is obtained by solving the equation

$$A^T A\vec{x} = A^T\vec{y}$$

Furthermore, there always exists a solution to this system of equations.

Proof

For $\vec{x}$ the minimizer of Theorem 4.11.5, $(\vec{y} - A\vec{x}) \cdot A\vec{u} = 0$ for all $\vec{u} \in \mathbb{R}^n$ and from Lemma 4.11.1, this is the same as saying

$$A^T(\vec{y} - A\vec{x}) \cdot \vec{u} = 0$$

for all $u \in \mathbb{R}^n$. This implies

$$A^T\vec{y} - A^T A\vec{x} = \vec{0}.$$

Therefore, there is a solution to the equation of this corollary, and it solves the minimization problem of Theorem 4.11.5

Note that $\vec{x}$ might not be unique but $A\vec{x}$, the closest point of $A\left(\mathbb{R}^n\right)$ to $\vec{y}$ is unique as was shown in the above argument.

Consider the following example.

Example 4.11.16: Least Squares Solution to a System

Find a least squares solution to the system

$$\begin{bmatrix} 2 & 1 \\ -1 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

Solution

First, consider whether there exists a real solution. To do so, set up the augmnented matrix given by

$$\left[\begin{array}{cc|c} 2 & 1 & 2 \\ -1 & 3 & 1 \\ 4 & 5 & 1 \end{array}\right]$$

The reduced row-echelon form of this augmented matrix is

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$$

It follows that there is no real solution to this system. Therefore we wish to find the least squares solution. The normal equations are

$$A^T A\vec{x} = A^T\vec{y}$$

$$\begin{bmatrix} 2 & -1 & 4 \\ 1 & 3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 3 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

and so we need to solve the system

$$\begin{bmatrix} 21 & 19 \\ 19 & 35 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \end{bmatrix}$$

This is a familiar exercise and the solution is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{5}{34} \\ \frac{7}{34} \end{bmatrix}$$

Consider another example.

Example 4.11.17: Least Squares Solution to a System

Find a least squares solution to the system

$$\begin{bmatrix} 2 & 1 \\ -1 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 9 \end{bmatrix}$$

Solution

First, consider whether there exists a real solution. To do so, set up the augmnented matrix given by

$$\left[\begin{array}{cc|c} 2 & 1 & 3 \\ -1 & 3 & 2 \\ 4 & 5 & 9 \end{array}\right]$$

The reduced row-echelon form of this augmented matrix is

$$\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$$

It follows that the system has a solution given by $x = y = 1$. However we can also use the normal equations and find the least squares solution.

$$\begin{bmatrix} 2 & -1 & 4 \\ 1 & 3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 3 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 3 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 9 \end{bmatrix}$$

Then

$$\begin{bmatrix} 21 & 19 \\ 19 & 35 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 40 \\ 54 \end{bmatrix}$$

The least squares solution is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

which is the same as the solution found above.

An important application of Corollary 4.11.1 is the problem of finding the least squares regression line in statistics. Suppose you are given points in the xy plane

$$\{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$$

and you would like to find constants m and b such that the line $\vec{y} = m\vec{x} + b$ goes through all these points. Of course this will be impossible in general. Therefore, we try to find m, b such that the line will be as close as possible. The desired system is

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix}$$

which is of the form $\vec{y} = A\vec{x}$. It is desired to choose m and b to make

$$\left\| A \begin{bmatrix} m \\ b \end{bmatrix} - \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\|^2$$

as small as possible. According to Theorem 4.11.5 and Corollary 4.11.1, the best values for m and b occur as the solution to

$$A^T A \begin{bmatrix} m \\ b \end{bmatrix} = A^T \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \text{where } A = \begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}$$

Thus, computing $A^T A$,

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}$$

Solving this system of equations for m and b (using Cramer's rule for example) yields:

$$m = \frac{-\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right) + \left(\sum_{i=1}^{n} x_i y_i\right) n}{\left(\sum_{i=1}^{n} x_i^2\right) n - \left(\sum_{i=1}^{n} x_i\right)^2}$$

and

$$b = \frac{-\left(\sum_{i=1}^{n} x_i\right)\sum_{i=1}^{n} x_i y_i + \left(\sum_{i=1}^{n} y_i\right)\sum_{i=1}^{n} x_i^2}{\left(\sum_{i=1}^{n} x_i^2\right) n - \left(\sum_{i=1}^{n} x_i\right)^2}.$$

Consider the following example.

Example 4.11.18:Least Squares Regression

Find the least squares regression line $\vec{y} = m\vec{x} + b$ for the following set of data points:

$$\{(0,1),(1,2),(2,2),(3,4),(4,5)\}$$

Solution

In this case we have $n = 5$ data points and we obtain:

$$\sum_{i=1}^{5} x_i = 10 \qquad \sum_{i=1}^{5} y_i = 14$$

$$\sum_{i=1}^{5} x_i y_i = 38 \quad \sum_{i=1}^{5} x_i^2 = 30$$

and hence

$$m = \frac{-10 * 14 + 5 * 38}{5 * 30 - 10^2} = 1.00$$

$$b = \frac{-10 * 38 + 14 * 30}{5 * 30 - 10^2} = 0.80$$

The least squares regression line for the set of data points is:

$$\vec{y} = \vec{x} + .8$$

One could use this line to approximate other values for the data. For example for $x = 6$ one could use $y(6) = 6 + .8 = 6.8$ as an approximate value for the data.

The following diagram shows the data points and the corresponding regression line.

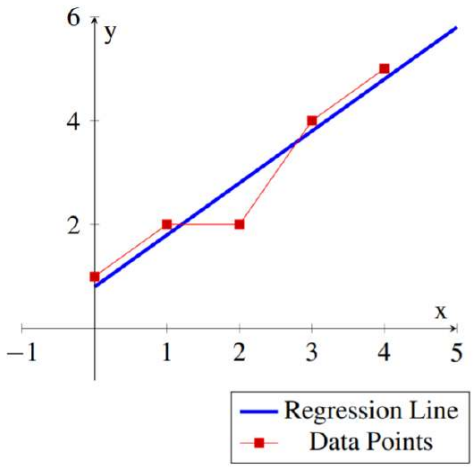


Figure 4.11.3

One could clearly do a least squares fit for curves of the form $y = ax^2 + bx + c$ in the same way. In this case you want to solve as well as possible for a, b, and c the system

$$\begin{bmatrix} x_1^2 & x_1 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

and one would use the same technique as above. Many other similar problems are important, including many in higher dimensions and they are all solved the same way.

This page titled 4.11: Orthogonality is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.11: Orthogonality** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.12: Applications

Applications

Outcomes

A. Apply the concepts of vectors in $\mathbb{R}^n$ to the applications of physics and work.

Vectors and Physics

Suppose you push on something. Then, your push is made up of two components, how hard you push and the direction you push. This illustrates the concept of force.

Definition 4.12.1: Force

Force is a vector. The magnitude of this vector is a measure of how hard it is pushing. It is measured in units such as Newtons or pounds or tons. The direction of this vector is the direction in which the push is taking place.

Vectors are used to model force and other physical vectors like velocity. As with all vectors, a vector modeling force has two essential ingredients, its magnitude and its direction.

Recall the special vectors which point along the coordinate axes. These are given by

$$\vec{e}_i = [0 \cdots 0 \ 1 \ 0 \cdots 0]^T$$

where the 1 is in the ith slot and there are zeros in all the other spaces. The direction of $\vec{e}_i$ is referred to as the ith direction.

Consider the following picture which illustrates the case of $\mathbb{R}^3$. Recall that in $\mathbb{R}^3$, we may refer to these vectors as $\vec{i}, \vec{j}$, and $\vec{k}$.

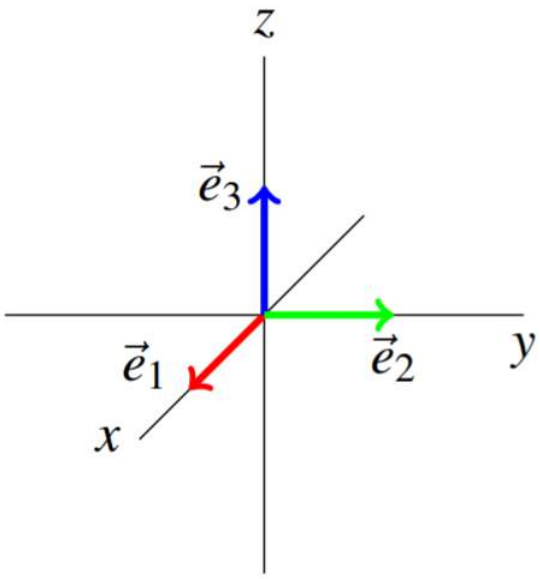


Figure 4.12.1

Given a vector $\vec{u} = [u_1 \cdots u_n]^T$, it follows that

$$\vec{u} = u_1\vec{e}_1 + \cdots + u_n\vec{e}_n = \sum_{k=1}^{n} u_i\vec{e}_i$$

What does addition of vectors mean physically? Suppose two forces are applied to some object. Each of these would be represented by a force vector and the two forces acting together would yield an overall force acting on the object which would also be a force vector known as the resultant. Suppose the two vectors are $\vec{u} = \sum_{k=1}^{n} u_i\vec{e}_i$ and $\vec{v} = \sum_{k=1}^{n} v_i\vec{e}_i$. Then the vector $\vec{u}$ involves a

component in the ith direction given by $u_i\vec{e}_i$, while the component in the ith direction of $\vec{v}$ is $v_i\vec{e}_i$. Then the vector $\vec{u}+\vec{v}$ should have a component in the ith direction equal to $(u_i+v_i)\,\vec{e}_i$. This is exactly what is obtained when the vectors, $\vec{u}$ and $\vec{v}$ are added.

$$\begin{aligned}\vec{u}+\vec{v} &= [u_1+v_1\cdots u_n+v_n]^T \\ &= \sum_{i=1}^{n}(u_i+v_i)\,\vec{e}_i\end{aligned}$$

Thus the addition of vectors according to the rules of addition in $\mathbb{R}^n$ which were presented earlier, yields the appropriate vector which duplicates the cumulative effect of all the vectors in the sum.

Below is a video on finding the horizontal and downward force on a lawnmower handle.



Consider now some examples of vector addition.

Example 4.12.1: The Resultant of Three Forces

There are three ropes attached to a car and three people pull on these ropes. The first exerts a force of $\vec{F}_1=2\vec{i}+3\vec{j}-2\vec{k}$ Newtons, the second exerts a force of $\vec{F}_2=3\vec{i}+5\vec{j}+\vec{k}$ Newtons and the third exerts a force of $5\vec{i}-\vec{j}+2\vec{k}$ Newtons. Find the total force in the direction of $\vec{i}$.

Solution

To find the total force, we add the vectors as described above. This is given by

$$\begin{aligned}&(2\vec{i}+3\vec{j}-2\vec{k})+(3\vec{i}+5\vec{j}+\vec{k})+(5\vec{i}-\vec{j}+2\vec{k}) \\ &=(2+3+5)\vec{i}+(3+5+-1)\vec{j}+(-2+1+2)\vec{k} \\ &=10\vec{i}+7\vec{j}+\vec{k}\end{aligned}$$

Hence, the total force is $10\vec{i}+7\vec{j}+\vec{k}$ Newtons. Therefore, the force in the $\vec{i}$ direction is 10 Newtons.

Below is a video on finding the net and opposite force of three vectors.



Below is a video on finding the resultant force and direction.



Consider another example.

Example 4.12.2:Finding a Vector from Geometric Description

An airplane flies North East at 100 miles per hour. Write this as a vector.

Solution

A picture of this situation follows.

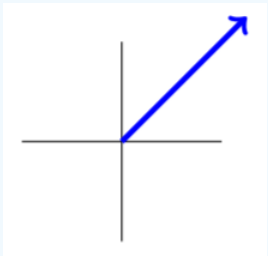


Figure 4.12.2

Therefore, we need to find the vector $\vec{u}$ which has length 100 and direction as shown in this diagram. We can consider the vector $\vec{u}$ as the hypotenuse of a right triangle having equal sides, since the direction of $\vec{u}$ corresponds with the 45° line. The sides, corresponding to the $\vec{i}$ and $\vec{j}$ directions, should be each of length $100/\sqrt{2}$. Therefore, the vector is given by

$$\vec{u} = \frac{100}{\sqrt{2}}\vec{i} + \frac{100}{\sqrt{2}}\vec{j} = \begin{bmatrix} 0.05in\frac{100}{\sqrt{2}} & 0.05in\frac{100}{\sqrt{2}} \end{bmatrix}^T$$

This example also motivates the concept of **velocity**, defined below.

Definition 4.12.2: Speed and Velocity

The **speed** of an object is a measure of how fast it is going. It is measured in units of length per unit time. For example, miles per hour, kilometers per minute, feet per second. The **velocity** is a vector having the speed as the magnitude but also specifying the direction.

Thus the velocity vector in the above example is $0.05in\frac{100}{\sqrt{2}}\vec{i}+0.05in\frac{100}{\sqrt{2}}\vec{j}$ while the speed is 100 miles per hour.

Consider the following example.

Example 4.12.3: Position From Velocity and Time

The velocity of an airplane is $100\vec{i}+\vec{j}+\vec{k}$ measured in kilometers per hour and at a certain instant of time its position is $(1,2,1)$.

Find the position of this airplane one minute later.

Solution

Here imagine a Cartesian coordinate system in which the third component is altitude and the first and second components are measured on a line from West to East and a line from South to North.

Consider the vector $\begin{bmatrix}1 & 2 & 1\end{bmatrix}^T$, which is the initial position vector of the airplane. As the plane moves, the position vector changes according to the velocity vector. After one minute (considered as $\frac{1}{60}$ of an hour) the airplane has moved in the $\vec{i}$ direction a distance of $100\times\frac{1}{60}=\frac{5}{3}$ kilometer. In the $\vec{j}$ direction it has moved $\frac{1}{60}$ kilometer during this same time, while it moves $\frac{1}{60}$ kilometer in the $\vec{k}$ direction. Therefore, the new displacement vector for the airplane is

$$\begin{bmatrix}1 & 2 & 1\end{bmatrix}^T+\begin{bmatrix}\frac{5}{3} & \frac{1}{60} & \frac{1}{60}\end{bmatrix}^T=\begin{bmatrix}\frac{8}{3} & \frac{121}{60} & \frac{121}{60}\end{bmatrix}^T$$

Below is a video on finding the direction and speed of a plane in wind.



Now consider an example which involves combining two velocities.

Example 4.12.4: Sum of Two Velocities

A certain river is one half kilometer wide with a current flowing at 4 kilometers per hour from East to West. A man swims directly toward the opposite shore from the South bank of the river at a speed of 3 kilometers per hour. How far down the river does he find himself when he has swam across? How far does he end up swimming?

Solution

Consider the following picture which demonstrates the above scenario.

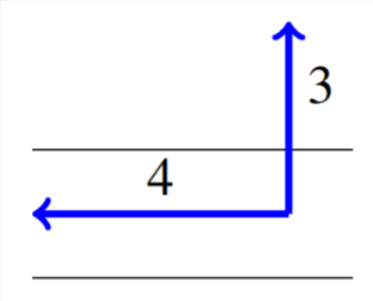


Figure 4.12.3

First we want to know the total time of the swim across the river. The velocity in the direction across the river is 3 kilometers per hour, and the river is $\frac{1}{2}$ kilometer wide. It follows the trip takes $1/6$ hour or 10 minutes.

Now, we can compute how far downstream he will end up. Since the river runs at a rate of 4 kilometers per hour, and the trip takes $1/6$ hour, the distance traveled downstream is given by $4\left(\frac{1}{6}\right)=\frac{2}{3}$ kilometers.

The distance traveled by the swimmer is given by the hypotenuse of a right triangle. The two arms of the triangle are given by the distance across the river, $\frac{1}{2}$km, and the distance traveled downstream, $\frac{2}{3}$ km. Then, using the Pythagorean Theorem, we can calculate the total distance d traveled.

$$d=\sqrt{\left(\frac{2}{3}\right)^2+\left(\frac{1}{2}\right)^2}=\frac{5}{6}\text{km}$$

Therefore, the swimmer travels a total distance of $\frac{5}{6}$ kilometers.

Work

The mathematical concept of work is an application of vectors in $\mathbb{R}^n$. The physical concept of work differs from the notion of work employed in ordinary conversation. For example, suppose you were to slide a 150 pound weight off a table which is three feet high and shuffle along the floor for 50 yards, keeping the height always three feet and then deposit this weight on another three foot high table. The physical concept of work would indicate that the force exerted by your arms did no work during this project. The reason for this definition is that even though your arms exerted considerable force on the weight, the direction of motion was at right angles to the force they exerted. The only part of a force which does work in the sense of physics is the component of the force in the direction of motion.

Work is defined to be the magnitude of the component of this force times the distance over which it acts, when the component of force points in the direction of motion. In the case where the force points in exactly the opposite direction of motion work is given by (-1) times the magnitude of this component times the distance. Thus the work done by a force on an object as the object moves from one point to another is a measure of the extent to which the force contributes to the motion. This is illustrated in the following picture in the case where the given force contributes to the motion.

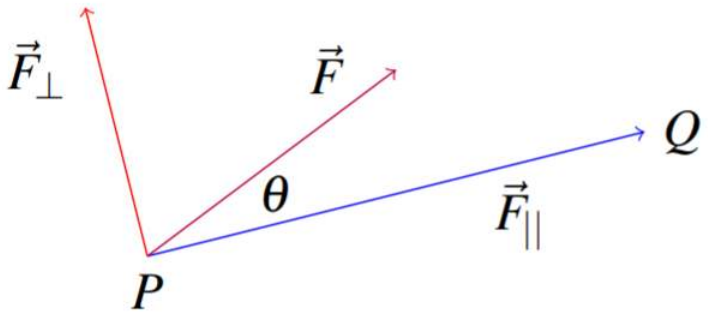


Figure 4.12.4

Recall that for any vector $\vec{u}$ in $\mathbb{R}^n$, we can write $\vec{u}$ as a sum of two vectors, as in

$$\vec{u} = \vec{u}_{||} + \vec{u}_\perp$$

For any force $\vec{F}$, we can write this force as the sum of a vector in the direction of the motion and a vector perpendicular to the motion. In other words,

$$\vec{F} = \vec{F}_{||} + \vec{F}_\perp$$

In the above picture the force, $\vec{F}$ is applied to an object which moves on the straight line from P to Q. There are two vectors shown, $\vec{F}_{||}$ and $\vec{F}_\perp$ and the picture is intended to indicate that when you add these two vectors you get $\vec{F}$. In other words, $\vec{F} = \vec{F}_{||} + \vec{F}_\perp$. Notice that $\vec{F}_{||}$ acts in the direction of motion and $\vec{F}_\perp$ acts perpendicular to the direction of motion. Only $\vec{F}_{||}$ contributes to the work done by $\vec{F}$ on the object as it moves from P to Q. $\vec{F}_{||}$ is called the component of the force in the direction of motion. From trigonometry, you see the magnitude of $\vec{F}_{||}$ should equal $\|\vec{F}\|\,|\cos\theta|$. Thus, since $\vec{F}_{||}$ points in the direction of the vector from P to Q, the total work done should equal

$$\|\vec{F}\|\|\overrightarrow{PQ}\|\cos\theta = \|\vec{F}\|\|\vec{q}-\vec{p}\|\cos\theta$$

Now, suppose the included angle had been obtuse. Then the work done by the force $\vec{F}$ on the object would have been negative because $\vec{F}_{||}$ would point in -1 times the direction of the motion. In this case, $\cos\theta$ would also be negative and so it is still the case that the work done would be given by the above formula. Thus from the geometric description of the dot product given above, the work equals

$$\|\vec{F}\|\|\vec{q}-\vec{p}\|\cos\theta = \vec{F}\bullet(\vec{q}-\vec{p})$$

This explains the following definition.

Definition 4.12.3:Work Done on an Object by a Force

Let $\vec{F}$ be a force acting on an object which moves from the point P to the point Q, which have position vectors given by $\vec{p}$ and $\vec{q}$ respectively. Then the **work** done on the object by the given force equals $\vec{F}\bullet(\vec{q}-\vec{p})$.

Consider the following example.

Example 4.12.4:Finding Work

Let $\vec{F} = [\,2 \quad 7 \quad -3\,]^T$ Newtons. Find the work done by this force in moving from the point $(1,2,3)$ to the point $(-9,-3,4)$ along the straight line segment joining these points where distances are measured in meters.

Solution

First, compute the vector $\vec{q}-\vec{p}$, given by

$$[\,-9 \quad -3 \quad 4\,]^T - [\,1 \quad 2 \quad 3\,]^T = [\,-10 \quad -5 \quad 1\,]^T$$

According to Definition 4.12.3 the work done is

$$\begin{aligned}\begin{bmatrix}2 & 7 & 3\end{bmatrix}^T \bullet \begin{bmatrix}-10 & -5 & 1\end{bmatrix}^T &= -20+(-35)+(-3)\\ &= -58 \text{ Newton meters}\end{aligned}$$

Note that if the force had been given in pounds and the distance had been given in feet, the units on the work would have been foot pounds. In general, work has units equal to units of a force times units of a length. Recall that 1 Newton meter is equal to 1 Joule. Also notice that the work done by the force can be negative as in the above example.

Below is a video on force and work.



This page titled 4.12: Applications is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.12: Applications** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.E: Exercises

Exercise 4.E.1

Find $-3\begin{bmatrix}5\\-1\\2\\-3\end{bmatrix}+5\begin{bmatrix}-8\\2\\-3\\6\end{bmatrix}$.

Answer

$$\begin{bmatrix}-55\\13\\-21\\39\end{bmatrix}$$

Exercise 4.E.2

Find $-7\begin{bmatrix}6\\0\\4\\-1\end{bmatrix}+6\begin{bmatrix}-13\\-1\\1\\6\end{bmatrix}$.

Exercise 4.E.3

Decide whether

$$\vec{v}=\begin{bmatrix}4\\4\\-3\end{bmatrix}$$

is a linear combination of the vectors

$$\vec{u}_1=\begin{bmatrix}3\\1\\-1\end{bmatrix} \quad \text{and} \quad \vec{u}_2=\begin{bmatrix}2\\-2\\1\end{bmatrix}.$$

Answer

$$\begin{bmatrix}4\\4\\-3\end{bmatrix}=2\begin{bmatrix}3\\1\\-1\end{bmatrix}-\begin{bmatrix}2\\-2\\1\end{bmatrix}$$

Exercise 4.E.4

Decide whether

$$\vec{v}=\begin{bmatrix}4\\4\\4\end{bmatrix}$$

is a linear combination of the vectors

$$\vec{u}_1 = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} \quad \text{and} \quad \vec{u}_2 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}.$$

Answer

The system

$$\begin{bmatrix} 4 \\ 4 \\ 4 \end{bmatrix} = a_1 \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} + a_2 \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

has no solution.

Exercise 4.E.5

Find the vector equation for the line through $(-7, 6, 0)$ and $(-1, 1, 4)$. Then, find the parametric equations for this line.

Exercise 4.E.6

Find parametric equations for the line through the point $(7, 7, 1)$ with a direction vector $\vec{d} = \begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix}$.

Exercise 4.E.7

Parametric equations of the line are

$$\begin{aligned} x &= t + 2 \\ y &= 6 - 3t \\ x &= -t = 6 \end{aligned}$$

Find a direction vector for the line and a point on the line.

Exercise 4.E.8

Find the vector equation for the line through the two points $(-5, 5, 1),\ (2, 2, 4)$. Then, find the parametric equations.

Exercise 4.E.9

The equation of a line in two dimensions is written as $y = x - 5$. Find parametric equations for this line.

Exercise 4.E.10

Find parametric equations for the line through $(6, 5, -2)$ and $(5, 1, 2)$.

Exercise 4.E.11

Find the vector equation and parametric equations for the line through the point $(-7, 10, -6)$ with a direction vector $\vec{d} = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$.

Exercise 4.E.12

Parametric equations of the line are

$$\begin{aligned} x &= 2t+2 \\ y &= 5-4t \\ z &= -t-3 \end{aligned}$$

Find a direction vector for the line and a point on the line, and write the vector equation of the line.

Exercise 4.E.13

Find the vector equation and parametric equations for the line through the two points $(4,10,0),\ (1,-5,-6)$

Exercise 4.E.14

Find the point on the line segment from $P=(-4,7,5)$ to $Q=(2,-2,-3)$ which is $\frac{1}{7}$ of the way from P to Q.

Exercise 4.E.15

Suppose a triangle in $\mathbb{R}^n$ has vertices at $P_1,\ P_2$, and P_3. Consider the lines which are drawn from a vertex to the mid point of the opposite side. Show these three lines intersect in a point and find the coordinates of this point.

Exercise 4.E.16

Find $\begin{bmatrix}1\\2\\3\\4\end{bmatrix} \bullet \begin{bmatrix}2\\0\\1\\3\end{bmatrix}$.

Answer

$$\begin{bmatrix}1\\2\\3\\4\end{bmatrix} \bullet \begin{bmatrix}2\\0\\1\\3\end{bmatrix} = 17$$

Exercise 4.E.17

Use the formula given in Proposition 4.7.2 to verify the Cauchy Schwarz inequality and to show that equality occurs if and only if one of the vectors is a scalar multiple of the other

Answer

This formula says that $\vec{u} \bullet \vec{v} = ||\vec{u}||\ ||\vec{v}|| \cos\theta$ where θ is the included angle between the two vectors. Thus

$$||\vec{u} \bullet \vec{v}|| = ||\vec{u}||\ ||\vec{v}||\ ||\cos\theta|| \leq ||\vec{u}||\ ||\vec{v}||$$

and equality holds if and only if $\theta = 0$ or π. This means that the two vectors either point in the same direction or opposite directions. Hence one is a multiple of the other.

Exercise 4.E.18

For $\vec{u}$, $\vec{v}$ vectors in $\mathbb{R}^3$, define the product, $\vec{u} * \vec{v} = u_1v_1 + 2u_2v_2 + 3u_3v_3$. Show the axioms for a dot product all hold for this product. Prove

$$||\vec{u} * \vec{v}|| \leq (\vec{u} * \vec{u})^{1/2}(\vec{v} * \vec{v})^{1/2}$$

Answer

This follows from the Cauchy Schwarz inequality and the proof of Theorem 4.7.1 which only used the properties of the dot product. Since this new product has the same properties the Cauchy Schwarz inequality holds for it as well.

Exercise 4.E.19

Let $\vec{a}$, $\vec{b}$ be vectors. Show that $\left(\vec{a} \bullet \vec{b}\right) = \frac{1}{4}\left(||\vec{a}+\vec{b}||^2 - ||\vec{a}-\vec{b}||^2\right)$.

Exercise 4.E.20

Using the axioms of the dot product, prove the parallelogram identity:

$$||\vec{a}+\vec{b}||^2 + ||\vec{a}-\vec{b}||^2 = 2||\vec{a}||^2 + 2||\vec{b}||^2$$

Exercise 4.E.21

Let A be a real $m \times n$ matrix and let $\vec{u} \in \mathbb{R}^n$ and $\vec{v} \in \mathbb{R}^m$. Show $A\vec{u} \bullet \vec{v} = \vec{u} \bullet A^T\vec{v}$. Hint: Use the definition of matrix multiplication to do this.

Answer

$A\vec{x} \bullet \vec{y} = \sum_k (A\vec{x})_k y_k = \sum_k \sum_i A_{ki}x_i y_k = \sum_i \sum_k A^T_{ik} x_i y_k = \vec{x} \bullet A^T\vec{y}$

Exercise 4.E.22

Use the result of Problem 4.E.21 to verify directly that $(AB)^T = B^TA^T$ without making any reference to subscripts.

Answer

$$\begin{aligned} AB\vec{x} \bullet \vec{y} &= B\vec{x} \bullet A^T\vec{y} \\ &= \vec{x} \bullet B^TA^T\vec{y} \\ &= \vec{x} \bullet (AB)^T\vec{y} \end{aligned}$$

Since this is true for all $\vec{x}$, it follows that, in particular, it holds for

$$\vec{x} = B^TA^T\vec{y} - (AB)^T\vec{y}$$

and so from the axioms of the dot product,

$$\left(B^TA^T\vec{y} - (AB)^T\vec{y}\right) \bullet \left(B^TA^T\vec{y} - (AB)^T\vec{y}\right) = 0$$

and so $B^TA^T\vec{y} - (AB)^T\vec{y} = \vec{0}$. However, this is true for all $\vec{y}$ and so $B^TA^T - (AB)^T = 0$.

Exercise 4.E.23

Find the angle between the vectors

$$\vec{u} = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}, \vec{v} = \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$$

Answer

$\frac{[\,3 \quad -1 \quad -1\,]^T \bullet [\,1 \quad 4 \quad 2\,]^T}{\sqrt{9+1+1}\sqrt{1+16+4}} = \frac{-3}{\sqrt{11}\sqrt{21}} = -0.19739 = \cos\theta$ Therefore we need to solve

$$-0.19739 = \cos\theta$$

Thus $\theta = 1.7695$ radians.

Exercise $4.E.24$

Find the angle between the vectors

$$\vec{u} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \vec{v} = \begin{bmatrix} 1 \\ 2 \\ -7 \end{bmatrix}$$

Answer

$\frac{-10}{\sqrt{1+4+1}\sqrt{1+4+49}} = -0.55555 = \cos\theta$ Therefore we need to solve $-0.55555 = \cos\theta$, which gives $\theta = 2.0313$ radians.

Exercise $4.E.25$

Find $\text{proj}_{\vec{v}}(\vec{w})$ where $\vec{w} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Answer

$$\frac{\vec{u}\bullet\vec{v}}{\vec{u}\bullet\vec{u}}\vec{u} = \frac{-5}{14}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{5}{14} \\ -\frac{5}{7} \\ -\frac{15}{14} \end{bmatrix}$$

Exercise $4.E.26$

Find $\text{proj}_{\vec{v}}(\vec{w})$ where $\vec{w} = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$.

Answer

$$\frac{\vec{u}\bullet\vec{v}}{\vec{u}\bullet\vec{u}}\vec{u} = \frac{-5}{10}\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ -\frac{3}{2} \end{bmatrix}$$

Exercise $4.E.27$

Find $\text{proj}_{\vec{v}}(\vec{w})$ where $\vec{w} = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix}$.

Answer

$$\frac{\vec{u} \bullet \vec{v}}{\vec{u} \bullet \vec{u}} \vec{u} = \frac{\begin{bmatrix} 1 & 2 & -2 & 1 \end{bmatrix}^T \bullet \begin{bmatrix} 1 & 2 & 3 & 0 \end{bmatrix}^T}{1+4+9} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{14} \\ -\frac{1}{7} \\ -\frac{3}{14} \\ 0 \end{bmatrix}$$

Exercise 4.E.28

Let $P = (1, 2, 3)$ be a point in $\mathbb{R}^3$. Let L be the line through the point $P_0 = (1, 4, 5)$ with direction vector $\vec{d} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$. Find the shortest distance from P to L, and find the point Q on L that is closest to P.

Exercise 4.E.29

Let $P = (0, 2, 1)$ be a point in $\mathbb{R}^3$. Let L be the line through the point $P_0 = (1, 1, 1)$ with direction vector $\vec{d} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$. Find the shortest distance from P to L, and find the point Q on L that is closest to P.

Exercise 4.E.30

Does it make sense to speak of $\text{proj}_{\vec{0}}(\vec{w})$?

Answer

No, it does not. The 0 vector has no direction. The formula for $\text{proj}_{\vec{0}}(\vec{w})$ doesn't make sense either.

Exercise 4.E.31

Prove the Cauchy Schwarz inequality in $\mathbb{R}^n$ as follows. For $\vec{u}$,$\vec{v}$ vectors, consider

$$(\vec{w} - \text{proj}_{\vec{v}}\vec{w}) \bullet (\vec{w} - \text{proj}_{\vec{v}}\vec{w}) \geq 0$$

Simplify using the axioms of the dot product and then put in the formula for the projection. Notice that this expression equals 0 and you get equality in the Cauchy Schwarz inequality if and only if $\vec{w} = \text{proj}_{\vec{v}}\vec{w}$. What is the geometric meaning of $\vec{w} = \text{proj}_{\vec{v}}\vec{w}$?

Answer

$$\left(\vec{u} - \frac{\vec{u} \bullet \vec{v}}{||\vec{v}||^2}\vec{v}\right) \bullet \left(\vec{u} - \frac{\vec{u} \bullet \vec{v}}{||\vec{v}||^2}\vec{v}\right) = ||\vec{u}||^2 - 2(\vec{u} \bullet \vec{v})^2 \frac{1}{||\vec{v}||^2} + (\vec{u} \bullet \vec{v})^2 \frac{1}{||\vec{v}||^2} \geq 0$$

And so

$$||\vec{u}||^2 ||\vec{v}||^2 \geq (\vec{u} \bullet \vec{v})^2$$

You get equality exactly when $\vec{u} = \text{proj}_{\vec{v}}\vec{u} = \frac{\vec{u} \bullet \vec{v}}{||\vec{v}||^2}\vec{v}$ in other words, when $\vec{u}$ is a multiple of $\vec{v}$.

Exercise 4.E.32

Let $\vec{v},\ \vec{w},\ \vec{u}$ be vectors. Show that $(\vec{w}+\vec{u})_\perp = \vec{w}_\perp + \vec{u}_\perp$ where $\vec{w}_\perp = \vec{w} - \text{proj}_{\vec{v}}(\vec{w})$.

Answer

$$\begin{aligned}\vec{w} - \text{proj}_{\vec{v}}(\vec{w}) + \vec{u} - \text{proj}_{\vec{v}}(\vec{u}) &= \vec{w} + \vec{u} - (\text{proj}_{\vec{v}}(\vec{w}) + \text{proj}_{\vec{v}}(\vec{u})) \\ &= \vec{w} + \vec{u} - \text{proj}_{\vec{v}}(\vec{w} + \vec{u})\end{aligned}$$

This follows because

$$\begin{aligned}\text{proj}_{\vec{v}}(\vec{w}) + \text{proj}_{\vec{v}}(\vec{u}) &= \frac{\vec{u} \bullet \vec{v}}{||\vec{v}||^2}\vec{v} + \frac{\vec{w} \bullet \vec{v}}{||\vec{v}||^2}\vec{v} \\ &= \frac{(\vec{u} + \vec{w}) \bullet \vec{v}}{||\vec{v}||^2}\vec{v} \\ &= \text{proj}_{\vec{v}}(\vec{w} + \vec{u})\end{aligned}$$

Exercise 4.E.33

Show that

$$(\vec{v} - \text{proj}_{\vec{u}}(\vec{v}), \vec{u}) = (\vec{v} - \text{proj}_{\vec{u}}(\vec{v})) \bullet \vec{u} = 0$$

and conclude every vector in $\mathbb{R}^n$ can be written as the sum of two vectors, one which is perpendicular and one which is parallel to the given vector.

Answer

$(\vec{v} - \text{proj}_{\vec{u}}(\vec{v})) \bullet \vec{u} = \vec{v} \bullet \vec{u} - \left(\frac{(\vec{v} \cdot \vec{u}}{||\vec{u}||^2}\vec{u}\right) \bullet \vec{u} = \vec{v} \bullet \vec{u} - \vec{v} \bullet \vec{u} = 0$. Therefore, $\vec{v} = \vec{v} - \text{proj}_{\vec{u}}(\vec{v}) + \text{proj}_{\vec{u}}(\vec{v})$. The first is perpendicular to $\vec{u}$ and the second is a multiple of $\vec{u}$ so it is parallel to $\vec{u}$.

Exercise 4.E.34

Show that if $\vec{a} \times \vec{u} = \vec{0}$ for any unit vector $\vec{u}$, then $\vec{a} = \vec{0}$.

Answer

If $\vec{a} \neq \vec{0}$, then the condition says that $||\vec{a} \times \vec{u}|| = ||\vec{a}|| \sin\theta = 0$ for all angles θ. Hence $\vec{a} = \vec{0}$ after all.

Exercise 4.E.35

Find the area of the triangle determined by the three points $(1,2,3),\ (4,2,0)$and $(-3,2,1)$.

Answer

$\begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix} \times \begin{bmatrix} -4 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 18 \\ 0 \end{bmatrix}$. So the area is 9.

Exercise 4.E.36

Find the area of the triangle determined by the three points $(1,0,3),\ (4,1,0)$and $(-3,1,1)$.

Answer

$\begin{bmatrix} 3 \\ 1 \\ -3 \end{bmatrix} \times \begin{bmatrix} -4 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 18 \\ 7 \end{bmatrix}$. The area is given by

$$\frac{1}{2}\sqrt{1+(18)^2+49} = \frac{1}{2}\sqrt{374}$$

Exercise 4.E.37

Find the area of the triangle determined by the three points, $(1,2,3)$, $(2,3,4)$and $(3,4,5)$. Did something interesting happen here? What does it mean geometrically?

Answer

$[1 \quad 1 \quad 1] \times [2 \quad 2 \quad 2] = [0 \quad 0 \quad 0]$. The area is 0. It means the three points are on the same line.

Exercise 4.E.38

Find the area of the parallelogram determined by the vectors $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}$.

Answer

$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \times \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 8 \\ -8 \end{bmatrix}$. The area is $8\sqrt{3}$.

Exercise 4.E.39

Find the area of the parallelogram determined by the vectors $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$.

Answer

$\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} \times \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \\ -2 \end{bmatrix}$. The area is $\sqrt{36+121+4} = \sqrt{161}$.

Exercise 4.E.40

Is $\vec{u} \times (\vec{v} \times \vec{w}) = (\vec{u} \times \vec{v}) \times \vec{w}$? What is the meaning of $\vec{u} \times \vec{v} \times \vec{w}$? Explain. Hint: Try $\left(\vec{i} \times \vec{j}\right) \times \vec{k}$.

Answer

$\left(\vec{i} \times \vec{j}\right) \times \vec{j} = \vec{k} \times \vec{j} = i\vec{i}$. However, $\vec{i} \times \left(\vec{j} \times \vec{j}\right) = \vec{0}$ and so the cross product is not associative.

Exercise 4.E.41

Verify directly that the coordinate description of the cross product, $\vec{u} \times \vec{v}$ has the property that it is perpendicular to both $\vec{u}$ and $\vec{v}$. Then show by direct computation that this coordinate description satisfies

$$\begin{aligned} ||\vec{u}\times\vec{v}||^2 &= ||\vec{u}||^2||\vec{v}||^2 - (\vec{u}\bullet\vec{v})^2 \\ &= ||\vec{u}||^2||\vec{v}||^2(1-\cos^2(\theta)) \end{aligned}$$

where θ is the angle included between the two vectors. Explain why $||\vec{u}\times\vec{v}||$ has the correct magnitude.

Answer

Verify directly from the coordinate description of the cross product that the right hand rule applies to the vectors $\vec{i},\vec{j},\vec{k}$. Next verify that the distributive law holds for the coordinate description of the cross product. This gives another way to approach the cross product. First define it in terms of coordinates and then get the geometric properties from this. However, this approach does not yield the right hand rule property very easily. From the coordinate description,

$$\vec{a}\times\vec{b}\cdot\vec{a} = \epsilon_{ijk}a_jb_ka_i = -\epsilon_{jik}a_kb_ka_i = -\epsilon_{jik}b_ka_ia_j = -\vec{a}\times\vec{b}\cdot\vec{a}$$

and so $\vec{a}\times\vec{b}$ is perpendicular to $\vec{a}$. Similarly, $\vec{a}\times\vec{b}$ is perpendicular to $\vec{b}$. Now we need that

$$||\vec{a}\times\vec{b}||^2 = ||\vec{a}||^2||\vec{b}||^2(1-\cos^2\theta) = ||\vec{a}||^2||\vec{b}||^2\sin^2\theta$$

and so $||\vec{a}\times\vec{b}|| = ||\vec{a}||\,||\vec{b}||\sin\theta$, the area of the parallelogram determined by $\vec{a}$, $\vec{b}$. Only the right hand rule is a little problematic. However, you can see right away from the component definition that the right hand rule holds for each of the standard unit vectors. Thus $\vec{i}\times\vec{j} = \vec{k}$ etc.

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \vec{k}$$

? Exercise 4.E. 42

Suppose A is a 3×3 skew symmetric matrix such that $A^T = -A$. Show there exists a vector $\vec{\Omega}$ such that for all $\vec{u}\in\mathbb{R}^3$

$$A\vec{u} = \vec{\Omega}\times\vec{u}$$

Hint: Explain why since A is skew symmetric it is of the form

$$A = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}$$

where the ω_i are numbers. Then consider $\omega_1\vec{i}+\omega_2\vec{j}+\omega_3\vec{k}$.

? Exercise 4.E. 43

Find the volume of the parallelepiped determined by the vectors $\begin{bmatrix} 1 \\ -7 \\ -5 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -2 \\ -6 \end{bmatrix}$, and $\begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}$.

Answer

$$\begin{vmatrix} 1 & -7 & -5 \\ 1 & -2 & -6 \\ 3 & 2 & 3 \end{vmatrix} = 113$$

Exercise 4.E.44

Suppose $\vec{u}$, $\vec{v}$, and $\vec{w}$ are three vectors whose components are all integers. Can you conclude the volume of the parallelepiped determined from these three vectors will always be an integer?

Answer

Yes. It will involve the sum of product of integers and so it will be an integer.

Exercise 4.E.45

What does it mean geometrically if the box product of three vectors gives zero?

Answer

It means that if you place them so that they all have their tails at the same point, the three will lie in the same plane.

Exercise 4.E.46

Using Problem 4.E.45, find an equation of a plane containing the two position vectors, $\vec{p}$ and $\vec{q}$ and the point 0. Hint: If (x, y, z) is a point on this plane, the volume of the parallelepiped determined by (x, y, z) and the vectors $\vec{p}$, $\vec{q}$ equals 0.

Answer

$$\vec{x} \bullet \left(\vec{a} \times \vec{b}\right) = 0$$

Exercise 4.E.47

Using the notion of the box product yielding either plus or minus the volume of the parallelepiped determined by the given three vectors, show that

$$(\vec{u} \times \vec{v}) \bullet \vec{w} = \vec{u} \bullet (\vec{v} \times \vec{w})$$

In other words, the dot and the cross can be switched as long as the order of the vectors remains the same. Hint: There are two ways to do this, by the coordinate description of the dot and cross product and by geometric reasoning.

Exercise 4.E.48

Simplify $(\vec{u} \times \vec{v}) \bullet (\vec{v} \times \vec{w}) \times (\vec{w} \times \vec{z})$.

Answer

Here $[\vec{v}, \vec{w}, \vec{z}]$ denotes the box product. Consider the cross product term. From the above,

$$\begin{aligned}(\vec{v} \times \vec{w}) \times (\vec{w} \times \vec{z}) &= [\vec{v}, \vec{w}, \vec{z}]\vec{w} - [\vec{w}, \vec{w}, \vec{z}]\vec{v} \\ &= [\vec{v}, \vec{w}, \vec{z}]\vec{w}\end{aligned}$$

Thus it reduces to

$$(\vec{u} \times \vec{v}) \bullet [\vec{v}, \vec{w}, \vec{z}]\vec{w} = [\vec{v}, \vec{w}, \vec{z}][\vec{u}, \vec{v}, \vec{w}]$$

Exercise 4.E.49

Simplify $||\vec{u} \times \vec{v}||^2 + (\vec{u} \bullet \vec{v})^2 - ||\vec{u}||^2||\vec{v}||^2$.

Answer

$$\begin{aligned} ||\vec{u}\times\vec{v}||^2 &= \epsilon_{ijk}u_jv_k\epsilon_{irs}u_rv_s = (\delta_{jr}\delta_{ks}-\delta_{kr}\delta_{js})u_rv_su_jv_k \\ &= u_jv_ku_jv_k - u_kv_ju_jv_k = ||\vec{u}||^2||\vec{v}||^2 - (\vec{u}\bullet\vec{v})^2 \end{aligned}$$

It follows that the expression reduces to 0. You can also do the following.

$$\begin{aligned} ||\vec{u}\times\vec{v}||^2 &= ||\vec{u}||^2||\vec{v}||^2\sin^2\theta \\ &= ||\vec{u}||^2||\vec{v}||^2(1-\cos^2\theta) \\ &= ||\vec{u}||^2||\vec{v}||^2 - ||\vec{u}||^2||\vec{v}||^2\cos^2\theta \\ &= ||\vec{u}||^2||\vec{v}||^2 - (\vec{u}\bullet\vec{v})^2 \end{aligned}$$

which implies the expression equals 0.

Exercise 4.E.50

For $\vec{u},\ \vec{v},\ \vec{w}$ functions of t, prove the following product rules:

$$\begin{aligned} (\vec{u}\times\vec{v})' &= \vec{u}'\times\vec{v}+\vec{u}\times\vec{v}' \\ (\vec{u}\bullet\vec{v})' &= \vec{u}'\bullet\vec{v}+\vec{u}\bullet\vec{v}' \end{aligned}$$

Answer

We will show it using the summation convention and permutation symbol

$$\begin{aligned} ((\vec{u}\times\vec{v})')_i &= ((\vec{u}\times\vec{v})_i)' = (\epsilon_{ijk}u_jv_k)' \\ &= \epsilon_{ijk}u_j'v_k+\epsilon_{ijk}u_kv_k' = (\vec{u}'\times\vec{v}+\vec{u}\times\vec{v}')_i \end{aligned}$$

and so $(\vec{u}\times\vec{v})' = \vec{u}'\times\vec{v}+\vec{u}\times\vec{v}'$.

Exercise 4.E.51

Here are some vectors.

$$\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 7 \\ -4 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \\ -10 \end{bmatrix}, \begin{bmatrix} 12 \\ 17 \\ -24 \end{bmatrix}$$

Describe the span of these vectors as the span of as few vectors as possible.

Exercise 4.E.52

Here are some vectors.

$$\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 12 \\ 29 \\ -24 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 9 \\ -4 \end{bmatrix}, \begin{bmatrix} 5 \\ 12 \\ -10 \end{bmatrix}.$$

Describe the span of these vectors as the span of as few vectors as possible.

Exercise 4.E.53

Here are some vectors.

$$\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$$

Describe the span of these vectors as the span of as few vectors as possible.

Exercise $4.E.54$

Here are some vectors.

$$\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

Now here is another vector:

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Is this vector in the span of the first four vectors? If it is, exhibit a linear combination of the first four vectors which equals this vector, using as few vectors as possible in the linear combination.

Exercise $4.E.55$

Here are some vectors.

$$\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$$

Now here is another vector:

$$\begin{bmatrix} 2 \\ -3 \\ -4 \end{bmatrix}$$

Is this vector in the span of the first four vectors? If it is, exhibit a linear combination of the first four vectors which equals this vector, using as few vectors as possible in the linear combination.

Exercise $4.E.56$

Here are some vectors.

$$\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Now here is another vector:

$$\begin{bmatrix} 1 \\ 9 \\ 1 \end{bmatrix}$$

Is this vector in the span of the first four vectors? If it is, exhibit a linear combination of the first four vectors which equals this vector, using as few vectors as possible in the linear combination.

Exercise $4.E.57$

Here are some vectors,

$$\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -5 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 2 \end{bmatrix}$$

Now here is another vector:

$$\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Is this vector in the span of the first four vectors? If it is, exhibit a linear combination of the first four vectors which equals this vector, using as few vectors as possible in the linear combination.

? Exercise 4.E.58

Here are some vectors.

$$\begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -5 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ 2 \end{bmatrix}$$

Now here is another vector:

$$\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

Is this vector in the span of the first four vectors? If it is, exhibit a linear combination of the first four vectors which equals this vector, using as few vectors as possible in the linear combination.

? Exercise 4.E.59

Here are some vectors.

$$\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ -3 \end{bmatrix}, \begin{bmatrix} -1 \\ 4 \\ 2 \end{bmatrix}$$

Now here is another vector:

$$\begin{bmatrix} -1 \\ -4 \\ 2 \end{bmatrix}$$

Is this vector in the span of the first four vectors? If it is, exhibit a linear combination of the first four vectors which equals this vector, using as few vectors as possible in the linear combination.

? Exercise 4.E.60

Suppose $\{\vec{x}_1, \cdots, \vec{x}_k\}$ is a set of vectors from $\mathbb{R}^n$. Show that $\vec{0}$ is in $span\{\vec{x}_1, \cdots, \vec{x}_k\}$.

Answer

$$\sum_{i=1}^{k} 0\vec{x}_k = \vec{0}$$

Exercise 4.E.61

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others. Also give a linearly independent set of vectors which has the same span as the given vectors.

$$\begin{bmatrix} 1 \\ 3 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 10 \\ 2 \\ 1 \end{bmatrix}$$

Exercise 4.E.62

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others. Also give a linearly independent set of vectors which has the same span as the given vectors.

$$\begin{bmatrix} -1 \\ -2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -3 \\ -4 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 6 \\ 4 \end{bmatrix}$$

Exercise 4.E.63

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others. Also give a linearly independent set of vectors which has the same span as the given vectors.

$$\begin{bmatrix} 1 \\ 5 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \\ -2 \\ 1 \end{bmatrix}$$

Exercise 4.E.64

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others. Also give a linearly independent set of vectors which has the same span as the given vectors.

$$\begin{bmatrix} 1 \\ -1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \\ 34 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 7 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 8 \\ 1 \end{bmatrix}$$

Exercise 4.E.65

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others.

$$\begin{bmatrix} 1 \\ 3 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ -10 \\ 3 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

Exercise 4.E.66

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others. Also give a linearly independent set of vectors which has the same span as the given vectors.

$$\begin{bmatrix} 1 \\ 3 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ -4 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 10 \\ -14 \\ 1 \end{bmatrix}$$

Exercise 4.E.67

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others. Also give a linearly independent set of vectors which has the same span as the given vectors.

$$\begin{bmatrix} 1 \\ 0 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 8 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 7 \\ 34 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 7 \\ 1 \end{bmatrix}$$

Exercise 4.E.68

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others. Also give a linearly independent set of vectors which has the same span as the given vectors.

$$\begin{bmatrix} 1 \\ 4 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 7 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -2 \\ 1 \end{bmatrix}$$

Exercise 4.E.69

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others.

$$\begin{bmatrix} 1 \\ 2 \\ 2 \\ -4 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 1 \\ -4 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -2 \\ 5 \end{bmatrix}$$

Exercise 4.E.70

Are the following vectors linearly independent? If they are, explain why and if they are not, exhibit one of them as a linear combination of the others. Also give a linearly independent set of vectors which has the same span as the given vectors.

$$\begin{bmatrix} 2 \\ 3 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} -5 \\ -6 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 0 \\ 4 \end{bmatrix}$$

Exercise 4.E.71

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise 4.E.72

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -2 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise 4.E.73

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -5 \\ -7 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise 4.E.74

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ -2 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise 4.E.75

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ 4 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 11 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -3 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise $4.E.76$

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ 3 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{3}{2} \\ -\frac{9}{2} \\ \frac{3}{2} \\ -\frac{3}{2} \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise $4.E.77$

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ 3 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 0 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors

Exercise $4.E.78$

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ 4 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ -2 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise $4.E.79$

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ -1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 7 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 8 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -9 \\ -6 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 8 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise 4.E.80

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -3 \\ 3 \\ 3 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -9 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise 4.E.81

Here are some vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ b+1 \\ a \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 3b+3 \\ 3a \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ b+2 \\ 2a+1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2b-5 \\ -5a-7 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ b+2 \\ 2a+2 \\ 1 \end{bmatrix}$$

These vectors can't possibly be linearly independent. Tell why. Next obtain a linearly independent subset of these vectors which has the same span as these vectors. In other words, find a basis for the span of these vectors.

Exercise 4.E.82

Let $H = span\left\{ \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -2 \\ -2 \end{bmatrix} \right\}$. Find the dimension of H and determine a basis.

Exercise 4.E.83

Let H denote $span\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 5 \\ -5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ -2 \end{bmatrix} \right\}$. Find the dimension of H and determine a basis.

Exercise 4.E.84

Let H denote $span\left\{ \begin{bmatrix} -2 \\ 1 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} -9 \\ 4 \\ 3 \\ -9 \end{bmatrix}, \begin{bmatrix} -33 \\ 15 \\ 12 \\ -36 \end{bmatrix}, \begin{bmatrix} -22 \\ 10 \\ 8 \\ -24 \end{bmatrix} \right\}$. Find the dimension of H and determine a basis.

Exercise 4.E.85

Let H denote $span\left\{ \begin{bmatrix} -1 \\ 1 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} -4 \\ 3 \\ -2 \\ -4 \end{bmatrix}, \begin{bmatrix} -3 \\ 2 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -2 \\ -4 \end{bmatrix}, \begin{bmatrix} -7 \\ 5 \\ -3 \\ -6 \end{bmatrix} \right\}$. Find the dimension of H and determine a basis.

Exercise 4.E.86

Let H denote $span\left\{\begin{bmatrix}2\\3\\2\\1\end{bmatrix},\begin{bmatrix}8\\15\\6\\3\end{bmatrix},\begin{bmatrix}3\\6\\2\\1\end{bmatrix},\begin{bmatrix}4\\6\\6\\3\end{bmatrix},\begin{bmatrix}8\\15\\6\\3\end{bmatrix}\right\}$. Find the dimension of H and determine a basis.

Exercise 4.E.87

Let H denote $span\left\{\begin{bmatrix}0\\2\\0\\-1\end{bmatrix},\begin{bmatrix}-1\\6\\0\\-2\end{bmatrix},\begin{bmatrix}-2\\16\\0\\-6\end{bmatrix},\begin{bmatrix}-3\\22\\0\\-8\end{bmatrix}\right\}$. Find the dimension of H and determine a basis.

Exercise 4.E.88

Let H denote $span\left\{\begin{bmatrix}5\\1\\1\\4\end{bmatrix},\begin{bmatrix}14\\3\\2\\8\end{bmatrix},\begin{bmatrix}38\\8\\6\\24\end{bmatrix},\begin{bmatrix}47\\10\\7\\28\end{bmatrix},\begin{bmatrix}10\\2\\3\\12\end{bmatrix}\right\}$. Find the dimension of H and determine a basis.

Exercise 4.E.89

Let H denote $span\left\{\begin{bmatrix}6\\1\\1\\5\end{bmatrix},\begin{bmatrix}17\\3\\2\\10\end{bmatrix},\begin{bmatrix}52\\9\\7\\35\end{bmatrix},\begin{bmatrix}18\\3\\4\\20\end{bmatrix}\right\}$. Find the dimension of H and determine a basis.

Exercise 4.E.90

Let $M=\left\{\vec{u}=\begin{bmatrix}u_1\\u_2\\u_3\\u_4\end{bmatrix}\in\mathbb{R}^4:\sin(u_1)=1\right\}$. Is M a subspace? Explain.

Answer

No. Let $\vec{u}=\begin{bmatrix}\frac{\pi}{2}\\0\\0\\0\end{bmatrix}$. Then $2\vec{u}\not\in M$ although $\vec{u}\in M$.

Exercise 4.E.91

Let $M=\left\{\vec{u}=\begin{bmatrix}u_1\\u_2\\u_3\\u_4\end{bmatrix}\in\mathbb{R}^4:||u_1||\leq 4\right\}$. Is M a subspace? Explain.

Answer

No. $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \in M$ but $10 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \not\in M$.

Exercise 4.E.92

Let $M = \left\{ \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \in \mathbb{R}^4 : u_1 \geq 0 \text{ for each } i = 1, 2, 3, 4 \right\}$. Is M a subspace? Explain.

Answer

This is not a subspace. $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ is in it. However, $(-1) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ is not.

Exercise 4.E.93

Let $\vec{w}$, $\vec{w}_1$ be given vectors in $\mathbb{R}^4$ and define

$$M = \left\{ \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \in \mathbb{R}^4 : \vec{w} \bullet \vec{u} = 0 \text{ and } \vec{w}_1 \bullet \vec{u} = 0 \right\}.$$

Is M a subspace? Explain.

Answer

This is a subspace because it is closed with respect to vector addition and scalar multiplication.

Exercise 4.E.94

Let $\vec{w} \in \mathbb{R}^4$ and let $M = \left\{ \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \in \mathbb{R}^4 : \vec{w} \bullet \vec{u} = 0 \right\}$. Is M a subspace? Explain.

Answer

Yes, this is a subspace because it is closed with respect to vector addition and scalar multiplication.

Exercise 4.E.95

Let $M = \left\{ \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \in \mathbb{R}^4 : u_3 \geq u_1 \right\}$. Is M a subspace? Explain.

Answer

This is not a subspace. $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ is in it. However $(-1)\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}$ is not.

? Exercise 4.E. 96

Let $M = \left\{ \vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \in \mathbb{R}^4 : u_3 = u_1 = 0 \right\}$. Is M a subspace? Explain.

Answer

This is a subspace. It is closed with respect to vector addition and scalar multiplication.

? Exercise 4.E. 97

Consider the set of vectors S given by

$$S = \left\{ \begin{bmatrix} 4u+v-5w \\ 12u+6v-6w \\ 4u+4v+4w \end{bmatrix} : u, v, w \in \mathbb{R} \right\}.$$

Is S a subspace of $\mathbb{R}^3$? If so, explain why, give a basis for the subspace and find its dimension.

? Exercise 4.E. 98

Consider the set of vectors S given by

$$S = \left\{ \begin{bmatrix} 2u+6v+7w \\ -3u-9v-12w \\ 2u+6v+6w \\ u+3v+3w \end{bmatrix} : u, v, w \in \mathbb{R} \right\}.$$

Is S a subspace of $\mathbb{R}^4$? If so, explain why, give a basis for the subspace and find its dimension.

? Exercise 4.E. 99

Consider the set of vectors S given by

$$S = \left\{ \begin{bmatrix} 2u+v \\ 6v-3u+3w \\ 3v-6u+3w \end{bmatrix} : u, v, w \in \mathbb{R} \right\}.$$

Is this set of vectors a subspace of $\mathbb{R}^3$? If so, explain why, give a basis for the subspace and find its dimension.

? Exercise 4.E. 100

Consider the vectors of the form

$$\left\{ \begin{bmatrix} 2u+v+7w \\ u-2v+w \\ -6v-6w \end{bmatrix} : u, v, w \in \mathbb{R} \right\}.$$

Is this set of vectors a subspace of $\mathbb{R}^3$? If so, explain why, give a basis for the subspace and find its dimension.

Exercise 4.E.101

Consider the vectors of the form

$$\left\{ \begin{bmatrix} 3u+v+11w \\ 18u+6v+66w \\ 28u+8v+100w \end{bmatrix} : u,v,w \in \mathbb{R} \right\}.$$

Is this set of vectors a subspace of $\mathbb{R}^3$? If so, explain why, give a basis for the subspace and find its dimension.

Exercise 4.E.102

Consider the vectors of the form

$$\left\{ \begin{bmatrix} 3u+v \\ 2w-4u \\ 2w-2v-8u \end{bmatrix} : u,v,w \in \mathbb{R} \right\}.$$

Is this set of vectors a subspace of $\mathbb{R}^3$? If so, explain why, give a basis for the subspace and find its dimension.

Exercise 4.E.103

Consider the set of vectors S given by

$$\left\{ \begin{bmatrix} u+v+w \\ 2u+2v+4w \\ u+v+w \\ 0 \end{bmatrix} : u,v,w \in \mathbb{R} \right\}.$$

Is S is a subspace of $\mathbb{R}^4$? If so, explain why, give a basis for the subspace and find its dimension.

Exercise 4.E.104

Consider the set of vectors S given by

$$\left\{ \begin{bmatrix} v \\ -3u-3w \\ 8u-4v+4w \end{bmatrix} : u,v,w \in \mathbb{R} \right\}.$$

Is S is a subspace of $\mathbb{R}^4$? If so, explain why, give a basis for the subspace and find its dimension.

Exercise 4.E.105

If you have 5 vectors in $\mathbb{R}^5$ and the vectors are linearly independent, can it always be concluded they span $\mathbb{R}^5$? Explain.

Answer

Yes. If not, there would exist a vector not in the span. But then you could add in this vector and obtain a linearly independent set of vectors with more vectors than a basis.

Exercise 4.E.106

If you have 6 vectors in $\mathbb{R}^5$, is it possible they are linearly independent? Explain.

Answer

They can't be.

Exercise 4.E.107

Suppose A is an $m \times n$ matrix and $\{\vec{w}_1, \cdots, \vec{w}_k\}$ is a linearly independent set of vectors in $A(\mathbb{R}^n) \subseteq \mathbb{R}^m$. Now suppose $A\vec{z}_i = \vec{w}_i$. Show $\{\vec{z}_1, \cdots, \vec{z}_k\}$ is also independent.

Answer

Say $\sum_{i=1}^{k} c_i \vec{z}_i = \vec{0}$. Then apply A to it as follows.

$$\sum_{i=1}^{k} c_a A\vec{z}_i = \sum_{i=1}^{k} c_i \vec{w}_i = \vec{0}$$

and so, by linear independence of the $\vec{w}_i$, it follows that each $c_i = 0$.

Exercise 4.E.108

Suppose V, W are subspaces of $\mathbb{R}^n$. Let $V \cap W$ be all vectors which are in both V and W. Show that $V \cap W$ is a subspace also.

Answer

If $\vec{x}, \vec{y} \in V \cap W$, then for scalars α, β, the linear combination $\alpha\vec{x} + \beta\vec{y}$ must be in both V and W since they are both subspaces.

Exercise 4.E.109

Suppose V and W both have dimension equal to 7 and they are subspaces of $\mathbb{R}^{10}$. What are the possibilities for the dimension of $V \cap W$? Hint: Remember that a linear independent set can be extended to form a basis.

Exercise 4.E.110

Suppose V has dimension p and W has dimension q and they are each contained in a subspace, U which has dimension equal to n where $n > \max(p, q)$. What are the possibilities for the dimension of $V \cap W$? Hint: Remember that a linearly independent set can be extended to form a basis.

Answer

Let $\{x_1, \cdots, x_k\}$ be a basis for $V \cap W$. Then there is a basis for V and W which are respectively

$$\{x_1, \cdots, x_k, y_{k+1}, \cdots, y_p\},\ \{x_1, \cdots, x_k, z_{k+1}, \cdots z_q\}$$

It follows that you must have $k + p - k + q - k \leq n$ and so you must have

$$p + q - n \leq k$$

Exercise 4.E.111

Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix. Show that

$$\dim(\ker(AB)) \leq \dim(\ker(A)) + \dim(\ker(B)).$$

Consider the subspace, $B(\mathbb{R}^p) \cap \ker(A)$ and suppose a basis for this subspace is $\{\vec{w}_1, \cdots, \vec{w}_k\}$. Now suppose $\{\vec{u}_1, \cdots, \vec{u}_r\}$ is a basis for $\ker(B)$. Let $\{\vec{z}_1, \cdots, \vec{z}_k\}$ be such that $B\vec{z}_1 = \vec{w}_i$ and argue that

$$\ker(AB) \subseteq span\{\vec{u}_1, \cdots, \vec{u}_r, \vec{z}_1, \cdots, \vec{z}_k\}.$$

Answer

Here is how you do this. Suppose $AB\vec{x} = \vec{0}$. Then $B\vec{x} \in \ker(A) \cap B(\mathbb{R}^p)$ and so $B\vec{x} = \sum_{i=1}^{k} B\vec{z}_i$ showing that

$$\vec{x} - \sum_{i=1}^{k} \vec{z}_i \in \ker(B)$$

Consider $B(\mathbb{R}^p) \cap \ker(A)$ and let a basis be $\{\vec{w}_1, \cdots, \vec{w}_k\}$. Then each $\vec{w}_i$ is of the form $B\vec{z}_i = \vec{w}_i$. Therefore, $\{\vec{z}_1, \cdots, \vec{z}_k\}$ is linearly independent and $AB\vec{z}_i = 0$. Now let $\{\vec{u}_1, \cdots, \vec{u}_r\}$ be a basis for $\ker(B)$. If $AB\vec{x} = \vec{0}$, then $B\vec{x} \in \ker(A) \cap B(\mathbb{R}^p)$ and so $B\vec{x} = \sum_{i=1}^{k} c_i B\vec{z}_1$ which implies

$$\vec{x} - \sum_{i=1}^{k} c_i \vec{z}_i \in \ker(B)$$

and so it is of the form

$$\vec{x} - \sum_{i=1}^{k} c_i \vec{z}_i = \sum_{j=1}^{r} d_j \vec{u}_j$$

It follows that if $AB\vec{x} = \vec{0}$ so that $\vec{x} \in \ker(AB)$, then

$$\vec{x} \in span(\vec{z}_1, \cdots, \vec{z}_k, \vec{u}_1, \cdots, \vec{u}_r).$$

Therefore,

$$\begin{aligned} \dim(\ker(AB)) &\leq k + r = \dim(B(\mathbb{R}^p) \cap \ker(A)) + \dim(\ker(B)) \\ &\leq \dim(\ker(A)) + \dim(\ker(B)) \end{aligned}$$

? Exercise 4.E.112

Show that if A is an $m \times n$ matrix, then $\ker(A)$ is a subspace of $\mathbb{R}^n$.

Answer

If $\vec{x}, \vec{y} \in \ker(A)$ then

$$A(a\vec{x} + b\vec{y}) = aA\vec{x} + bA\vec{y} = a\vec{0} + b\vec{0} = \vec{0}$$

and so $\ker(A)$ is closed under linear combinations. Hence it is a subspace.

? Exercise 4.E.113

Find the rank of the following matrix. Also find a basis for the row and column spaces.

$$\begin{bmatrix} 1 & 3 & 0 & -2 & 0 & 3 \\ 3 & 9 & 1 & -7 & 0 & 8 \\ 1 & 3 & 1 & -3 & 1 & -1 \\ 1 & 3 & -1 & -1 & -2 & 10 \end{bmatrix}$$

Exercise 4.E.114

Find the rank of the following matrix. Also find a basis for the row and column spaces.

$$\begin{bmatrix} 1 & 3 & 0 & -2 & 7 & 3 \\ 3 & 9 & 1 & -7 & 23 & 8 \\ 1 & 3 & 1 & -3 & 9 & 2 \\ 1 & 3 & -1 & -1 & 5 & 4 \end{bmatrix}$$

Exercise 4.E.115

Find the rank of the following matrix. Also find a basis for the row and column spaces.

$$\begin{bmatrix} 1 & 0 & 3 & 0 & 7 & 0 \\ 3 & 1 & 10 & 0 & 23 & 0 \\ 1 & 1 & 4 & 1 & 7 & 0 \\ 1 & -1 & 2 & -2 & 9 & 1 \end{bmatrix}$$

Exercise 4.E.116

Find the rank of the following matrix. Also find a basis for the row and column spaces.

$$\begin{bmatrix} 1 & 0 & 3 \\ 3 & 1 & 10 \\ 1 & 1 & 4 \\ 1 & -1 & 2 \end{bmatrix}$$

Exercise 4.E.117

Find the rank of the following matrix. Also find a basis for the row and column spaces.

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 1 \\ 1 & 2 & 3 & -2 & -18 \\ 1 & 2 & 2 & -1 & -11 \\ -1 & -2 & -2 & 1 & 11 \end{bmatrix}$$

Exercise 4.E.118

Find the rank of the following matrix. Also find a basis for the row and column spaces.

$$\begin{bmatrix} 1 & 0 & 3 & 0 \\ 3 & 1 & 10 & 0 \\ -1 & 1 & -2 & 1 \\ 1 & -1 & 2 & -2 \end{bmatrix}$$

Exercise 4.E.119

Find $\ker(A)$ for the following matrices.

a. $A = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}$

b. $A = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 3 \\ 3 & 2 & 1 \end{bmatrix}$

c. $A = \begin{bmatrix} 2 & 4 & 0 \\ 3 & 6 & -2 \\ 1 & 2 & -2 \end{bmatrix}$

d. $A = \begin{bmatrix} 2 & -1 & 3 & 5 \\ 2 & 0 & 1 & 2 \\ 6 & 4 & -5 & -6 \\ 0 & 2 & -4 & -6 \end{bmatrix}$

Exercise 4.E.120

Determine whether the following set of vectors is orthogonal. If it is orthogonal, determine whether it is also orthonormal.

$$\begin{bmatrix} \frac{1}{6}\sqrt{2}\sqrt{3} \\ \frac{1}{3}\sqrt{2}\sqrt{3} \\ -\frac{1}{6}\sqrt{2}\sqrt{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{2}\sqrt{2} \\ 0 \\ \frac{1}{2}\sqrt{2} \end{bmatrix}, \begin{bmatrix} -\frac{1}{3}\sqrt{3} \\ \frac{1}{3}\sqrt{3} \\ \frac{1}{3}\sqrt{3} \end{bmatrix}$$

If the set of vectors is orthogonal but not orthonormal, give an orthonormal set of vectors which has the same span.

Exercise 4.E.121

Determine whether the following set of vectors is orthogonal. If it is orthogonal, determine whether it is also orthonormal.

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

If the set of vectors is orthogonal but not orthonormal, give an orthonormal set of vectors which has the same span.

Exercise 4.E.122

Determine whether the following set of vectors is orthogonal. If it is orthogonal, determine whether it is also orthonormal.

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

If the set of vectors is orthogonal but not orthonormal, give an orthonormal set of vectors which has the same span.

Exercise 4.E.123

Determine whether the following set of vectors is orthogonal. If it is orthogonal, determine whether it is also orthonormal.

$$\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

If the set of vectors is orthogonal but not orthonormal, give an orthonormal set of vectors which has the same span.

Exercise 4.E.124

Determine whether the following set of vectors is orthogonal. If it is orthogonal, determine whether it is also orthonormal.

$$\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\-1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

If the set of vectors is orthogonal but not orthonormal, give an orthonormal set of vectors which has the same span.

? Exercise 4.E.125

Here are some matrices. Label according to whether they are symmetric, skew symmetric, or orthogonal.

a. $\begin{bmatrix}1 & 0 & 0\\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\end{bmatrix}$

b. $\begin{bmatrix}1 & 2 & -3\\ 2 & 1 & 4\\ -3 & 4 & 7\end{bmatrix}$

c. $\begin{bmatrix}0 & -2 & -3\\ 2 & 0 & -4\\ 3 & 4 & 0\end{bmatrix}$

Answer

a. Orthogonal
b. Symmetric
c. Skew Symmetric

? Exercise 4.E.126

For U an orthogonal matrix, explain why $||U\vec{x}|| = ||\vec{x}||$ for any vector $\vec{x}$. Next explain why if U is an $n \times n$ matrix with the property that $||U\vec{x}|| = ||\vec{x}||$ for all vectors, $\vec{x}$, then U must be orthogonal. Thus the orthogonal matrices are exactly those which preserve length.

Answer

$||U\vec{x}||^2 = U\vec{x} \bullet U\vec{x} = U^T U\vec{x} \bullet \vec{x} = I\vec{x} \bullet \vec{x} = ||\vec{x}||^2$. Next suppose distance is preserved by U. Then

$$\begin{aligned}(U(\vec{x}+\vec{y})) \bullet (U(\vec{x}+\vec{y})) &= ||Ux||^2 + ||Uy||^2 + 2(Ux \bullet Uy)\\ &= ||\vec{x}||^2 + ||\vec{y}||^2 + 2(U^T U\vec{x} \bullet \vec{y})\end{aligned}$$

But since U preserves distances, it is also the case that

$$(U(\vec{x}+\vec{y}) \bullet U(\vec{x}+\vec{y})) = ||\vec{x}||^2 + ||\vec{y}||^2 + 2(\vec{x} \bullet \vec{y})$$

Hence

$$\vec{x} \bullet \vec{y} = U^T U\vec{x} \bullet \vec{y}$$

and so

$$((U^T U - I)\vec{x}) \bullet \vec{y} = 0$$

Since y is arbitrary, it follows that $U^T U - I = 0$. Thus U is orthogonal.

https://math.libretexts.org/@go/page/128002

Exercise 4.E.127

Suppose U is an orthogonal $n \times n$ matrix. Explain why $rank(U) = n$.

Answer

You could observe that $\det(UU^T) = (\det(U))^2 - 1 \quad$ so $\det(U) \neq 0$.

Exercise 4.E.128

Fill in the missing entries to make the matrix orthogonal.

$$\begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & - & - \\ - & \frac{\sqrt{6}}{3} & - \end{bmatrix}.$$

Answer

$$\begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & a \\ 0 & \frac{\sqrt{6}}{3} & b \end{bmatrix} \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & a \\ 0 & \frac{\sqrt{6}}{3} & b \end{bmatrix}^T$$
$$= \begin{bmatrix} 1 & \frac{1}{3}\sqrt{3}a - \frac{1}{3} & \frac{1}{3}\sqrt{3}b - \frac{1}{3} \\ \frac{1}{3}\sqrt{3}a - \frac{1}{3} & a^2 + \frac{2}{3} & ab - \frac{1}{3} \\ \frac{1}{3}\sqrt{3}b - \frac{1}{3} & ab - \frac{1}{3} & b^2 + \frac{2}{3} \end{bmatrix}$$

This requires, $a = 1/\sqrt{3}, b = 1/\sqrt{3}$.

$$\begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & 1/\sqrt{3} \\ 0 & \frac{\sqrt{6}}{3} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} & 1/\sqrt{3} \\ 0 & \frac{\sqrt{6}}{3} & 1/\sqrt{3} \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Exercise 4.E.129

Fill in the missing entries to make the matrix orthogonal.

$$\begin{bmatrix} \frac{2}{3} & \frac{\sqrt{2}}{2} & \frac{1}{6}\sqrt{2} \\ \frac{2}{3} & - & - \\ - & 0 & - \end{bmatrix}$$

Answer

$$\begin{bmatrix} \frac{2}{3} & \frac{\sqrt{2}}{2} & \frac{1}{6}\sqrt{2} \\ \frac{2}{3} & \frac{-\sqrt{2}}{2} & a \\ -\frac{1}{3} & 0 & b \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{\sqrt{2}}{2} & \frac{1}{6}\sqrt{2} \\ \frac{2}{3} & \frac{-\sqrt{2}}{2} & a \\ -\frac{1}{3} & 0 & b \end{bmatrix}^T = \begin{bmatrix} 1 & \frac{1}{6}\sqrt{2}a - \frac{1}{18} & \frac{1}{6}\sqrt{2}b - \frac{2}{9} \\ \frac{1}{6}\sqrt{2}a - \frac{1}{18} & a^2 + \frac{17}{18} & ab - \frac{2}{9} \\ \frac{1}{6}\sqrt{2}b - \frac{2}{9} & ab - \frac{2}{9} & b^2 + \frac{1}{9} \end{bmatrix}$$

This requires $a = \frac{1}{3\sqrt{2}}$, $b = \frac{4}{3\sqrt{2}}$.

https://math.libretexts.org/@go/page/128002

$$\begin{bmatrix} \frac{2}{3} & \frac{\sqrt{2}}{2} & \frac{1}{6}\sqrt{2} \\ \frac{2}{3} & \frac{-\sqrt{2}}{2} & \frac{1}{3\sqrt{2}} \\ -\frac{1}{3} & 0 & \frac{4}{3\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{2}{3} & \frac{\sqrt{2}}{2} & \frac{1}{6}\sqrt{2} \\ \frac{2}{3} & \frac{-\sqrt{2}}{2} & \frac{1}{3\sqrt{2}} \\ -\frac{1}{3} & 0 & \frac{4}{3\sqrt{2}} \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Exercise 4.E.130

Fill in the missing entries to make the matrix orthogonal.

$$\begin{bmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & - \\ \frac{2}{3} & 0 & - \\ - & - & \frac{4}{15}\sqrt{5} \end{bmatrix}$$

Answer

Try

$$\begin{bmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & c \\ \frac{2}{3} & 0 & d \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{15}\sqrt{5} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & c \\ \frac{2}{3} & 0 & d \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{15}\sqrt{5} \end{bmatrix}^T$$
$$= \begin{bmatrix} c^2 + \frac{41}{45} & cd + \frac{2}{9} & \frac{4}{15}\sqrt{5}c - \frac{8}{45} \\ cd + \frac{2}{9} & d^2 + \frac{4}{9} & \frac{4}{15}\sqrt{5}d + \frac{4}{9} \\ \frac{4}{15}\sqrt{5}c - \frac{8}{45} & \frac{4}{15}\sqrt{5}d + \frac{4}{9} & 1 \end{bmatrix}$$

This requires that $c = \frac{2}{3\sqrt{5}}, d = \frac{-5}{3\sqrt{5}}$.

$$\begin{bmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{-5}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{15}\sqrt{5} \end{bmatrix} \begin{bmatrix} \frac{1}{3} & -\frac{2}{\sqrt{5}} & \frac{2}{3\sqrt{5}} \\ \frac{2}{3} & 0 & \frac{-5}{3\sqrt{5}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{15}\sqrt{5} \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Exercise 4.E.131

Find an orthonormal basis for the span of each of the following sets of vectors.

a. $\begin{bmatrix} 3 \\ -4 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 7 \\ 1 \end{bmatrix}$

b. $\begin{bmatrix} 3 \\ 0 \\ -4 \end{bmatrix}, \begin{bmatrix} 11 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 7 \end{bmatrix}$

c. $\begin{bmatrix} 3 \\ 0 \\ -4 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ 10 \end{bmatrix}, \begin{bmatrix} -7 \\ 1 \\ 1 \end{bmatrix}$

Answer

a. $\begin{bmatrix} \frac{3}{5} \\ -\frac{4}{5} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

b. $\begin{bmatrix} \frac{3}{5} \\ 0 \\ -\frac{4}{5} \end{bmatrix}, \begin{bmatrix} \frac{4}{5} \\ 0 \\ \frac{3}{5} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

c. $\begin{bmatrix} \frac{3}{5} \\ 0 \\ -\frac{4}{5} \end{bmatrix}, \begin{bmatrix} \frac{4}{5} \\ 0 \\ \frac{3}{5} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

Exercise $4.E.132$

Using the Gram Schmidt process find an orthonormal basis for the following span:

$$span\left\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}$$

Answer

A solution is

$$\begin{bmatrix} \frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{6}\sqrt{6} \end{bmatrix}, \begin{bmatrix} \frac{3}{10}\sqrt{2} \\ -\frac{2}{5}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix}, \begin{bmatrix} \frac{7}{15}\sqrt{3} \\ -\frac{1}{15}\sqrt{3} \\ -\frac{1}{3}\sqrt{3} \end{bmatrix}$$

Exercise $4.E.133$

Using the Gram Schmidt process find an orthonormal basis for the following span:

$$span\left\{\begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$$

Answer

Then a solution is

$$\begin{bmatrix} \frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{6} \\ \frac{1}{6}\sqrt{6} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{6}\sqrt{2}\sqrt{3} \\ -\frac{2}{9}\sqrt{2}\sqrt{3} \\ \frac{5}{18}\sqrt{2}\sqrt{3} \\ \frac{1}{9}\sqrt{2}\sqrt{3} \end{bmatrix}, \begin{bmatrix} \frac{5}{111}\sqrt{3}\sqrt{37} \\ \frac{1}{133}\sqrt{3}\sqrt{37} \\ -\frac{17}{333}\sqrt{3}\sqrt{37} \\ \frac{22}{333}\sqrt{3}\sqrt{37} \end{bmatrix}$$

Exercise $4.E.134$

The set $V = \left\{\begin{bmatrix} x \\ y \\ z \end{bmatrix} : 2x+3y-z=0\right\}$ is a subspace of $\mathbb{R}^3$. Find an orthonormal basis for this subspace.

Answer

The subspace is of the form

https://math.libretexts.org/@go/page/128002

$$\begin{bmatrix} x \\ y \\ 2x+3y \end{bmatrix}$$

and a basis is $\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$. Therefore, an orthonormal basis is

$$\begin{bmatrix} \frac{1}{5}\sqrt{5} \\ 0 \\ \frac{2}{5}\sqrt{5} \end{bmatrix}, \begin{bmatrix} -\frac{3}{35}\sqrt{5}\sqrt{14} \\ \frac{1}{14}\sqrt{5}\sqrt{14} \\ \frac{3}{70}\sqrt{5}\sqrt{14} \end{bmatrix}$$

? Exercise 4.E. 135

Consider the following scalar equation of a plane.

$$2x-3y+z=0$$

Find the orthogonal complement of the vector $\vec{v}=\begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$. Also find the point on the plane which is closest to $(3,4,1)$.

? Exercise 4.E. 136

Consider the following scalar equation of a plane.

$$x+3y+z=0$$

Find the orthogonal complement of the vector $\vec{v}=\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$. Also find the point on the plane which is closest to $(3,4,1)$.

? Exercise 4.E. 137

Let $\vec{v}$ be a vector and let $\vec{n}$ be a normal vector for a plane through the origin. Find the equation of the line through the point determined by $\vec{v}$ which has direction vector $\vec{n}$. Show that it intersects the plane at the point determined by $\vec{v}-proj_{\vec{n}}\vec{v}$. Hint: The line:$\vec{v}+t\vec{n}$. It is in the plane if $\vec{n}\bullet(\vec{v}+t\vec{n})=0$. Determine t. Then substitute in to the equation of the line.

? Exercise 4.E. 138

As shown in the above problem, one can find the closest point to~v in a plane through the origin by finding the intersection of the line through $\vec{v}$ having direction vector equal to the normal vector to the plane with the plane. If the plane does not pass through the origin, this will still work to find the point on the plane closest to the point determined by $\vec{v}$. Here is a relation which defines a plane

$$2x+y+z=11$$

and here is a point: $(1,1,2)$. Find the point on the plane which is closest to this point. Then determine the distance from the point to the plane by taking the distance between these two points. Hint: Line: $(x,y,z)=(1,1,2)+t(2,1,1)$. Now require that it intersect the plane.

Exercise 4.E.139

In general, you have a point (x_0, y_0, z_0) and a scalar equation for a plane $ax+by+cz=d$ where $a^2+b^2+c^2>0$. Determine a formula for the closest point on the plane to the given point. Then use this point to get a formula for the distance from the given point to the plane. Hint: Find the line perpendicular to the plane which goes through the given point: $(x,y,z)=(x_0,y_0,z_0)+t(a,b,c)$. Now require that this point satisfy the equation for the plane to determine t.

Exercise 4.E.140

Find the least squares solution to the following system.

$$\begin{aligned} x+2y &= 1 \\ 2x+3y &= 2 \\ 3x+5y &= 4 \end{aligned}$$

Answer

$$\begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 5 \end{bmatrix}^T \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 14 & 23 \\ 23 & 38 \end{bmatrix} \begin{bmatrix} 14 & 23 \\ 23 & 38 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 5 \end{bmatrix}^T \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 17 \\ 28 \end{bmatrix}$$

$$\begin{bmatrix} 14 & 23 \\ 23 & 38 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 17 \\ 28 \end{bmatrix}$$

$$\begin{bmatrix} 14 & 23 \\ 23 & 38 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 17 \\ 28 \end{bmatrix},$$

Solution is: $\begin{bmatrix} \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$

Exercise 4.E.141

You are doing experiments and have obtained the ordered pairs,

$$(0,1),(1,2),(2,3.5),(3,4)$$

Find m and b such that $\vec{y}=m\vec{x}+b$ approximates these four points as well as possible.

Exercise 4.E.142

Suppose you have several ordered triples, (x_i, y_i, z_i). Describe how to find a polynomial such as

$$z=a+bx+cy+dxy+ex^2+fy^2$$

giving the best fit to the given ordered triples.

Exercise 4.E.143

The wind blows from the South at 20 kilometers per hour and an airplane which flies at 600 kilometers per hour in still air is heading East. Find the velocity of the airplane and its location after two hours.

Exercise 4.E.144

The wind blows from the West at 30 kilometers per hour and an airplane which flies at 400 kilometers per hour in still air is heading North East. Find the velocity of the airplane and its position after two hours.

Exercise 4.E.145

The wind blows from the North at 10 kilometers per hour. An airplane which flies at 300 kilometers per hour in still air is supposed to go to the point whose coordinates are at $(100, 100)$. In what direction should the airplane fly?

Exercise 4.E.146

Three forces act on an object. Two are $\begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -3 \\ 4 \end{bmatrix}$ Newtons. Find the third force if the object is not to move.

Exercise 4.E.147

Three forces act on an object. Two are $\begin{bmatrix} 6 \\ -3 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ Newtons. Find the third force if the total force on the object is to be $\begin{bmatrix} 7 \\ 1 \\ 3 \end{bmatrix}$.

Exercise 4.E.148

A river flows West at the rate of b miles per hour. A boat can move at the rate of 8 miles per hour. Find the smallest value of b such that it is not possible for the boat to proceed directly across the river.

Exercise 4.E.149

The wind blows from West to East at a speed of 50 miles per hour and an airplane which travels at 400 miles per hour in still air is heading North West. What is the velocity of the airplane relative to the ground? What is the component of this velocity in the direction North?

Answer

The velocity is the sum of two vectors. $50\vec{i} + \frac{300}{\sqrt{2}}\left(\vec{i} + \vec{j}\right) = \left(50 + \frac{300}{\sqrt{2}}\right)\vec{i} + \frac{300}{\sqrt{2}}\vec{j}$. The component in the direction of North is then $\frac{300}{\sqrt{2}} = 150\sqrt{2}$ and the velocity relative to the ground is

$$\left(50 + \frac{300}{\sqrt{2}}\right)\vec{i} + \frac{300}{\sqrt{2}}\vec{j}$$

Exercise 4.E.150

The wind blows from West to East at a speed of 60 miles per hour and an airplane can travel travels at 100 miles per hour in still air. How many degrees West of North should the airplane head in order to travel exactly North?

Exercise 4.E.151

The wind blows from West to East at a speed of 50 miles per hour and an airplane which travels at 400 miles per hour in still air heading somewhat West of North so that, with the wind, it is flying due North. It uses 30.0 gallons of gas every hour. If it has to travel 600.0 miles due North, how much gas will it use in flying to its destination?

Exercise 4.E.152

An airplane is flying due north at 150.0 miles per hour but it is not actually going due North because there is a wind which is pushing the airplane due east at 40.0 miles per hour. After one hour, the plane starts flying 30° East of North. Assuming the plane starts at $(0,0)$, where is it after 2 hours? Let North be the direction of the positive y axis and let East be the direction of the positive x axis.

Answer

Velocity of plane for the first hour: $[0 \quad 150]+[40 \quad 0]=[40 \quad 150]$. After one hour it is at $(40,150)$. Next the velocity of the plane is $150\left[\frac{1}{2} \quad \frac{\sqrt{3}}{2}\right]+[40 \quad 0]$ in miles per hour. After two hours it is then at $(40,150)+150\left[\frac{1}{2} \quad \frac{\sqrt{3}}{2}\right]+[40 \quad 0]=[155 \quad 75\sqrt{3}+150]=[155.0 \quad 279.9]$

Exercise 4.E.153

City A is located at the origin $(0,0)$ while city B is located at $(300,500)$where distances are in miles. An airplane flies at 250 miles per hour in still air. This airplane wants to fly from city A to city B but the wind is blowing in the direction of the positive y axis at a speed of 50 miles per hour. Find a unit vector such that if the plane heads in this direction, it will end up at city B having flown the shortest possible distance. How long will it take to get there?

Answer

Wind: $[0 \quad 50]$. Direction it needs to travel: $(3,5)\frac{1}{\sqrt{34}}$. Then you need $250[a \quad b]+[0 \quad 50]$ to have this direction where $[a \quad b]$ is an appropriate unit vector. Thus you need

$$\begin{aligned} a^2+b^2 &= 1 \\ \frac{250b+50}{250a} &= \frac{5}{3} \end{aligned}$$

Thus $a=\frac{3}{5}, b=\frac{4}{5}$. The velocity of the plane relative to the ground is $[150 \quad 250]$. The speed of the plane relative to the ground is given by

$$\sqrt{(150)^2+(250)^2}=291.55 \text{ miles per hour}$$

It has to go a distance of $\sqrt{(300)^2+(500)^2}=583.10$miles. Therefore, it takes

$$\frac{583.1}{291.55}=2 \text{ hours}$$

Exercise 4.E.154

A certain river is one half mile wide with a current flowing at 2 miles per hour from East to West. A man swims directly toward the opposite shore from the South bank of the river at a speed of 3 miles per hour. How far down the river does he find himself when he has swam across? How far does he end up traveling?

Answer

Water:$[-2 \quad 0]$ Swimmer:$[0 \quad 3]$ Speed relative to earth: $[-2 \quad 3]$. It takes him $1/6$ of an hour to get across. Therefore, he ends up traveling $\frac{1}{6}\sqrt{4+9}=\frac{1}{6}\sqrt{13}$ miles. He ends up $1/3$mile down stream.

Exercise 4.E.155

A certain river is one half mile wide with a current flowing at 2 miles per hour from East to West. A man can swim at 3 miles per hour in still water. In what direction should he swim in order to travel directly across the river? What would the answer to this problem be if the river flowed at 3 miles per hour and the man could swim only at the rate of 2 miles per hour?

Answer

Man: $3\left[\begin{array}{cc} a & b \end{array}\right]$ Water: $\left[\begin{array}{cc} -2 & 0 \end{array}\right]$ Then you need $3a=2$ and so $a=2/3$ and hence $b=\sqrt{5}/3$. The vector is then $\left[\begin{array}{cc} \frac{2}{3} & \frac{\sqrt{5}}{3} \end{array}\right]$.

In the second case, he could not do it. You would need to have a unit vector $\left[\begin{array}{cc} a & b \end{array}\right]$ such that $2a=3$ which is not possible.

Exercise 4.E.156

Three forces are applied to a point which does not move. Two of the forces are $2\vec{i}+2\vec{j}-6\vec{k}$ Newtons and $8\vec{i}+8\vec{j}+3\vec{k}$ Newtons. Find the third force.

Exercise 4.E.157

The total force acting on an object is to be $4\vec{i}+2\vec{j}-3\vec{k}$ Newtons. A force of $-3\vec{i}-1\vec{j}+8\vec{k}$ Newtons is being applied. What other force should be applied to achieve the desired total force?

Exercise 4.E.158

A bird flies from its nest 8 km in the direction $\frac{5}{6}\pi$ north of east where it stops to rest on a tree. It then flies 1 km in the direction due southeast and lands atop a telephone pole. Place an xy coordinate system so that the origin is the bird's nest, and the positive x axis points east and the positive y axis points north. Find the displacement vector from the nest to the telephone pole.

Exercise 4.E.159

If $\vec{F}$ is a force and $\vec{D}$ is a vector, show $\text{proj}_{\vec{D}}\left(\vec{F}\right)=\left(\|\vec{F}\|\cos\theta\right)\vec{u}$ where $\vec{u}$ is the unit vector in the direction of $\vec{D}$, where $\vec{u}=\vec{D}/\|\vec{D}\|$ and θ is the included angle between the two vectors, $\vec{F}$ and $\vec{D}$. $\|\vec{F}\|\cos\theta$ is sometimes called the component of the force, $\vec{F}$ in the direction, $\vec{D}$.

Answer

$$\text{proj}_{\vec{D}}\left(\vec{F}\right)=\frac{\vec{F}\bullet\vec{D}}{\|\vec{D}\|}\frac{\vec{D}}{\|\vec{D}\|}=\left(\|\vec{F}\|\cos\theta\right)\frac{\vec{D}}{\|\vec{D}\|}=\left(\|\vec{F}\|\cos\theta\right)\vec{u}$$

Exercise 4.E.160

A boy drags a sled for 100 feet along the ground by pulling on a rope which is 20 degrees from the horizontal with a force of 40 pounds. How much work does this force do?

Answer

$$40\cos\left(\frac{20}{180}\pi\right)100=3758.8$$

Exercise 4.E.161

A girl drags a sled for 200 feet along the ground by pulling on a rope which is 30 degrees from the horizontal with a force of 20 pounds. How much work does this force do?

Answer

$20\cos\left(\frac{\pi}{6}\right)200 = 3464.1$

Exercise 4.E.162

A large dog drags a sled for 300 feet along the ground by pulling on a rope which is 45 degrees from the horizontal with a force of 20 pounds. How much work does this force do?

Answer

$20\left(\cos\frac{\pi}{4}\right)300 = 4242.6$

Exercise 4.E.163

How much work does it take to slide a crate 20 meters along a loading dock by pulling on it with a 200 Newton force at an angle of 30° from the horizontal? Express your answer in Newton meters.

Answer

$200\left(\cos\left(\frac{\pi}{6}\right)\right)20 = 3464.1$

Exercise 4.E.164

An object moves 10 meters in the direction of $\vec{j}$. There are two forces acting on this object, $\vec{F}_1 = \vec{i}+\vec{j}+2\vec{k}$, and $\vec{F}_2 = -5\vec{i}+2\vec{j}-6\vec{k}$. Find the total work done on the object by the two forces. **Hint:** You can take the work done by the resultant of the two forces or you can add the work done by each force. Why?

Answer

$\begin{bmatrix}-4\\3\\-4\end{bmatrix}\bullet\begin{bmatrix}0\\1\\0\end{bmatrix}\times 10 = 30$ You can consider the resultant of the two forces because of the properties of the dot product.

Exercise 4.E.165

An object moves 10 meters in the direction of $\vec{j}+\vec{i}$. There are two forces acting on this object, $\vec{F}_1 = \vec{i}+2\vec{j}+2\vec{k}$, and $\vec{F}_2 = 5\vec{i}+2\vec{j}-6\vec{k}$. Find the total work done on the object by the two forces. **Hint:** You can take the work done by the resultant of the two forces or you can add the work done by each force. Why?

Answer

$$\vec{F}_1 \bullet \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} 10 + \vec{F}_2 \bullet \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} 10 = \left(\vec{F}_1 + \vec{F}_2\right) \bullet \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} 10$$
$$= \begin{bmatrix} 6 \\ 4 \\ -4 \end{bmatrix} \bullet \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} 10$$
$$= 50\sqrt{2}$$

? Exercise 4.E.166

An object moves 20 meters in the direction of $\vec{k}+\vec{j}$. There are two forces acting on this object, $\vec{F}_1 = \vec{i}+\vec{j}+2\vec{k}$, and $\vec{F}_2 = \vec{i}+2\vec{j}-6\vec{k}$. Find the total work done on the object by the two forces. **Hint:** You can take the work done by the resultant of the two forces or you can add the work done by each force.

Answer

$$\begin{bmatrix} 2 \\ 3 \\ -4 \end{bmatrix} \bullet \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} 20 = -10\sqrt{2}$$

This page titled 4.E: Exercises is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **4.12: Applications** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

CHAPTER OVERVIEW

5: Linear Transformations

5.1: Linear Transformations
5.2: The Matrix of a Linear Transformation I
5.3: Properties of Linear Transformations
5.4: Special Linear Transformations in R^2
5.5: One-to-One and Onto Transformations
5.6: Isomorphisms
5.7: The Kernel and Image of A Linear Map
5.8: The Matrix of a Linear Transformation II
5.9: The General Solution of a Linear System
5.E: Exercises

This page titled 5: Linear Transformations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx)
.

5.1: Linear Transformations

Outcomes

A. Understand the definition of a linear transformation, and that all linear transformations are determined by matrix multiplication.

Recall that when we multiply an $m \times n$ matrix by an $n \times 1$ column vector, the result is an $m \times 1$ column vector. In this section we will discuss how, through matrix multiplication, an $m \times n$ matrix **transforms** an $n \times 1$ column vector into an $m \times 1$ column vector.

Recall that the $n \times 1$ vector given by

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

is said to belong to $\mathbb{R}^n$, which is the set of all $n \times 1$ vectors. In this section, we will discuss transformations of vectors in $\mathbb{R}^n$.

Consider the following example.

Example 5.1.1:A Function Which Transforms Vectors

Consider the matrix $A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{bmatrix}$. Show that by matrix multiplication A transforms vectors in $\mathbb{R}^3$ into vectors in $\mathbb{R}^2$.

Solution

First, recall that vectors in $\mathbb{R}^3$ are vectors of size 3×1, while vectors in $\mathbb{R}^2$ are of size 2×1. If we multiply A, which is a 2×3 matrix, by a 3×1 vector, the result will be a 2×1 vector. This what we mean when we say that A *transforms* vectors.

Now, for $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in $\mathbb{R}^3$, multiply on the left by the given matrix to obtain the new vector. This product looks like

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y \\ 2x+y \end{bmatrix}$$

The resulting product is a 2×1 vector which is determined by the choice of x and y. Here are some numerical examples.

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

Here, the vector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ in $\mathbb{R}^3$ was transformed by the matrix into the vector $\begin{bmatrix} 5 \\ 4 \end{bmatrix}$ in $\mathbb{R}^2$.

Here is another example:

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 10 \\ 5 \\ -3 \end{bmatrix} = \begin{bmatrix} 20 \\ 25 \end{bmatrix}$$

The idea is to define a function which takes vectors in $\mathbb{R}^3$ and delivers new vectors in $\mathbb{R}^2$. In this case, that function is multiplication by the matrix A.

Let T denote such a function. The notation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ means that the function T transforms vectors in $\mathbb{R}^n$ into vectors in $\mathbb{R}^m$. The notation $T(\vec{x})$ means the transformation T applied to the vector $\vec{x}$. The above example demonstrated a transformation achieved by matrix multiplication. In this case, we often write

$$T_A(\vec{x}) = A\vec{x}$$

Therefore, T_A is the transformation determined by the matrix A. In this case we say that T is a matrix transformation.

Below is a video on matrix transformations.



Recall the property of matrix multiplication that states that for k and p scalars,

$$A(kB+pC) = kAB+pAC$$

In particular, for A an $m \times n$ matrix and B and C, $n \times 1$ vectors in $\mathbb{R}^n$, this formula holds.

In other words, this means that matrix multiplication gives an example of a linear transformation, which we will now define.

Definition 5.1.1:Linear Transformation

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a function, where for each $\vec{x} \in \mathbb{R}^n, T(\vec{x}) \in \mathbb{R}^m$. Then T is a **linear transformation** if whenever k, p are scalars and $\vec{x}_1$ and $\vec{x}_2$ are vectors in $\mathbb{R}^n$ ($n \times 1$ vectors),

$$T(k\vec{x}_1 + p\vec{x}_2) = kT(\vec{x}_1) + pT(\vec{x}_2)$$

Consider the following example.

Example 5.1.2:Linear Transformation

Let T be a transformation defined by $T : \mathbb{R}^3 \to \mathbb{R}^2$ is defined by

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ x-z \end{bmatrix} \text{ for all } \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3$$

Show that T is a linear transformation.

Solution

By Definition 5.1.1 we need to show that $T(k\vec{x}_1 + p\vec{x}_2) = kT(\vec{x}_1) + pT(\vec{x}_2)$ for all scalars k, p and vectors $\vec{x}_1, \vec{x}_2$. Let

$$\vec{x}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$$

Then

$$
\begin{aligned}
T\left(k\vec{x}_1+p\vec{x}_2\right) &= T\left(k\begin{bmatrix}x_1\\y_1\\z_1\end{bmatrix}+p\begin{bmatrix}x_2\\y_2\\z_2\end{bmatrix}\right)\\
&= T\left(\begin{bmatrix}kx_1\\ky_1\\kz_1\end{bmatrix}+\begin{bmatrix}px_2\\py_2\\pz_2\end{bmatrix}\right)\\
&= T\left(\begin{bmatrix}kx_1+px_2\\ky_1+py_2\\kz_1+pz_2\end{bmatrix}\right)\\
&= \begin{bmatrix}(kx_1+px_2)+(ky_1+py_2)\\(kx_1+px_2)-(kz_1+pz_2)\end{bmatrix}\\
&= \begin{bmatrix}(kx_1+ky_1)+(px_2+py_2)\\(kx_1-kz_1)+(px_2-pz_2)\end{bmatrix}\\
&= \begin{bmatrix}kx_1+ky_1\\kx_1-kz_1\end{bmatrix}+\begin{bmatrix}px_2+py_2\\px_2-pz_2\end{bmatrix}\\
&= k\begin{bmatrix}x_1+y_1\\x_1-z_1\end{bmatrix}+p\begin{bmatrix}x_2+y_2\\x_2-z_2\end{bmatrix}\\
&= kT(\vec{x}_1)+pT(\vec{x}_2)
\end{aligned}
$$

Therefore T is a linear transformation.

Two important examples of linear transformations are the zero transformation and identity transformation. The zero transformation defined by $T\left(\vec{x}\right)=\vec{0}$ for all $\vec{x}$ is an example of a linear transformation. Similarly the identity transformation defined by $T\left(\vec{x}\right)=\vec{x}$ is also linear. Take the time to prove these using the method demonstrated in Example 5.1.2.

We began this section by discussing matrix transformations, where multiplication by a matrix transforms vectors. These matrix transformations are in fact linear transformations.

Theorem 5.1.1:Matrix Transformations are Linear Transformations

Let $T:\mathbb{R}^n\mapsto\mathbb{R}^m$ be a transformation defined by $T(\vec{x})=A\vec{x}$. Then T is a linear transformation.

It turns out that every linear transformation can be expressed as a matrix transformation, and thus linear transformations are exactly the same as matrix transformations.

Below is a video on finding the domain and codomain of a linear transformation given the transformation matrix.



This page titled 5.1: Linear Transformations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **5.1: Linear Transformations** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

5.2: The Matrix of a Linear Transformation I

Outcomes

A. Find the matrix of a linear transformation with respect to the standard basis.
B. Determine the action of a linear transformation on a vector in $\mathbb{R}^n$.

In the above examples, the action of the linear transformations was to multiply by a matrix. It turns out that this is always the case for linear transformations. If T is **any** linear transformation which maps $\mathbb{R}^n$ to $\mathbb{R}^m$, there is **always** an $m \times n$ matrix A with the property that

$$T\left(\vec{x}\right) = A\vec{x} \tag{5.2.1}$$

for all $\vec{x} \in \mathbb{R}^n$.

Theorem 5.2.1:Matrix of a Linear Transformation

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. Then we can find a matrix A such that $T(\vec{x}) = A\vec{x}$. In this case, we say that T is *determined* or *induced* by the matrix A.

Here is why. Suppose $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is a linear transformation and you want to find the matrix defined by this linear transformation as described in (5.2.1). Note that

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + x_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \sum_{i=1}^{n} x_i \vec{e}_i$$

where $\vec{e}_i$ is the ith column of I_n, that is the $n \times 1$ vector which has zeros in every slot but the ith and a 1 in this slot.

Then since T is linear,

$$\begin{aligned} T\left(\vec{x}\right) &= \sum_{i=1}^{n} x_i T\left(\vec{e}_i\right) \\ &= \begin{bmatrix} | & & | \\ T\left(\vec{e}_1\right) & \cdots & T\left(\vec{e}_n\right) \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \end{aligned}$$

The desired matrix is obtained from constructing the ith column as $T\left(\vec{e}_i\right)$. Recall that the set $\{\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n\}$ is called the standard basis of $\mathbb{R}^n$. Therefore the matrix of T is found by applying T to the standard basis. We state this formally as the following theorem.

Theorem 5.2.2:Matrix of a Linear Transformation

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. Then the matrix A satisfying $T\left(\vec{x}\right) = A\vec{x}$ is given by

$$A = \begin{bmatrix} | & & | \\ T\left(\vec{e}_1\right) & \cdots & T\left(\vec{e}_n\right) \\ | & & | \end{bmatrix}$$

where $\vec{e}_i$ is the ith column of I_n, and then $T\left(\vec{e}_i\right)$ is the ith column of A.

The following Corollary is an essential result.

Corollary 5.2.1:Matrix and Linear Transformation

A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if it is a matrix transformation.

Consider the following example.

Example 5.2.1:The Matrix of a Linear Transformation

Suppose T is a linear transformation, $T : \mathbb{R}^3 \to \mathbb{R}^2$ where

$$T\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix}, T\begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}9\\-3\end{bmatrix}, T\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix}$$

Find the matrix A of T such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x}$.

Solution

By Theorem 5.2.2 we construct A as follows:

$$A = \begin{bmatrix} | & & | \\ T(\vec{e}_1) & \cdots & T(\vec{e}_n) \\ | & & | \end{bmatrix}$$

In this case, A will be a 2×3 matrix, so we need to find $T(\vec{e}_1), T(\vec{e}_2)$, and $T(\vec{e}_3)$. Luckily, we have been given these values so we can fill in A as needed, using these vectors as the columns of A. Hence,

$$A = \begin{bmatrix}1 & 9 & 1\\2 & -3 & 1\end{bmatrix}$$

In this example, we were given the resulting vectors of $T(\vec{e}_1), T(\vec{e}_2)$, and $T(\vec{e}_3)$. Constructing the matrix A was simple, as we could simply use these vectors as the columns of A. The next example shows how to find A when we are not given the $T(\vec{e}_i)$ so clearly.

Example 5.2.2: The Matrix of Linear Transformation: Inconveniently Defined

Suppose T is a linear transformation, $T : \mathbb{R}^2 \to \mathbb{R}^2$ and

$$T\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1\\2\end{bmatrix}, T\begin{bmatrix}0\\-1\end{bmatrix} = \begin{bmatrix}3\\2\end{bmatrix}$$

Find the matrix A of T such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x}$.

Solution

By Theorem 5.2.2 to find this matrix, we need to determine the action of T on $\vec{e}_1$ and $\vec{e}_2$. In Example 9.9.2, we were given these resulting vectors. However, in this example, we have been given T of two different vectors. How can we find out the action of T on $\vec{e}_1$ and $\vec{e}_2$? In particular for $\vec{e}_1$, suppose there exist x and y such that

$$\begin{bmatrix}1\\0\end{bmatrix} = x\begin{bmatrix}1\\1\end{bmatrix} + y\begin{bmatrix}0\\-1\end{bmatrix} \tag{5.2.2}$$

Then, since T is linear,

$$T\begin{bmatrix}1\\0\end{bmatrix} = xT\begin{bmatrix}1\\1\end{bmatrix} + yT\begin{bmatrix}0\\-1\end{bmatrix}$$

Substituting in values, this sum becomes

$$T\begin{bmatrix}1\\0\end{bmatrix}=x\begin{bmatrix}1\\2\end{bmatrix}+y\begin{bmatrix}3\\2\end{bmatrix} \tag{5.2.3}$$

Therefore, if we know the values of x and y which satisfy (5.2.2), we can substitute these into equation (5.2.3). By doing so, we find $T\left(\vec{e}_1\right)$ which is the first column of the matrix A.

We proceed to find x and y. We do so by solving (5.2.2), which can be done by solving the system

$$\begin{aligned}x&=1\\x-y&=0\end{aligned}$$

We see that $x=1$ and $y=1$ is the solution to this system. Substituting these values into equation (5.2.3), we have

$$T\begin{bmatrix}1\\0\end{bmatrix}=1\begin{bmatrix}1\\2\end{bmatrix}+1\begin{bmatrix}3\\2\end{bmatrix}=\begin{bmatrix}1\\2\end{bmatrix}+\begin{bmatrix}3\\2\end{bmatrix}=\begin{bmatrix}4\\4\end{bmatrix}$$

Therefore $\begin{bmatrix}4\\4\end{bmatrix}$ is the first column of A.

Computing the second column is done in the same way, and is left as an exercise.

The resulting matrix A is given by

$$A=\begin{bmatrix}4&-3\\4&-2\end{bmatrix}$$

This example illustrates a very long procedure for finding the matrix of A.

Below is a video on finding a transformation matrix given T(i) and T(j).



Below is a video on finding a transformation matrix given T(i), T(j) and T(k).



Below is a video on finding a transformation matrix given T(x) = b.



While this method is reliable and will always result in the correct matrix A, the following procedure provides an alternative method.

Procedure 5.2.1: Finding the Matrix of Inconveniently Defined Linear Transformation

Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Suppose there exist vectors $\{\vec{a}_1, \cdots, \vec{a}_n\}$ in $\mathbb{R}^n$ such that $\begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}^{-1}$ exists, and

$$T(\vec{a}_i) = \vec{b}_i$$

Then the matrix of T must be of the form

$$\begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_n \end{bmatrix} \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}^{-1}$$

We will illustrate this procedure in the following example. You may also find it useful to work through Example 5.2.2 using this procedure.

Example 5.2.3: Matrix of a Linear Transformation Given Inconveniently

Suppose $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation and

$$T\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, T\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, T\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Find the matrix of this linear transformation.

Solution

By Procedure 5.2.1, $A = \begin{bmatrix} 1 & 0 & 1 \\ 3 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}^{-1}$ and $B = \begin{bmatrix} 0 & 2 & 0 \\ 1 & 1 & 0 \\ 1 & 3 & 1 \end{bmatrix}$

Then, Procedure 5.2.1 claims that the matrix of T is

$$C = BA^{-1} = \begin{bmatrix} 2 & -2 & 4 \\ 0 & 0 & 1 \\ 4 & -3 & 6 \end{bmatrix}$$

Indeed you can first verify that $T(\vec{x}) = C\vec{x}$ for the 3 vectors above:

$$\begin{bmatrix} 2 & -2 & 4 \\ 0 & 0 & 1 \\ 4 & -3 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 & -2 & 4 \\ 0 & 0 & 1 \\ 4 & -3 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -2 & 4 \\ 0 & 0 & 1 \\ 4 & -3 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

But more generally $T(\vec{x}) = C\vec{x}$ for any $\vec{x}$. To see this, let $\vec{y} = A^{-1}\vec{x}$ and then using linearity of T:

$$T(\vec{x}) = T(A\vec{y}) = T\left(\sum_i \vec{y}_i \vec{a}_i\right) = \sum \vec{y}_i T(\vec{a}_i) \sum \vec{y}_i \vec{b}_i = B\vec{y} = BA^{-1}\vec{x} = C\vec{x}$$

Recall the dot product discussed earlier. Consider the map $\vec{v} \mapsto \text{proj}_{\vec{u}}(\vec{v})$ which takes a vector a transforms it to its projection onto a given vector $\vec{u}$. It turns out that this map is linear, a result which follows from the properties of the dot product. This is shown as follows.

$$\begin{aligned} \text{proj}_{\vec{u}}(k\vec{v} + p\vec{w}) &= \left(\frac{(k\vec{v} + p\vec{w}) \bullet \vec{u}}{\vec{u} \bullet \vec{u}}\right) \vec{u} \\ &= k\left(\frac{\vec{v} \bullet \vec{u}}{\vec{u} \bullet \vec{u}}\right) \vec{u} + p\left(0.05in\frac{\vec{w} \bullet \vec{u}}{\vec{u} \bullet \vec{u}}\right) \vec{u} \\ &= k \text{ proj}_{\vec{u}}(\vec{v}) + p \text{ proj}_{\vec{u}}(\vec{w}) \end{aligned}$$

Below is a video on finding a transformation matrix given where it sends three specified vectors.



Consider the following example.

Example 5.2.4: Matrix of a Projection Map

Let $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and let T be the projection map $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$ defined by

$$T(\vec{v}) = \text{proj}_{\vec{u}}(\vec{v})$$

for any $\vec{v} \in \mathbb{R}^3$.

1. Does this transformation come from multiplication by a matrix?
2. If so, what is the matrix?

Solution

1. First, we have just seen that $T(\vec{v}) = \text{proj}_{\vec{u}}(\vec{v})$ is linear. Therefore by Theorem 5.2.1, we can find a matrix A such that $T(\vec{x}) = A\vec{x}$.

2. The columns of the matrix for T are defined above as $T(\vec{e}_i)$. It follows that $T(\vec{e}_i) = \text{proj}_{\vec{u}}(\vec{e}_i)$ gives the ith column of the desired matrix. Therefore, we need to find

$$\text{proj}_{\vec{u}}(\vec{e}_i) = \left(\frac{\vec{e}_i \bullet \vec{u}}{\vec{u} \bullet \vec{u}}\right)\vec{u}$$

For the given vector $\vec{u}$, this implies the columns of the desired matrix are

$$\frac{1}{14}\begin{bmatrix}1\\2\\3\end{bmatrix}, \frac{2}{14}\begin{bmatrix}1\\2\\3\end{bmatrix}, \frac{3}{14}\begin{bmatrix}1\\2\\3\end{bmatrix}$$

which you can verify. Hence the matrix of T is

$$\frac{1}{14}\begin{bmatrix}1 & 2 & 3\\2 & 4 & 6\\3 & 6 & 9\end{bmatrix}$$

Below is a video on matching transformation matrices to their transformation descriptions.



This page titled 5.2: The Matrix of a Linear Transformation I is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **5.2: The Matrix of a Linear Transformation I** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

5.3: Properties of Linear Transformations

Outcomes

A. Use properties of linear transformations to solve problems.
B. Find the composite of transformations and the inverse of a transformation.

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. Then there are some important properties of T which will be examined in this section. Consider the following theorem.

Theorem 5.3.1: Properties of Linear Transformations

Properties of Linear Transformationsproperties Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation and let $\vec{x} \in \mathbb{R}^n$.

- T preserves the zero vector.

$$T(0\vec{x}) = 0T(\vec{x}). \text{ Hence } T(\vec{0}) = \vec{0}$$

- T preserves the negative of a vector:

$$T((-1)\vec{x}) = (-1)T(\vec{x}). \text{ Hence } T(-\vec{x}) = -T(\vec{x}).$$

- T preserves linear combinations:

$$\text{Let } \vec{x}_1, \ldots, \vec{x}_k \in \mathbb{R}^n \text{ and } a_1, \ldots, a_k \in \mathbb{R}.$$

$$\text{Then if } \vec{y} = a_1\vec{x}_1 + a_2\vec{x}_2 + \ldots + a_k\vec{x}_k, \text{ it follows that}$$

$$T(\vec{y}) = T(a_1\vec{x}_1 + a_2\vec{x}_2 + \ldots + a_k\vec{x}_k) = a_1T(\vec{x}_1) + a_2T(\vec{x}_2) + \ldots + a_kT(\vec{x}_k).$$

These properties are useful in determining the action of a transformation on a given vector. Consider the following example.

Example 5.3.1:Linear Combination

Let $T : \mathbb{R}^3 \mapsto \mathbb{R}^4$ be a linear transformation such that

$$T\begin{bmatrix}1\\3\\1\end{bmatrix} = \begin{bmatrix}4\\4\\0\\-2\end{bmatrix}, T\begin{bmatrix}4\\0\\5\end{bmatrix} = \begin{bmatrix}4\\5\\-1\\5\end{bmatrix}$$

Find $T\begin{bmatrix}-7\\3\\-9\end{bmatrix}$.

Solution

Using the third property in Theorem 9.6.1, we can find $T\begin{bmatrix}-7\\3\\-9\end{bmatrix}$ by writing $\begin{bmatrix}-7\\3\\-9\end{bmatrix}$ as a linear combination of $\begin{bmatrix}1\\3\\1\end{bmatrix}$ and $\begin{bmatrix}4\\0\\5\end{bmatrix}$.

Therefore we want to find $a, b \in \mathbb{R}$ such that

$$\begin{bmatrix}-7\\3\\-9\end{bmatrix} = a\begin{bmatrix}1\\3\\1\end{bmatrix} + b\begin{bmatrix}4\\0\\5\end{bmatrix}$$

The necessary augmented matrix and resulting reduced row-echelon form are given by:

$$\left[\begin{array}{cc|c} 1 & 4 & -7 \\ 3 & 0 & 3 \\ 1 & 5 & -9 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{array}\right]$$

Hence $a = 1, b = -2$ and

$$\begin{bmatrix} -7 \\ 3 \\ -9 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + (-2) \begin{bmatrix} 4 \\ 0 \\ 5 \end{bmatrix}$$

Now, using the third property above, we have

$$\begin{aligned} T \begin{bmatrix} -7 \\ 3 \\ -9 \end{bmatrix} &= T \left(1 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + (-2) \begin{bmatrix} 4 \\ 0 \\ 5 \end{bmatrix} \right) \\ &= 1T \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} - 2T \begin{bmatrix} 4 \\ 0 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} 4 \\ 4 \\ 0 \\ -2 \end{bmatrix} - 2 \begin{bmatrix} 4 \\ 5 \\ -1 \\ 5 \end{bmatrix} \\ &= \begin{bmatrix} -4 \\ -6 \\ 2 \\ -12 \end{bmatrix} \end{aligned}$$

Therefore, $T \begin{bmatrix} -7 \\ 3 \\ -9 \end{bmatrix} = \begin{bmatrix} -4 \\ -6 \\ 2 \\ -12 \end{bmatrix}$.

Below is a video on finding linear combinations of a given linear transformation using properties.



Suppose two linear transformations act in the same way on $\vec{x}$ for all vectors. Then we say that these transformations are equal.

Definition 5.3.1: Equal Transformations

Let S and T be linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. Then $S = T$ if and only if for every $\vec{x} \in \mathbb{R}^n$,

$$S(\vec{x}) = T(\vec{x})$$

Suppose two linear transformations act on the same vector $\vec{x}$, first the transformation T and then a second transformation given by S. We can find the **composite** transformation that results from applying both transformations.

Definition 5.3.2:Composition of Linear Transformations

Let $T : \mathbb{R}^k \mapsto \mathbb{R}^n$ and $S : \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear transformations. Then the **composite** of S and T is

$$S \circ T : \mathbb{R}^k \mapsto \mathbb{R}^m$$

The action of $S \circ T$ is given by

$$(S \circ T)(\vec{x}) = S(T(\vec{x})) \text{ for all } \vec{x} \in \mathbb{R}^k$$

Notice that the resulting vector will be in $\mathbb{R}^m$. Be careful to observe the order of transformations. We write $S \circ T$ but apply the transformation T first, followed by S.

Theorem 5.3.2:Composition of Transformations

Let $T : \mathbb{R}^k \mapsto \mathbb{R}^n$ and $S : \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear transformations such that T is induced by the matrix A and S is induced by the matrix B. Then $S \circ T$ is a linear transformation which is induced by the matrix BA.

Consider the following example.

Example 5.3.2:Composition of Transformations

Let T be a linear transformation induced by the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}$$

and S a linear transformation induced by the matrix

$$B = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$$

Find the matrix of the composite transformation $S \circ T$. Then, find $(S \circ T)(\vec{x})$ for $\vec{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$.

Solution

By Theorem 5.3.2, the matrix of $S \circ T$ is given by BA.

$$BA = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 8 & 4 \\ 2 & 0 \end{bmatrix}$$

To find $(S \circ T)(\vec{x})$, multiply $\vec{x}$ by BA as follows

$$\begin{bmatrix} 8 & 4 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 24 \\ 2 \end{bmatrix}$$

To check, first determine $T(\vec{x})$:

$$\begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ 2 \end{bmatrix}$$

Then, compute $S(T(\vec{x}))$ as follows:

$$\begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 9 \\ 2 \end{bmatrix} = \begin{bmatrix} 24 \\ 2 \end{bmatrix}$$

Below is a video on finding standard matrices for transformations and a composition of transformations.



Consider a composite transformation $S \circ T$, and suppose that this transformation acted such that $(S \circ T)(\vec{x}) = \vec{x}$. That is, the transformation S took the vector $T(\vec{x})$ and returned it to $\vec{x}$. In this case, S and T are inverses of each other. Consider the following definition.

Definition 5.3.3:Inverse of a Transformation

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ and $S : \mathbb{R}^n \mapsto \mathbb{R}^n$ be linear transformations. Suppose that for each $\vec{x} \in \mathbb{R}^n$,

$$(S \circ T)(\vec{x}) = \vec{x}$$

and

$$(T \circ S)(\vec{x}) = \vec{x}$$

Then, S is called an inverse of T and T is called an inverse of S. Geometrically, they reverse the action of each other.

The following theorem is crucial, as it claims that the above inverse transformations are unique.

Theorem 5.3.3:Inverse of a Transformation

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ be a linear transformation induced by the matrix A. Then T has an inverse transformation if and only if the matrix A is invertible. In this case, the inverse transformation is unique and denoted $T^{-1} : \mathbb{R}^n \mapsto \mathbb{R}^n$. T^{-1} is induced by the matrix A^{-1}.

Consider the following example.

Example 5.3.3:Inverse of a Transformation

Let $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be a linear transformation induced by the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$$

Show that T^{-1} exists and find the matrix B which it is induced by.

Solution

Since the matrix A is invertible, it follows that the transformation T is invertible. Therefore, T^{-1} exists.

You can verify that A^{-1} is given by:

$$A^{-1} = \begin{bmatrix} -4 & 3 \\ 3 & -2 \end{bmatrix}$$

Therefore the linear transformation T^{-1} is induced by the matrix A^{-1}.

Below is a video on finding the transformation matrix in R^2 given two transformations using an inverse matrix.



Below is a video on finding the transformation matrix in R^3 given three transformations using an inverse matrix.



This page titled 5.3: Properties of Linear Transformations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **5.3: Properties of Linear Transformations** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

5.4: Special Linear Transformations in R²

Outcomes

A. Find the matrix of rotations and reflections in $\mathbb{R}^2$ and determine the action of each on a vector in $\mathbb{R}^2$.

In this section, we will examine some special examples of linear transformations in $\mathbb{R}^2$ including rotations and reflections. We will use the geometric descriptions of vector addition and scalar multiplication discussed earlier to show that a rotation of vectors through an angle and reflection of a vector across a line are examples of linear transformations.

More generally, denote a transformation given by a rotation by T. Why is such a transformation linear? Consider the following picture which illustrates a rotation. Let $\vec{u}, \vec{v}$ denote vectors.

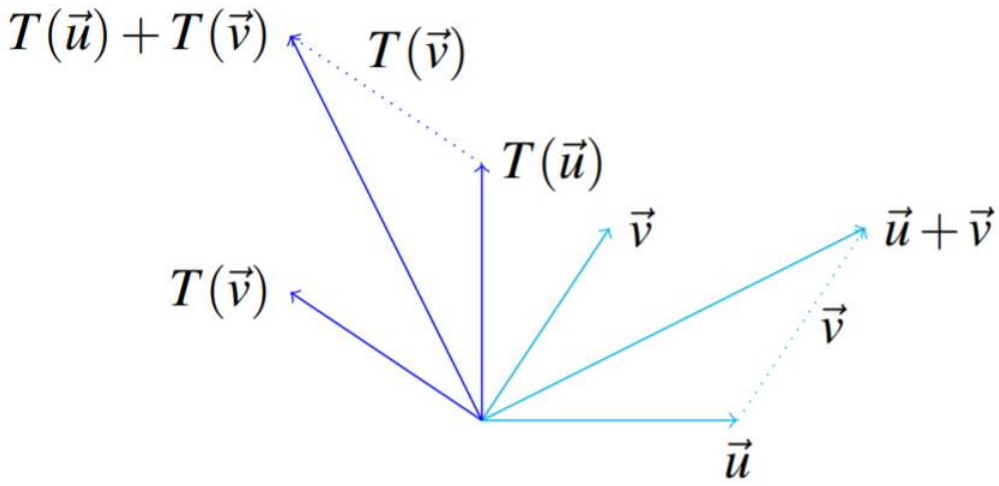


Figure 5.4.1

Let's consider how to obtain $T[\vec{u}+\vec{v}]$. Simply, you add $T(\vec{u})$ and $T(\vec{v})$. Here is why. If you add $T(\vec{u})$ to $T(\vec{v})$ you get the diagonal of the parallelogram determined by $T(\vec{u})$ and $T(\vec{v})$, as this action is our usual vector addition. Now, suppose we first add $\vec{u}$ and $\vec{v}$, and then apply the transformation T to $\vec{u}+\vec{v}$. Hence, we find $T(\vec{u}+\vec{v})$. As shown in the diagram, this will result in the same vector. In other words, $T(\vec{u}+\vec{v})=T(\vec{u})+T(\vec{v})$.

This is because the rotation preserves all angles between the vectors as well as their lengths. In particular, it preserves the shape of this parallelogram. Thus both $T[\vec{u}]+T[\vec{v}]$ and $T[\vec{u}+\vec{v}]$ give the same vector. It follows that T distributes across addition of the vectors of $\mathbb{R}^2$.

Similarly, if k is a scalar, it follows that $T[k\vec{u}]=kT[\vec{u}]$. Thus rotations are an example of a linear transformation by Definition 9.6.1.

The following theorem gives the matrix of a linear transformation which rotates all vectors through an angle of θ.

Theorem 5.4.1:Rotation

Let $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation given by rotating vectors through an angle of θ. Then the matrix A of R_θ is given by

$$\begin{bmatrix} \cos[\theta] & -\sin[\theta] \\ \sin[\theta] & \cos[\theta] \end{bmatrix}$$

Proof

Let $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. These identify the geometric vectors which point along the positive x axis and positive y axis as shown.

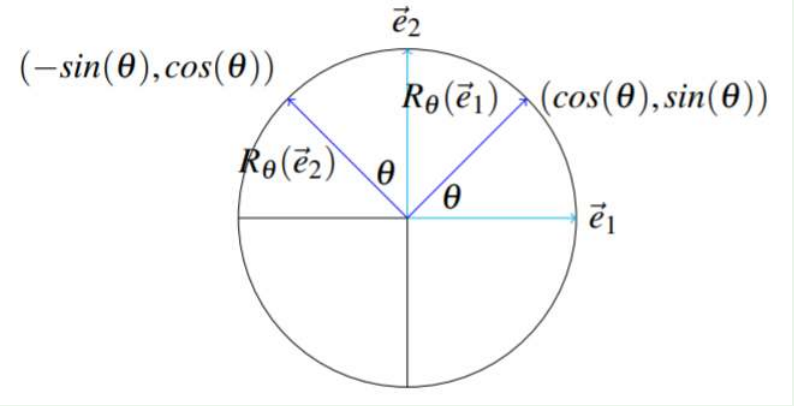


Figure 5.4.2

From Theorem 5.2.2, we need to find $R_\theta(\vec{e}_1)$ and $R_\theta(\vec{e}_2)$, and use these as the columns of the matrix A of T. We can use cos, sin of the angle θ to find the coordinates of $R_\theta(\vec{e}_1)$ as shown in the above picture. The coordinates of $R_\theta(\vec{e}_2)$ also follow from trigonometry. Thus

$$R_\theta(\vec{e}_1) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, R_\theta(\vec{e}_2) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}$$

Therefore, from Theorem 5.2.2,

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

We can also prove this algebraically without the use of the above picture. The definition of $[\cos[\theta], \sin[\theta]]$ is as the coordinates of the point of $R_\theta(\vec{e}_1)$. Now the point of the vector $\vec{e}_2$ is exactly $\pi/2$ further along the unit circle from the point of $\vec{e}_1$, and therefore after rotation through an angle of θ the coordinates x and y of the point of $R_\theta(\vec{e}_2)$ are given by

$$[x, y] = [\cos[\theta + \pi/2], \sin[\theta + \pi/2]] = [-\sin\theta, \cos\theta]$$

Consider the following example.

Example 5.4.1:Rotation in $\mathbb{R}^2$

Let $R_{\frac{\pi}{2}} : \mathbb{R}^2 \to \mathbb{R}^2$ denote rotation through $\pi/2$. Find the matrix of $R_{\frac{\pi}{2}}$. Then, find $R_{\frac{\pi}{2}}(\vec{x})$ where $\vec{x} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$.

Solution

By Theorem 5.4.1, the matrix of $R_{\frac{\pi}{2}}$ is given by

$$\begin{bmatrix} \cos[\theta] & -\sin[\theta] \\ \sin[\theta] & \cos[\theta] \end{bmatrix} = \begin{bmatrix} \cos[\pi/2] & -\sin[\pi/2] \\ \sin[\pi/2] & \cos[\pi/2] \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

To find $R_{\frac{\pi}{2}}(\vec{x})$, we multiply the matrix of $R_{\frac{\pi}{2}}$ by $\vec{x}$ as follows

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Below is a video on finding the transformation matrix ifor a 2D rotation.



We now look at an example of a linear transformation involving two angles.

Example 5.4.2: The Rotation Matrix of the Sum of Two Angles

Find the matrix of the linear transformation which is obtained by first rotating all vectors through an angle of ϕ and then through an angle θ. Hence the linear transformation rotates all vectors through an angle of $\theta+\phi$.

Solution

Let $R_{\theta+\phi}$ denote the linear transformation which rotates every vector through an angle of $\theta+\phi$. Then to obtain $R_{\theta+\phi}$, we first apply R_ϕ and then R_θ where R_ϕ is the linear transformation which rotates through an angle of ϕ and R_θ is the linear transformation which rotates through an angle of θ. Denoting the corresponding matrices by $A_{\theta+\phi}$, A_ϕ, and A_θ, it follows that for every $\vec{u}$

$$R_{\theta+\phi}\left[\vec{u}\right] = A_{\theta+\phi}\vec{u} = A_\theta A_\phi \vec{u} = R_\theta R_\phi \left[\vec{u}\right]$$

Notice the order of the matrices here!

Consequently, you must have

$$\begin{aligned} A_{\theta+\phi} &= \begin{bmatrix} \cos[\theta+\phi] & -\sin[\theta+\phi] \\ \sin[\theta+\phi] & \cos[\theta+\phi] \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} = A_\theta A_\phi \end{aligned}$$

The usual matrix multiplication yields

$$\begin{aligned} A_{\theta+\phi} &= \begin{bmatrix} \cos[\theta+\phi] & -\sin[\theta+\phi] \\ \sin[\theta+\phi] & \cos[\theta+\phi] \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & -\cos\theta\sin\phi - \sin\theta\cos\phi \\ \sin\theta\cos\phi + \cos\theta\sin\phi & \cos\theta\cos\phi - \sin\theta\sin\phi \end{bmatrix} \\ &= A_\theta A_\phi \end{aligned}$$

Don't these look familiar? They are the usual trigonometric identities for the sum of two angles derived here using linear algebra concepts.

Here we have focused on rotations in two dimensions. However, you can consider rotations and other geometric concepts in any number of dimensions. This is one of the major advantages of linear algebra. You can break down a difficult geometrical procedure into small steps, each corresponding to multiplication by an appropriate matrix. Then by multiplying the matrices, you can obtain a single matrix which can give you numerical information on the results of applying the given sequence of simple procedures.

Linear transformations which reflect vectors across a line are a second important type of transformations in $\mathbb{R}^2$. Consider the following theorem.

Theorem 5.4.2:Reflection

Let $Q_m : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation given by reflecting vectors over the line $\vec{y} = m\vec{x}$. Then the matrix of Q_m is given by

$$\frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}$$

Consider the following example.

Example 5.4.3:Reflection in $\mathbb{R}^2$

Let $Q_2 : \mathbb{R}^2 \to \mathbb{R}^2$ denote reflection over the line $\vec{y} = 2\vec{x}$. Then Q_2 is a linear transformation. Find the matrix of Q_2. Then, find $Q_2(\vec{x})$ where $\vec{x} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$.

Solution

By Theorem 5.4.2, the matrix of Q_2 is given by

$$\frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix} = \frac{1}{1+(2)^2}\begin{bmatrix} 1-(2)^2 & 2(2) \\ 2(2) & (2)^2-1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -3 & 8 \\ 8 & 3 \end{bmatrix}$$

To find $Q_2(\vec{x})$ we multiply $\vec{x}$ by the matrix of Q_2 as follows:

$$\frac{1}{5}\begin{bmatrix} -3 & 8 \\ 8 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -\frac{19}{5} \\ \frac{2}{5} \end{bmatrix}$$

Below is a video on sketching a linear transformation of a rectangle given the transformation matrix (reflection).



Consider the following example which incorporates a reflection as well as a rotation of vectors.

Example 5.4.4:Rotation Followed by a Reflection

Find the matrix of the linear transformation which is obtained by first rotating all vectors through an angle of $\pi/6$ and then reflecting through the x axis.

Solution

By Theorem 5.4.1, the matrix of the transformation which involves rotating through an angle of $\pi/6$ is

$$\begin{bmatrix} \cos[\pi/6] & -\sin[\pi/6] \\ \sin[\pi/6] & \cos[\pi/6] \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}\sqrt{3} \end{bmatrix}$$

Reflecting across the x axis is the same action as reflecting vectors over the line $\vec{y} = m\vec{x}$ with $m = 0$. By Theorem 5.4.2, the matrix for the transformation which reflects all vectors through the x axis is

$$\frac{1}{1+m^2}\begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix} = \frac{1}{1+(0)^2}\begin{bmatrix} 1-(0)^2 & 2(0) \\ 2(0) & (0)^2-1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Therefore, the matrix of the linear transformation which first rotates through $\pi/6$ and then reflects through the x axis is given by

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}\sqrt{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \end{bmatrix}$$

Below is a video on describing the range or image of a linear transformation (line reflection).



This page titled 5.4: Special Linear Transformations in R² is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **5.4: Special Linear Transformations in R²** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

5.5: One-to-One and Onto Transformations

Outcomes

A. Determine if a linear transformation is onto or one to one.

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. We define the **range** or **image** of T as the set of vectors of $\mathbb{R}^m$ which are of the form $T(\vec{x})$ (equivalently, $A\vec{x}$) for some $\vec{x} \in \mathbb{R}^n$. It is common to write $T\mathbb{R}^n$, $T(\mathbb{R}^n)$, or $\text{Im}(T)$ to denote these vectors.

Lemma 5.5.1:Range of a Matrix Transformation

Let A be an $m \times n$ matrix where $A_1, \cdots, A_n$ denote the columns of A. Then, for a vector $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ in $\mathbb{R}^n$,

$$A\vec{x} = \sum_{k=1}^{n} x_k A_k$$

Therefore, $A(\mathbb{R}^n)$ is the collection of all linear combinations of these products.

Proof

This follows from the definition of matrix multiplication.

Below is a video on determining if function mappings are one to one and/or onto.



This section is devoted to studying two important characterizations of linear transformations, called one to one and onto. We define them now.

Definition 5.5.1: One to One

Suppose $\vec{x}_1$ and $\vec{x}_2$ are vectors in $\mathbb{R}^n$. A linear transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is called **one to one** (often written as $1-1$) if whenever $\vec{x}_1 \neq \vec{x}_2$ it follows that :

$$T(\vec{x}_1) \neq T(\vec{x}_2)$$

Equivalently, if $T(\vec{x}_1) = T(\vec{x}_2)$, then $\vec{x}_1 = \vec{x}_2$. Thus, T is one to one if it never takes two different vectors to the same vector.

Below is a video on one to one transformations.



The second important characterization is called onto.

Definition 5.5.2: Onto

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. Then T is called **onto** if whenever $\vec{x}_2 \in \mathbb{R}^m$ there exists $\vec{x}_1 \in \mathbb{R}^n$ such that $T(\vec{x}_1) = \vec{x}_2$.

We often call a linear transformation which is one-to-one an **injection**. Similarly, a linear transformation which is onto is often called a **surjection**.

Below is a video on one to one and onto transformation definitions.



Below is a video on onto transformations.



The following proposition is an important result.

Proposition 5.5.1:One to One

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. Then T is one to one if and only if $T(\vec{x}) = \vec{0}$ implies $\vec{x} = \vec{0}$.

Proof

We need to prove two things here. First, we will prove that if T is one to one, then $T(\vec{x}) = \vec{0}$ implies that $\vec{x} = \vec{0}$. Second, we will show that if $T(\vec{x}) = \vec{0}$ implies that $\vec{x} = \vec{0}$, then it follows that T is one to one. Recall that a linear transformation has the property that $T(\vec{0}) = \vec{0}$.

Suppose first that T is one to one and consider $T(\vec{0})$.

$$T(\vec{0}) = T\left(\vec{0} + \vec{0}\right) = T(\vec{0}) + T(\vec{0})$$

and so, adding the additive inverse of $T(\vec{0})$ to both sides, one sees that $T(\vec{0}) = \vec{0}$. If $T(\vec{x}) = \vec{0}$ it must be the case that $\vec{x} = \vec{0}$ because it was just shown that $T(\vec{0}) = \vec{0}$ and T is assumed to be one to one.

Now assume that if $T(\vec{x}) = \vec{0}$, then it follows that $\vec{x} = \vec{0}$. If $T(\vec{v}) = T(\vec{u})$, then

$$T(\vec{v}) - T(\vec{u}) = T(\vec{v} - \vec{u}) = \vec{0}$$

which shows that $\vec{v} - \vec{u} = 0$. In other words, $\vec{v} = \vec{u}$, and T is one to one.

Note that this proposition says that if $A = \begin{bmatrix} A_1 & \cdots & A_n \end{bmatrix}$ then A is one to one if and only if whenever

$$0 = \sum_{k=1}^{n} c_k A_k$$

it follows that each scalar $c_k = 0$.

Below is a video on comparing one to one and onto matrix transformations.



Below is a video on determining if transformation mappings are one to one and/or onto.



We will now take a look at an example of a one to one and onto linear transformation.

Example 5.5.1: A One to One and Onto Linear Transformation

Suppose

$$T\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}1&1\\1&2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

Then, $T:\mathbb{R}^2\to\mathbb{R}^2$ is a linear transformation. Is T onto? Is it one to one?

Solution

Recall that because T can be expressed as matrix multiplication, we know that T is a linear transformation. We will start by looking at onto. So suppose $\begin{bmatrix}a\\b\end{bmatrix}\in\mathbb{R}^2$. Does there exist $\begin{bmatrix}x\\y\end{bmatrix}\in\mathbb{R}^2$ such that $T\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}a\\b\end{bmatrix}$? If so, then since $\begin{bmatrix}a\\b\end{bmatrix}$ is an arbitrary vector in $\mathbb{R}^2$, it will follow that T is onto.

This question is familiar to you. It is asking whether there is a solution to the equation

$$\begin{bmatrix}1&1\\1&2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}a\\b\end{bmatrix}$$

This is the same thing as asking for a solution to the following system of equations.

$$\begin{aligned}x+y&=a\\x+2y&=b\end{aligned}$$

Set up the augmented matrix and row reduce.

$$\left[\begin{array}{cc|c}1&1&a\\1&2&b\end{array}\right]\to\left[\begin{array}{cc|c}1&0&2a-b\\0&1&b-a\end{array}\right] \tag{5.5.1}$$

You can see from this point that the system has a solution. Therefore, we have shown that for any a,b, there is a $\begin{bmatrix}x\\y\end{bmatrix}$ such that $T\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}a\\b\end{bmatrix}$. Thus T is onto.

Now we want to know if T is one to one. By Proposition 5.5.1 it is enough to show that $A\vec{x}=0$ implies $\vec{x}=0$. Consider the system $A\vec{x}=0$ given by:

$$\begin{bmatrix}1&1\\1&2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

This is the same as the system given by

$$x+y=0$$
$$x+2y=0$$

We need to show that the solution to this system is $x=0$ and $y=0$. By setting up the augmented matrix and row reducing, we end up with

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right]$$

This tells us that $x=0$ and $y=0$. Returning to the original system, this says that if

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

then

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

In other words, $A\vec{x}=0$ implies that $\vec{x}=0$. By Proposition 5.5.1, A is one to one, and so T is also one to one.

We also could have seen that T is one to one from our above solution for onto. By looking at the matrix given by (5.5.1), you can see that there is a **unique** solution given by $x=2a-b$ and $y=b-a$. Therefore, there is only one vector, specifically $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2a-b \\ b-a \end{bmatrix}$ such that $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$. Hence by Definition 5.5.1, T is one to one.

Example 5.5.2:An Onto Transformation

Let $T:\mathbb{R}^4 \mapsto \mathbb{R}^2$ be a linear transformation defined by

$$T\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a+d \\ b+c \end{bmatrix} \text{ for all } \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4$$

Prove that T is onto but not one to one.

Solution

You can prove that T is in fact linear.

To show that T is onto, let $\begin{bmatrix} x \\ y \end{bmatrix}$ be an arbitrary vector in $\mathbb{R}^2$. Taking the vector $\begin{bmatrix} x \\ y \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^4$ we have

$$T\begin{bmatrix} x \\ y \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} x+0 \\ y+0 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

This shows that T is onto.

By Proposition 5.5.1 T is one to one if and only if $T(\vec{x})=\vec{0}$ implies that $\vec{x}=\vec{0}$. Observe that

$$T\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1+-1 \\ 0+0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

There exists a nonzero vector $\vec{x}$ in $\mathbb{R}^4$ such that $T(\vec{x}) = \vec{0}$. It follows that T is not one to one.

Below is a video on determining if a linear transformation is one to one and/or onto.



Below is another video on determining if a linear transformation is one to one and/or onto.



The above examples demonstrate a method to determine if a linear transformation T is one to one or onto. It turns out that the matrix A of T can provide this information.

Theorem 5.5.1:Matrix of a One to One or Onto Transformation

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation induced by the $m \times n$ matrix A. Then T is one to one if and only if the rank of A is n. T is onto if and only if the rank of A is m.

Consider Example 5.5.2. Above we showed that T was onto but not one to one. We can now use this theorem to determine this fact about T.

Example 5.5.3:An Onto Transformation

Let $T : \mathbb{R}^4 \mapsto \mathbb{R}^2$ be a linear transformation defined by

$$T \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a+d \\ b+c \end{bmatrix} \text{ for all } \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4$$

Prove that T is onto but not one to one.

Solution

Using Theorem 5.5.1 we can show that T is onto but not one to one from the matrix of T. Recall that to find the matrix A of T, we apply T to each of the standard basis vectors $\vec{e}_i$ of $\mathbb{R}^4$. The result is the 2×4 matrix A given by

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

Fortunately, this matrix is already in reduced row-echelon form. The rank of A is 2. Therefore by the above theorem T is onto but not one to one.

Below is a video on classifying matrix multiplication: onto and one to one.



Below is a video on determining if a 4x4 matrix represents a one to one and/or onto transformation.



Recall that if S and T are linear transformations, we can discuss their composite denoted $S \circ T$. The following examines what happens if both S and T are onto.

Example 5.5.4: Composite of Onto Transformations

Let $T : \mathbb{R}^k \mapsto \mathbb{R}^n$ and $S : \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear transformations. If T and S are onto, then $S \circ T$ is onto.

Solution

Let $\vec{z} \in \mathbb{R}^m$. Since S is onto, there exists a vector $\vec{y} \in \mathbb{R}^n$ such that $S(\vec{y}) = \vec{z}$. Furthermore, since T is onto, there exists a vector $\vec{x} \in \mathbb{R}^k$ such that $T(\vec{x}) = \vec{y}$. Thus

$$\vec{z} = S(\vec{y}) = S(T(\vec{x})) = (ST)(\vec{x}),$$

showing that for each $\vec{z} \in \mathbb{R}^m$ there exists and $\vec{x} \in \mathbb{R}^k$ such that $(ST)(\vec{x}) = \vec{z}$. Therefore, $S \circ T$ is onto.

The next example shows the same concept with regards to one-to-one transformations.

Example 5.5.5: Composite of One to One Transformations

Let $T : \mathbb{R}^k \mapsto \mathbb{R}^n$ and $S : \mathbb{R}^n \mapsto \mathbb{R}^m$ be linear transformations. Prove that if T and S are one to one, then $S \circ T$ is one-to-one.

Solution

To prove that $S \circ T$ is one to one, we need to show that if $S(T(\vec{v})) = \vec{0}$ it follows that $\vec{v} = \vec{0}$. Suppose that $S(T(\vec{v})) = \vec{0}$. Since S is one to one, it follows that $T(\vec{v}) = \vec{0}$. Similarly, since T is one to one, it follows that $\vec{v} = \vec{0}$. Hence $S \circ T$ is one to one.

This page titled 5.5: One-to-One and Onto Transformations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **5.5: One-to-One and Onto Transformations** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

5.6: Isomorphisms

Outcomes

A. Determine if a linear transformation is an isomorphism.
B. Determine if two subspaces of $\mathbb{R}^n$ are isomorphic.

Recall the definition of a linear transformation. Let V and W be two subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. A mapping $T : V \to W$ is called a **linear transformation** or **linear map** if it preserves the algebraic operations of addition and scalar multiplication. Specifically, if a, b are scalars and $\vec{x}, \vec{y}$ are vectors,

$$T\left(a\vec{x} + b\vec{y}\right) = aT(\vec{x}) + bT(\vec{y})$$

Consider the following important definition.

Definition 5.6.1: Isomorphism

A linear map T is called an **isomorphism** if the following two conditions are satisfied.

- T is one to one. That is, if $T(\vec{x}) = T(\vec{y})$, then $\vec{x} = \vec{y}$.
- T is onto. That is, if $\vec{w} \in W$, there exists $\vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$.

Two such subspaces which have an isomorphism as described above are said to be **isomorphic.**

Consider the following example of an isomorphism.

Example 5.6.1:Isomorphism

Let $T : \mathbb{R}^2 \mapsto \mathbb{R}^2$ be defined by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x-y \end{bmatrix}$$

Show that T is an isomorphism.

Solution

To prove that T is an isomorphism we must show

1. T is a linear transformation;
2. T is one to one;
3. T is onto.

We proceed as follows.

1. T is a linear transformation:

   Let k, p be scalars.

$$\begin{aligned} T\left(k\begin{bmatrix}x_1\\y_1\end{bmatrix}+p\begin{bmatrix}x_2\\y_2\end{bmatrix}\right) &= T\left(\begin{bmatrix}kx_1\\ky_1\end{bmatrix}+\begin{bmatrix}px_2\\py_2\end{bmatrix}\right)\\ &= T\left(\begin{bmatrix}kx_1+px_2\\ky_1+py_2\end{bmatrix}\right)\\ &= \begin{bmatrix}(kx_1+px_2)+(ky_1+py_2)\\(kx_1+px_2)-(ky_1+py_2)\end{bmatrix}\\ &= \begin{bmatrix}(kx_1+ky_1)+(px_2+py_2)\\(kx_1-ky_1)+(px_2-py_2)\end{bmatrix}\\ &= \begin{bmatrix}kx_1+ky_1\\kx_1-ky_1\end{bmatrix}+\begin{bmatrix}px_2+py_2\\px_2-py_2\end{bmatrix}\\ &= k\begin{bmatrix}x_1+y_1\\x_1-y_1\end{bmatrix}+p\begin{bmatrix}x_2+y_2\\x_2-y_2\end{bmatrix}\\ &= kT\left(\begin{bmatrix}x_1\\y_1\end{bmatrix}\right)+pT\left(\begin{bmatrix}x_2\\y_2\end{bmatrix}\right) \end{aligned}$$

Therefore T is linear.

2. T is one to one:

We need to show that if $T(\vec{x})=\vec{0}$ for a vector $\vec{x}\in\mathbb{R}^2$, then it follows that $\vec{x}=\vec{0}$. Let $\vec{x}=\begin{bmatrix}x\\y\end{bmatrix}$.

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}x+y\\x-y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

This provides a system of equations given by

$$\begin{aligned} x+y&=0\\ x-y&=0 \end{aligned}$$

You can verify that the solution to this system if $x=y=0$. Therefore

$$\vec{x}=\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

and T is one to one.

3. T is onto:

Let a,b be scalars. We want to check if there is always a solution to

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}x+y\\x-y\end{bmatrix}=\begin{bmatrix}a\\b\end{bmatrix}$$

This can be represented as the system of equations

$$\begin{aligned} x+y&=a\\ x-y&=b \end{aligned}$$

Setting up the augmented matrix and row reducing gives

$$\left[\begin{array}{cc|c}1&1&a\\1&-1&b\end{array}\right]\rightarrow\cdots\rightarrow\left[\begin{array}{cc|c}1&0&\frac{a+b}{2}\\0&1&\frac{a-b}{2}\end{array}\right]$$

This has a solution for all a,b and therefore T is onto.

Therefore T is an isomorphism.

Below is a video on one to one and onto functions (isomorphisms).



An important property of isomorphisms is that its inverse is also an isomorphism.

Proposition 5.6.1:Inverse of an Isomorphism

Let $T : V \to W$ be an isomorphism and V, W be subspaces of $\mathbb{R}^n$. Then $T^{-1} : W \to V$ is also an isomorphism.

Proof

Let T be an isomorphism. Since T is onto, a typical vector in W is of the form $T(\vec{v})$ where $\vec{v} \in V$. Consider then for a, b scalars,

$$T^{-1}\left(aT(\vec{v}_1) + bT(\vec{v}_2)\right)$$

where $\vec{v}_1, \vec{v}_2 \in V$. Is this equal to

$$aT^{-1}\left(T(\vec{v}_1)\right) + bT^{-1}\left(T(\vec{v}_2)\right) = a\vec{v}_1 + b\vec{v}_2?$$

Since T is one to one, this will be so if

$$T\left(a\vec{v}_1 + b\vec{v}_2\right) = T\left(T^{-1}\left(aT(\vec{v}_1) + bT(\vec{v}_2)\right)\right) = aT(\vec{v}_1) + bT(\vec{v}_2).$$

However, the above statement is just the condition that T is a linear map. Thus T^{-1} is indeed a linear map. If $\vec{v} \in V$ is given, then $\vec{v} = T^{-1}\left(T(\vec{v})\right)$ and so T^{-1} is onto. If $T^{-1}(\vec{v}) = 0$, then

$$\vec{v} = T\left(T^{-1}(\vec{v})\right) = T(\vec{0}) = \vec{0}$$

and so T^{-1} is one to one.

Another important result is that the composition of multiple isomorphisms is also an isomorphism.

Proposition 5.6.2:Composition of Isomorphisms

Let $T : V \to W$ and $S : W \to Z$ be isomorphisms where V, W, Z are subspaces of $\mathbb{R}^n$. Then $S \circ T$ defined by $(S \circ T)(\vec{v}) = S\left(T(\vec{v})\right)$ is also an isomorphism.

Proof

Suppose $T : V \to W$ and $S : W \to Z$ are isomorphisms. Why is $S \circ T$ a linear map? For a, b scalars,

$$\begin{aligned} S \circ T\left(a\vec{v}_1 + b(\vec{v}_2)\right) &= S\left(T\left(a\vec{v}_1 + b\vec{v}_2\right)\right) = S\left(aT\vec{v}_1 + bT\vec{v}_2\right) \\ &= aS\left(T\vec{v}_1\right) + bS\left(T\vec{v}_2\right) = a\left(S \circ T\right)(\vec{v}_1) + b\left(S \circ T\right)(\vec{v}_2) \end{aligned}$$

Hence $S \circ T$ is a linear map. If $(S \circ T)(\vec{v}) = 0$, then $S\left(T(\vec{v})\right) = 0$ and it follows that $T(\vec{v}) = \vec{0}$ and hence by this lemma again, $\vec{v} = \vec{0}$. Thus $S \circ T$ is one to one. It remains to verify that it is onto. Let $\vec{z} \in Z$. Then since S is onto, there

exists $\vec{w} \in W$ such that $S(\vec{w}) = \vec{z}$. Also, since T is onto, there exists $\vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$. It follows that $S(T(\vec{v})) = \vec{z}$ and so $S \circ T$ is also onto.

Consider two subspaces V and W, and suppose there exists an isomorphism mapping one to the other. In this way the two subspaces are related, which we can write as $V \sim W$. Then the previous two propositions together claim that $\sim$ is an equivalence relation. That is: $\sim$ satisfies the following conditions:

- $V \sim V$
- If $V \sim W$, it follows that $W \sim V$
- If $V \sim W$ and $W \sim Z$, then $V \sim Z$

We leave the verification of these conditions as an exercise.

Consider the following example.

Example 5.6.2:Matrix Isomorphism

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be defined by $T(\vec{x}) = A(\vec{x})$ where A is an invertible $n \times n$ matrix. Then T is an isomorphism.

Solution

The reason for this is that, since A is invertible, the only vector it sends to $\vec{0}$ is the zero vector. Hence if $A(\vec{x}) = A(\vec{y})$, then $A(\vec{x} - \vec{y}) = \vec{0}$ and so $\vec{x} = \vec{y}$. It is onto because if

$$\vec{y} \in \mathbb{R}^n, A\left(A^{-1}(\vec{y})\right) = \left(AA^{-1}\right)(\vec{y}) = \vec{y}.$$

In fact, all isomorphisms from $\mathbb{R}^n$ to $\mathbb{R}^n$ can be expressed as $T(\vec{x}) = A(\vec{x})$ where A is an invertible $n \times n$ matrix. One simply considers the matrix whose i^{th} column is $T\vec{e}_i$.

Recall that a basis of a subspace V is a set of linearly independent vectors which span V. The following fundamental lemma describes the relation between bases and isomorphisms.

Lemma 5.6.1:Mapping Bases

Let $T : V \to W$ be a linear transformation where V, W are subspaces of $\mathbb{R}^n$. If T is one to one, then it has the property that if $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is linearly independent, so is $\{T(\vec{u}_1), \cdots, T(\vec{u}_k)\}$.

More generally, T is an isomorphism if and only if whenever $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for V, it follows that $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is a basis for W.

Proof

First suppose that T is a linear transformation and is one to one and $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is linearly independent. It is required to show that $\{T(\vec{u}_1), \cdots, T(\vec{u}_k)\}$ is also linearly independent. Suppose then that

$$\sum_{i=1}^{k} c_i T(\vec{u}_i) = \vec{0}$$

Then, since T is linear,

$$T\left(\sum_{i=1}^{n} c_i \vec{u}_i\right) = \vec{0}$$

Since T is one to one, it follows that

$$\sum_{i=1}^{n} c_i \vec{u}_i = 0$$

Now the fact that $\{\vec{u}_1, \cdots, \vec{u}_n\}$ is linearly independent implies that each $c_i = 0$. Hence $\{T(\vec{u}_1), \cdots, T(\vec{u}_n)\}$ is linearly independent.

Now suppose that T is an isomorphism and $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for V. It was just shown that $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is linearly independent. It remains to verify that span$\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\} = W$. If $\vec{w} \in W$, then since T is onto there exists $\vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$. Since $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis, it follows that there exists scalars $\{c_i\}_{i=1}^n$ such that

$$\sum_{i=1}^n c_i \vec{v}_i = \vec{v}.$$

Hence,

$$\vec{w} = T(\vec{v}) = T\left(\sum_{i=1}^n c_i \vec{v}_i\right) = \sum_{i=1}^n c_i T(\vec{v}_i)$$

It follows that span$\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\} = W$ showing that this set of vectors is a basis for W.

Next suppose that T is a linear transformation which takes a basis to a basis. This means that if $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for V, it follows $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is a basis for W. Then if $w \in W$, there exist scalars c_i such that $w = \sum_{i=1}^n c_i T(\vec{v}_i) = T\left(\sum_{i=1}^n c_i \vec{v}_i\right)$ showing that T is onto. If $T\left(\sum_{i=1}^n c_i \vec{v}_i\right) = \vec{0}$ then $\sum_{i=1}^n c_i T(\vec{v}_i) = \vec{0}$ and since the vectors $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ are linearly independent, it follows that each $c_i = 0$. Since $\sum_{i=1}^n c_i \vec{v}_i$ is a typical vector in V, this has shown that if $T(\vec{v}) = \vec{0}$ then $\vec{v} = \vec{0}$ and so T is also one to one. Thus T is an isomorphism.

The following theorem illustrates a very useful idea for defining an isomorphism. Basically, if you know what it does to a basis, then you can construct the isomorphism.

Theorem 5.6.1:Isomorphic Subspaces

Suppose V and W are two subspaces of $\mathbb{R}^n$. Then the two subspaces are isomorphic if and only if they have the same dimension. In the case that the two subspaces have the same dimension, then for a linear map $T : V \to W$, the following are equivalent.

1. T is one to one.
2. T is onto.
3. T is an isomorphism.

Proof

Suppose first that these two subspaces have the same dimension. Let a basis for V be $\{\vec{v}_1, \cdots, \vec{v}_n\}$ and let a basis for W be $\{\vec{w}_1, \cdots, \vec{w}_n\}$. Now define T as follows.

$$T(\vec{v}_i) = \vec{w}_i$$

for $\sum_{i=1}^n c_i \vec{v}_i$ an arbitrary vector of V,

$$T\left(\sum_{i=1}^n c_i \vec{v}_i\right) = \sum_{i=1}^n c_i T\vec{v}_i = \sum_{i=1}^n c_i \vec{w}_i.$$

It is necessary to verify that this is well defined. Suppose then that

$$\sum_{i=1}^n c_i \vec{v}_i = \sum_{i=1}^n \hat{c}_i \vec{v}_i$$

Then

$$\sum_{i=1}^n (c_i - \hat{c}_i)\, \vec{v}_i = \vec{0}$$

and since $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis, $c_i = \hat{c}_i$ for each i. Hence

$$\sum_{i=1}^{n} c_i \vec{w}_i = \sum_{i=1}^{n} \hat{c}_i \vec{w}_i$$

and so the mapping is well defined. Also if a, b are scalars,

$$\begin{aligned} T\left(a\sum_{i=1}^{n} c_i\vec{v}_i + b\sum_{i=1}^{n} \hat{c}_i\vec{v}_i\right) &= T\left(\sum_{i=1}^{n} (ac_i + b\hat{c}_i)\,\vec{v}_i\right) = \sum_{i=1}^{n} (ac_i + b\hat{c}_i)\,\vec{w}_i \\ &= a\sum_{i=1}^{n} c_i\vec{w}_i + b\sum_{i=1}^{n} \hat{c}_i\vec{w}_i \\ &= aT\left(\sum_{i=1}^{n} c_i\vec{v}_i\right) + bT\left(\sum_{i=1}^{n} \hat{c}_i\vec{v}_i\right) \end{aligned}$$

Thus T is a linear transformation.

Now if

$$T\left(\sum_{i=1}^{n} c_i\vec{v}_i\right) = \sum_{i=1}^{n} c_i\vec{w}_i = \vec{0},$$

then since the $\{\vec{w}_1, \cdots, \vec{w}_n\}$ are independent, each $c_i = 0$ and so $\sum_{i=1}^{n} c_i\vec{v}_i = \vec{0}$ also. Hence T is one to one. If $\sum_{i=1}^{n} c_i\vec{w}_i$ is a vector in W, then it equals

$$\sum_{i=1}^{n} c_i T(\vec{v}_i) = T\left(\sum_{i=1}^{n} c_i\vec{v}_i\right)$$

showing that T is also onto. Hence T is an isomorphism and so V and W are isomorphic.

Next suppose $T : V \mapsto W$ is an isomorphism, so these two subspaces are isomorphic. Then for $\{\vec{v}_1, \cdots, \vec{v}_n\}$ a basis for V, it follows that a basis for W is $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ showing that the two subspaces have the same dimension.

Now suppose the two subspaces have the same dimension. Consider the three claimed equivalences.

First consider the claim that $1.) \Rightarrow 2.)$. If T is one to one and if $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for V, then $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is linearly independent. If it is not a basis, then it must fail to span W. But then there would exist $\vec{w} \notin \text{span}\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ and it follows that $\{T(\vec{v}_1), \cdots, T(\vec{v}_n), \vec{w}\}$ would be linearly independent which is impossible because there exists a basis for W of n vectors.

Hence $\text{span}\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\} = W$ and so $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is a basis. If $\vec{w} \in W$, there exist scalars c_i such that

$$\vec{w} = \sum_{i=1}^{n} c_i T(\vec{v}_i) = T\left(\sum_{i=1}^{n} c_i\vec{v}_i\right)$$

showing that T is onto. This shows that $1.) \Rightarrow 2.)$.

Next consider the claim that $2.) \Rightarrow 3.)$. Since $2.)$ holds, it follows that T is onto. It remains to verify that T is one to one. Since T is onto, there exists a basis of the form $\{T(\vec{v}_i), \cdots, T(\vec{v}_n)\}$. Then it follows that $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is linearly independent. Suppose

$$\sum_{i=1}^{n} c_i\vec{v}_i = \vec{0}$$

Then

$$\sum_{i=1}^{n} c_i T(\vec{v}_i) = \vec{0}$$

Hence each $c_i = 0$ and so, $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for V. Now it follows that a typical vector in V is of the form $\sum_{i=1}^{n} c_i\vec{v}_i$. If $T\left(\sum_{i=1}^{n} c_i\vec{v}_i\right) = \vec{0}$, it follows that

$$\sum_{i=1}^{n} c_i T(\vec{v}_i) = \vec{0}$$

and so, since $\{T(\vec{v}_i), \cdots, T(\vec{v}_n)\}$ is independent, it follows each $c_i = 0$ and hence $\sum_{i=1}^{n} c_i \vec{v}_i = \vec{0}$. Thus T is one to one as well as onto and so it is an isomorphism.

If T is an isomorphism, it is both one to one and onto by definition so $3.)$ implies both $1.)$ and $2.)$.

Note the interesting way of defining a linear transformation in the first part of the argument by describing what it does to a basis and then "extending it linearly" to the entire subspace.

Example 5.6.4: Isomorphic Subspaces

Let $V = \mathbb{R}^3$ and let W denote

$$\text{span}\left\{\begin{bmatrix}1\\2\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\2\\0\end{bmatrix}\right\}$$

Show that V and W are isomorphic.

Solution

First observe that these subspaces are both of dimension 3 and so they are isomorphic by Theorem 5.6.1. The three vectors which span W are easily seen to be linearly independent by making them the columns of a matrix and row reducing to the reduced row-echelon form.

You can exhibit an isomorphism of these two spaces as follows.

$$T(\vec{e}_1) = \begin{bmatrix}1\\2\\1\\1\end{bmatrix}, T(\vec{e}_2) = \begin{bmatrix}0\\1\\0\\1\end{bmatrix}, T(\vec{e}_3) = \begin{bmatrix}1\\1\\2\\0\end{bmatrix}$$

and extend linearly. Recall that the matrix of this linear transformation is just the matrix having these vectors as columns. Thus the matrix of this isomorphism is

$$\begin{bmatrix}1 & 0 & 1\\2 & 1 & 1\\1 & 0 & 2\\1 & 1 & 0\end{bmatrix}$$

You should check that multiplication on the left by this matrix does reproduce the claimed effect resulting from an application by T.

Consider the following example.

Example 5.6.5:Finding the Matrix of an Isomorphism

Let $V = \mathbb{R}^3$ and let W denote

$$\text{span}\left\{\begin{bmatrix}1\\2\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\2\\0\end{bmatrix}\right\}$$

Let $T : V \mapsto W$ be defined as follows.

$$T\begin{bmatrix}1\\1\\0\end{bmatrix}=\begin{bmatrix}1\\2\\1\\1\end{bmatrix},T\begin{bmatrix}0\\1\\1\end{bmatrix}=\begin{bmatrix}0\\1\\0\\1\end{bmatrix},T\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}1\\1\\2\\0\end{bmatrix}$$

Find the matrix of this isomorphism T.

Solution

First note that the vectors

$$\begin{bmatrix}1\\1\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}1\\1\\1\end{bmatrix}$$

are indeed a basis for $\mathbb{R}^3$ as can be seen by making them the columns of a matrix and using the reduced row-echelon form.

Now recall the matrix of T is a 4×3 matrix A which gives the same effect as T. Thus, from the way we multiply matrices,

$$A\begin{bmatrix}1&0&1\\1&1&1\\0&1&1\end{bmatrix}=\begin{bmatrix}1&0&1\\2&1&1\\1&0&2\\1&1&0\end{bmatrix}$$

Hence,

$$A=\begin{bmatrix}1&0&1\\2&1&1\\1&0&2\\1&1&0\end{bmatrix}\begin{bmatrix}1&0&1\\1&1&1\\0&1&1\end{bmatrix}^{-1}=\begin{bmatrix}1&0&0\\0&2&-1\\2&-1&1\\-1&2&-1\end{bmatrix}$$

Note how the span of the columns of this new matrix must be the same as the span of the vectors defining W.

This idea of defining a linear transformation by what it does on a basis works for linear maps which are not necessarily isomorphisms.

Example 5.6.6: Finding the Matrix of an Isomorphism

Let $V=\mathbb{R}^3$ and let W denote

$$\text{span}\left\{\begin{bmatrix}1\\0\\1\\1\end{bmatrix},\begin{bmatrix}0\\1\\0\\1\end{bmatrix},\begin{bmatrix}1\\1\\1\\2\end{bmatrix}\right\}$$

Let $T:V\mapsto W$ be defined as follows.

$$T\begin{bmatrix}1\\1\\0\end{bmatrix}=\begin{bmatrix}1\\0\\1\\1\end{bmatrix},T\begin{bmatrix}0\\1\\1\end{bmatrix}=\begin{bmatrix}0\\1\\0\\1\end{bmatrix},T\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}1\\1\\1\\2\end{bmatrix}$$

Find the matrix of this linear transformation.

Solution

Note that in this case, the three vectors which span W are not linearly independent. Nevertheless the above procedure will still work. The reasoning is the same as before. If A is this matrix, then

$$A\begin{bmatrix}1&0&1\\1&1&1\\0&1&1\end{bmatrix}=\begin{bmatrix}1&0&1\\0&1&1\\1&0&1\\1&1&2\end{bmatrix}$$

and so

$$A=\begin{bmatrix}1&0&1\\0&1&1\\1&0&1\\1&1&2\end{bmatrix}\begin{bmatrix}1&0&1\\1&1&1\\0&1&1\end{bmatrix}^{-1}=\begin{bmatrix}1&0&0\\0&0&1\\1&0&0\\1&0&1\end{bmatrix}$$

The columns of this last matrix are obviously not linearly independent.

Below is a video on determining a function output given an isomorphic transformation.



This page titled 5.6: Isomorphisms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **5.6: Isomorphisms** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

5.7: The Kernel and Image of A Linear Map

Outcomes

A. Describe the kernel and image of a linear transformation, and find a basis for each.

In this section we will consider the case where the linear transformation is not necessarily an isomorphism. First consider the following important definition.

Definition 5.7.1: Kernel and Image

Let V and W be subspaces of $\mathbb{R}^n$ and let $T : V \mapsto W$ be a linear transformation. Then the image of T denoted as $\text{im}\,(T)$ is defined to be the set

$$\text{im}\,(T) = \{T(\vec{v}) : \vec{v} \in V\}$$

In words, it consists of all vectors in W which equal $T(\vec{v})$ for some $\vec{v} \in V$.

The kernel of T, written $\ker(T)$, consists of all $\vec{v} \in V$ such that $T(\vec{v}) = \vec{0}$. That is,

$$\ker(T) = \left\{\vec{v} \in V : T(\vec{v}) = \vec{0}\right\}$$

Below is a video on the kernel and image of a linear transformation.



Below is a video on which vectors are in the range and kernel of a linear transformation.



Below is a video on which vectors are in the kernel of a 2x3 matrix.



Below is a video on finding a nonzero vector in the kernel of a linear transformation.



Below is another video on finding a nonzero vector in the kernel of a linear transformation.



It follows that $\text{im}\,(T)$ and $\ker(T)$ are subspaces of W and V respectively.

Proposition 5.7.1: Kernel and Image as Subspaces

Let V, W be subspaces of $\mathbb{R}^n$ and let $T : V \to W$ be a linear transformation. Then $\ker(T)$ is a subspace of V and $\text{im}\,(T)$ is a subspace of W.

Proof

First consider $\ker(T)$. It is necessary to show that if $\vec{v}_1, \vec{v}_2$ are vectors in $\ker(T)$ and if a, b are scalars, then $a\vec{v}_1 + b\vec{v}_2$ is also in $\ker(T)$. But

$$T\left(a\vec{v}_1 + b\vec{v}_2\right) = aT(\vec{v}_1) + bT(\vec{v}_2) = a\vec{0} + b\vec{0} = \vec{0}$$

Thus $\ker(T)$ is a subspace of V.

Next suppose $T(\vec{v}_1), T(\vec{v}_2)$ are two vectors in $\mathrm{im}\,(T)$. Then if a, b are scalars,

$$aT(\vec{v}_2) + bT(\vec{v}_2) = T\left(a\vec{v}_1 + b\vec{v}_2\right)$$

and this last vector is in $\mathrm{im}\,(T)$ by definition.

We will now examine how to find the kernel and image of a linear transformation and describe the basis of each.

Example 5.7.1: Kernel and Image of a Linear Transformation

Let $T : \mathbb{R}^4 \mapsto \mathbb{R}^2$ be defined by

$$T\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a-b \\ c+d \end{bmatrix}$$

Then T is a linear transformation. Find a basis for $\ker(T)$ and $\mathrm{im}(T)$.

Solution

You can verify that T is a linear transformation.

First we will find a basis for $\ker(T)$. To do so, we want to find a way to describe all vectors $\vec{x} \in \mathbb{R}^4$ such that $T(\vec{x}) = \vec{0}$. Let $\vec{x} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$ be such a vector. Then

$$T\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a-b \\ c+d \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The values of a, b, c, d that make this true are given by solutions to the system

$$\begin{aligned} a - b &= 0 \\ c + d &= 0 \end{aligned}$$

The solution to this system is $a = s, b = s, c = t, d = -t$ where s, t are scalars. We can describe $\ker(T)$ as follows.

$$\ker(T) = \left\{ \begin{bmatrix} s \\ s \\ t \\ -t \end{bmatrix} \right\} = \mathrm{span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix} \right\}$$

Notice that this set is linearly independent and therefore forms a basis for $\ker(T)$.

We move on to finding a basis for $\mathrm{im}(T)$. We can write the image of T as

$$\mathrm{im}(T) = \left\{ \begin{bmatrix} a-b \\ c+d \end{bmatrix} \right\}$$

We can write this in the form

$$\text{span} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

This set is clearly not linearly independent. By removing unnecessary vectors from the set we can create a linearly independent set with the same span. This gives a basis for $\text{im}(T)$ as

$$\text{im}(T) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Below is a video on finding the kernel of a transformation matrix.



Below is a video on finding the basis for the kernel of a linear transformation.



Below is a video on finding the kernel of a linear transformation given a matrix.



Recall that a linear transformation T is called one to one if and only if $T(\vec{x}) = \vec{0}$ implies $\vec{x} = \vec{0}$. Using the concept of kernel, we can state this theorem in another way.

Theorem 5.7.1: One to One and Kernel

Let T be a linear transformation where $\ker(T)$ is the kernel of T. Then T is one to one if and only if $\ker(T)$ consists of **only** the zero vector.

A major result is the relation between the dimension of the kernel and dimension of the image of a linear transformation. In the previous example $\ker(T)$ had dimension 2, and $\text{im}(T)$ also had dimension of 2. Is it a coincidence that the dimension of $\mathbb{M}_{22}$ is $4 = 2 + 2$? Consider the following theorem.

Theorem 5.7.2: Dimension of Kernel and Image

Let $T : V \to W$ be a linear transformation where V, W are subspaces of $\mathbb{R}^n$. Suppose the dimension of V is m. Then

$$m = \dim(\ker(T)) + \dim(\text{im}\,(T))$$

Proof

From Proposition 5.7.1, $\text{im}\,(T)$ is a subspace of W. We know that there exists a basis for $\text{im}\,(T)$, $\{T(\vec{v}_1), \cdots, T(\vec{v}_r)\}$. Similarly, there is a basis for $\ker(T)$, $\{\vec{u}_1, \cdots, \vec{u}_s\}$. Then if $\vec{v} \in V$, there exist scalars c_i such that

$$T(\vec{v}) = \sum_{i=1}^{r} c_i T(\vec{v}_i)$$

Hence $T\left(\vec{v} - \sum_{i=1}^{r} c_i \vec{v}_i\right) = 0$. It follows that $\vec{v} - \sum_{i=1}^{r} c_i \vec{v}_i$ is in $\ker(T)$. Hence there are scalars a_i such that

$$\vec{v} - \sum_{i=1}^{r} c_i \vec{v}_i = \sum_{j=1}^{s} a_j \vec{u}_j$$

Hence $\vec{v} = \sum_{i=1}^{r} c_i \vec{v}_i + \sum_{j=1}^{s} a_j \vec{u}_j$. Since $\vec{v}$ is arbitrary, it follows that

$$V = \text{span}\,\{\vec{u}_1, \cdots, \vec{u}_s, \vec{v}_1, \cdots, \vec{v}_r\}$$

If the vectors $\{\vec{u}_1, \cdots, \vec{u}_s, \vec{v}_1, \cdots, \vec{v}_r\}$ are linearly independent, then it will follow that this set is a basis. Suppose then that

$$\sum_{i=1}^{r} c_i \vec{v}_i + \sum_{j=1}^{s} a_j \vec{u}_j = 0$$

Apply T to both sides to obtain

$$\sum_{i=1}^{r} c_i T(\vec{v}_i) + \sum_{j=1}^{s} a_j T(\vec{u})_j = \sum_{i=1}^{r} c_i T(\vec{v}_i) = 0$$

Since $\{T(\vec{v}_1), \cdots, T(\vec{v}_r)\}$ is linearly independent, it follows that each $c_i = 0$. Hence $\sum_{j=1}^{s} a_j \vec{u}_j = 0$ and so, since the $\{\vec{u}_1, \cdots, \vec{u}_s\}$ are linearly independent, it follows that each $a_j = 0$ also. Therefore $\{\vec{u}_1, \cdots, \vec{u}_s, \vec{v}_1, \cdots, \vec{v}_r\}$ is a basis for V and so

$$n = s + r = \dim(\ker(T)) + \dim(\text{im}\,(T))$$

The above theorem leads to the next corollary.

Corollary 5.7.1

Let $T : V \to W$ be a linear transformation where V, W are subspaces of $\mathbb{R}^n$. Suppose the dimension of V is m. Then

$$\dim(\ker(T)) \leq m$$

$$\dim(\text{im}\,(T)) \leq m$$

This follows directly from the fact that $n = \dim(\ker(T)) + \dim(\text{im}\,(T))$.

Consider the following example.

Example 5.7.2

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by

$$T(\vec{x}) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \vec{x}$$

Then $\text{im}\,(T) = V$ is a subspace of $\mathbb{R}^3$ and T is an isomorphism of $\mathbb{R}^2$ and V. Find a 2×3 matrix A such that the restriction of multiplication by A to $V = \text{im}\,(T)$ equals T^{-1}.

Solution

Since the two columns of the above matrix are linearly independent, we conclude that $\dim(\text{im}(T)) = 2$ and therefore $\dim(\ker(T)) = 2 - \dim(\text{im}(T)) = 2 - 2 = 0$ by Theorem 5.7.2. Then by Theorem 5.7.1 it follows that T is one to one.

Thus T is an isomorphism of $\mathbb{R}^2$ and the two dimensional subspace of $\mathbb{R}^3$ which is the span of the columns of the given matrix. Now in particular,

$$T(\vec{e}_1) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \; T(\vec{e}_2) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Thus

$$T^{-1} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \vec{e}_1, \; T^{-1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \vec{e}_2$$

Extend T^{-1} to all of $\mathbb{R}^3$ by defining

$$T^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \vec{e}_1$$

Notice that the vectors

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

are linearly independent so T^{-1} can be extended linearly to yield a linear transformation defined on $\mathbb{R}^3$. The matrix of T^{-1} denoted as A needs to satisfy

$$A\begin{bmatrix}1&0&0\\1&0&1\\0&1&0\end{bmatrix}=\begin{bmatrix}1&0&1\\0&1&0\end{bmatrix}$$

and so

$$A=\begin{bmatrix}1&0&1\\0&1&0\end{bmatrix}\begin{bmatrix}1&0&0\\1&0&1\\0&1&0\end{bmatrix}^{-1}=\begin{bmatrix}0&1&0\\0&0&1\end{bmatrix}$$

Note that

$$\begin{bmatrix}0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1\\1\\0\end{bmatrix}=\begin{bmatrix}1\\0\end{bmatrix}$$

$$\begin{bmatrix}0&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}0\\1\end{bmatrix}$$

so the restriction to V of matrix multiplication by this matrix yields T^{-1}.

Below is a video on describing the kernel of a linear transformation: projection onto y=x.



Below is a video on describing the kernel of a linear transformation: reflection across the y-axis.



This page titled 5.7: The Kernel and Image of A Linear Map is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **5.7: The Kernel and Image of A Linear Map** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

5.8: The Matrix of a Linear Transformation II

Outcomes

A. Find the matrix of a linear transformation with respect to general bases.

We begin this section with an important lemma.

Lemma 5.8.1:Mapping of a Basis

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ be an isomorphism. Then T maps any basis of $\mathbb{R}^n$ to another basis for $\mathbb{R}^n$.

Conversely, if $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a linear transformation which maps a basis of $\mathbb{R}^n$ to another basis of $\mathbb{R}^n$, then it is an isomorphism.

Proof

First, suppose $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a linear transformation which is one to one and onto. Let $\{\vec{v}_1, \cdots, \vec{v}_n\}$ be a basis for $\mathbb{R}^n$. We wish to show that $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is also a basis for $\mathbb{R}^n$.

First consider why it is linearly independent. Suppose $\sum_{k=1}^n a_k T(\vec{v}_k) = \vec{0}$. Then by linearity we have $T\left(\sum_{k=1}^n a_k \vec{v}_k\right) = \vec{0}$ and since T is one to one, it follows that $\sum_{k=1}^n a_k \vec{v}_k = \vec{0}$. This requires that each $a_k = 0$ because $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is independent, and it follows that $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is linearly independent.

Next take $\vec{w} \in \mathbb{R}^n$. Since T is onto, there exists $\vec{v} \in \mathbb{R}^n$ such that $T(\vec{v}) = \vec{w}$. Since $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis, in particular it is a spanning set and there are scalars b_k such that $T\left(\sum_{k=1}^n b_k \vec{v}_k\right) = T(\vec{v}) = \vec{w}$. Therefore $\vec{w} = \sum_{k=1}^n b_k T(\vec{v}_k)$ which is in the span $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$. Therefore, $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is a basis as claimed.

Suppose now that $T : \mathbb{R}^n \mapsto \mathbb{R}^n$ is a linear transformation such that $T(\vec{v}_i) = \vec{w}_i$ where $\{\vec{v}_1, \cdots, \vec{v}_n\}$ and $\{\vec{w}_1, \cdots, \vec{w}_n\}$ are two bases for $\mathbb{R}^n$.

To show that T is one to one, let $T\left(\sum_{k=1}^n c_k \vec{v}_k\right) = \vec{0}$. Then $\sum_{k=1}^n c_k T(\vec{v}_k) = \sum_{k=1}^n c_k \vec{w}_k = \vec{0}$. It follows that each $c_k = 0$ because it is given that $\{\vec{w}_1, \cdots, \vec{w}_n\}$ is linearly independent. Hence $T\left(\sum_{k=1}^n c_k \vec{v}_k\right) = \vec{0}$ implies that $\sum_{k=1}^n c_k \vec{v}_k = \vec{0}$ and so T is one to one.

To show that T is onto, let $\vec{w}$ be an arbitrary vector in $\mathbb{R}^n$. This vector can be written as $\vec{w} = \sum_{k=1}^n d_k \vec{w}_k = \sum_{k=1}^n d_k T(\vec{v}_k) = T\left(\sum_{k=1}^n d_k \vec{v}_k\right)$. Therefore, T is also onto.

Consider now an important definition.

Definition 5.8.1: Coordinate Vector

Let $B = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$ be a basis for $\mathbb{R}^n$ and let $\vec{x}$ be an arbitrary vector in $\mathbb{R}^n$. Then $\vec{x}$ is uniquely represented as $\vec{x} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n$ for scalars $a_1, \cdots, a_n$.

The **coordinate vector** of $\vec{x}$ with respect to the basis B, written $C_B(\vec{x})$ or $[\vec{x}]_B$, is given by

$$C_B(\vec{x}) = C_B\left(a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n\right) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Below is a video on change of basis between the standard basis and a nonstandard basis.



Below is a video on change of basis between two nonstandard bases.



Consider the following example.

Example 5.8.1: Coordinate Vector

Let $B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$ be a basis of $\mathbb{R}^2$ and let $\vec{x} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ be a vector in $\mathbb{R}^2$. Find $C_B(\vec{x})$.

Solution

First, note the order of the basis is important so label the vectors in the basis B as

$$B = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} = \{\vec{v}_1, \vec{v}_2\}$$

Now we need to find a_1, a_2 such that $\vec{x} = a_1\vec{v}_1 + a_2\vec{v}_2$, that is:

$$\begin{bmatrix} 3 \\ -1 \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Solving this system gives $a_1 = 2, a_2 = -1$. Therefore the coordinate vector of $\vec{x}$ with respect to the basis B is

$$C_B(\vec{x}) = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

Below is a video on finding a vector given basis vectors and its coordinates with respect to this basis.



Below is a video on finding the coordinates of a vector with respect to a given basis knowing its coordinates with respect to the standard basis.



Below is another video on finding a vector given basis vectors and its coordinates with respect to this basis.



Below is a video on finding a vector given basis vectors and its coordinates with respect to this basis.



Below is a video on finding the coordinates of a vector with respect to the standard basis knowing its coordinates with respect to a given basis.



Given any basis B, one can easily verify that the coordinate function is actually an isomorphism.

Theorem 5.8.1:C_B Transformation is a Linear

For any basis B of $\mathbb{R}^n$, the coordinate function

$$C_B : \mathbb{R}^n \to \mathbb{R}^n$$

is a linear transformation, and moreover an isomorphism.

We now discuss the main result of this section, that is how to represent a linear transformation with respect to different bases.

Theorem 5.8.2:The Matrix of a Linear

Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation, and let B_1 and B_2 be bases of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively.

Then the following holds

$$C_{B_2} T = M_{B_2 B_1} C_{B_1} \tag{5.8.1}$$

where $M_{B_2 B_1}$ is a unique $m \times n$ matrix.

If the basis B_1 is given by $B_1 = \{\vec{v}_1, \cdots, \vec{v}_n\}$ in this order, then

$$M_{B_2 B_1} = [C_{B_2}(T(\vec{v}_1)) \; C_{B_2}(T(\vec{v}_2)) \; \cdots C_{B_2}(T(\vec{v}_n))]$$

Proof

The above equation (5.8.1) can be represented by the following diagram.

$$\begin{array}{ccc} & T & \\ \mathbb{R}^n & \rightarrow & \mathbb{R}^m \\ C_{B_1}\downarrow & \circ & \downarrow C_{B_2} \\ \mathbb{R}^n & \rightarrow & \mathbb{R}^m \\ & M_{B_2B_1} & \end{array}$$

Since C_{B_1} is an isomorphism, then the matrix we are looking for is the matrix of the linear transformation

$$C_{B_2}TC_{B_1}^{-1}:\mathbb{R}^n \mapsto \mathbb{R}^m.$$

By Theorem 5.2.2, the columns are given by the image of the standard basis $\{\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n\}$. But since $C_{B_1}^{-1}(\vec{e}_i)=\vec{v}_i$, we readily obtain that

$$\begin{aligned} M_{B_2B_1} &= \left[C_{B_2}TC_{B_1}^{-1}(\vec{e}_1)\ \ C_{B_2}TC_{B_1}^{-1}(\vec{2}_2)\ \ \cdots\ \ C_{B_2}TC_{B_1}^{-1}(\vec{e}_n)\right] \\ &= \left[C_{B_2}(T(\vec{v}_1))\ \ C_{B_2}(T(\vec{v}_2))\ \ \cdots\ \ C_{B_2}(T(\vec{v}_n))\right] \end{aligned}$$

and this completes the proof.

Consider the following example.

Example 5.8.2:Matrix of a Linear

Let $T:\mathbb{R}^2\mapsto\mathbb{R}^2$ be a linear transformation defined by $T\left(\begin{bmatrix}a\\b\end{bmatrix}\right)=\begin{bmatrix}b\\a\end{bmatrix}$.

Consider the two bases

$$B_1=\{\vec{v}_1,\vec{v}_2\}=\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}-1\\1\end{bmatrix}\right\}$$

and

$$B_2=\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\-1\end{bmatrix}\right\}$$

Find the matrix M_{B_2,B_1} of T with respect to the bases B_1 and B_2.

Solution

By Theorem 5.8.2, the columns of $M_{B_2B_1}$ are the coordinate vectors of $T(\vec{v}_1),T(\vec{v}_2)$ with respect to B_2.

Since

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}0\\1\end{bmatrix},$$

a standard calculation yields

$$\begin{bmatrix}0\\1\end{bmatrix}=\left(\frac{1}{2}\right)\begin{bmatrix}1\\1\end{bmatrix}+\left(-\frac{1}{2}\right)\begin{bmatrix}1\\-1\end{bmatrix},$$

the first column of $M_{B_2B_1}$ is $\begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\end{bmatrix}$.

The second column is found in a similar way. We have

$$T\left(\begin{bmatrix}-1\\1\end{bmatrix}\right)=\begin{bmatrix}1\\-1\end{bmatrix},$$

and with respect to B_2 calculate:

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Hence the second column of $M_{B_2B_1}$ is given by $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. We thus obtain

$$M_{B_2B_1} = \begin{bmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{bmatrix}$$

We can verify that this is the correct matrix $M_{B_2B_1}$ on the specific example

$$\vec{v} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

First applying T gives

$$T(\vec{v}) = T\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} \right) = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

and one can compute that

$$C_{B_2}\left(\begin{bmatrix} -1 \\ 3 \end{bmatrix} \right) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

On the other hand, one compute $C_{B_1}(\vec{v})$ as

$$C_{B_1}\left(\begin{bmatrix} 3 \\ -1 \end{bmatrix} \right) = \begin{bmatrix} 2 \\ -1 \end{bmatrix},$$

and finally applying $M_{B_1B_2}$ gives

$$\begin{bmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

as above.

We see that the same vector results from either method, as suggested by Theorem 5.8.2.

If the bases B_1 and B_2 are equal, say B, then we write M_B instead of M_{BB}. The following example illustrates how to compute such a matrix. Note that this is what we did earlier when we considered only $B_1 = B_2$ to be the standard basis.

Example 5.8.3:Matrix of a Linear Transformation with respect to an Arbitrary Basis

Consider the basis B of $\mathbb{R}^3$ given by

$$B = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

And let $T : \mathbb{R}^3 \mapsto \mathbb{R}^3$ be the linear transformation defined on B as:

$$T \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, T \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, T \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

1. Find the matrix M_B of T relative to the basis B.
2. Then find the usual matrix of T with respect to the standard basis of $\mathbb{R}^3$.

Solution

Equation (5.8.1) gives $C_B T = M_B C_B$, and thus $M_B = C_B T C_B^{-1}$.

Now $C_B(\vec{v}_i) = \vec{e}_i$, so the matrix of C_B^{-1} (with respect to the standard basis) is given by

$$\left[C_B^{-1}(\vec{e}_1) \;\; C_B^{-1}(\vec{e}_2) \;\; C_B^{-1}(\vec{e}_2)\right] = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Moreover the matrix of TC_B^{-1} is given by

$$\left[TC_B^{-1}(\vec{e}_1) \;\; TC_B^{-1}(\vec{e}_2) \;\; TC_B^{-1}(\vec{e}_2)\right] = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

Thus

$$\begin{aligned} M_B &= C_B T C_B^{-1} = [C_B^{-1}]^{-1}[TC_B^{-1}] \\ &= \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -5 & 1 \\ -1 & 4 & 0 \\ 0 & -2 & 1 \end{bmatrix} \end{aligned}$$

Consider how this works. Let $\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ be an arbitrary vector in $\mathbb{R}^3$.

Apply C_B^{-1} to $\vec{b}$ to get

$$b_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + b_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + b_3 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

Apply T to this linear combination to obtain

$$b_1 \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + b_2 \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + b_3 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} b_1 + b_2 \\ -b_1 + 2b_2 + b_3 \\ b_1 - b_2 + b_3 \end{bmatrix}$$

Now take the matrix M_B of the transformation (as found above) and multiply it by $\vec{b}$.

$$\begin{bmatrix} 2 & -5 & 1 \\ -1 & 4 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 2b_1 - 5b_2 + b_3 \\ -b_1 + 4b_2 \\ -2b_2 + b_3 \end{bmatrix}$$

Is this the coordinate vector of the above relative to the given basis? We check as follows.

$$\begin{aligned} &(2b_1 - 5b_2 + b_3) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (-b_1 + 4b_2) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + (-2b_2 + b_3) \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} b_1 + b_2 \\ -b_1 + 2b_2 + b_3 \\ b_1 - b_2 + b_3 \end{bmatrix} \end{aligned}$$

You see it is the same thing.

Now lets find the matrix of T with respect to the standard basis. Let A be this matrix. That is, multiplication by A is the same as doing T. Thus

$$A\begin{bmatrix}1 & 1 & -1\\ 0 & 1 & 1\\ 1 & 1 & 0\end{bmatrix} = \begin{bmatrix}1 & 1 & 0\\ -1 & 2 & 1\\ 1 & -1 & 1\end{bmatrix}$$

Hence

$$A = \begin{bmatrix}1 & 1 & 0\\ -1 & 2 & 1\\ 1 & -1 & 1\end{bmatrix}\begin{bmatrix}1 & 1 & -1\\ 0 & 1 & 1\\ 1 & 1 & 0\end{bmatrix}^{-1} = \begin{bmatrix}0 & 0 & 1\\ 2 & 3 & -3\\ -3 & -2 & 4\end{bmatrix}$$

Of course this is a very different matrix than the matrix of the linear transformation with respect to the non standard basis.

This page titled 5.8: The Matrix of a Linear Transformation II is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **5.8: The Matrix of a Linear Transformation II** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

5.9: The General Solution of a Linear System

Outcomes

A. Use linear transformations to determine the particular solution and general solution to a system of equations.
B. Find the kernel of a linear transformation.

Recall the definition of a linear transformation discussed above. T is a **linear transformation** if whenever $\vec{x}, \vec{y}$ are vectors and k, p are scalars,

$$T\left(k\vec{x}+p\vec{y}\right)=kT\left(\vec{x}\right)+pT\left(\vec{y}\right)$$

Thus linear transformations distribute across addition and pass scalars to the outside.

It turns out that we can use linear transformations to solve linear systems of equations. Indeed given a system of linear equations of the form $A\vec{x}=\vec{b}$, one may rephrase this as $T(\vec{x})=\vec{b}$ where T is the linear transformation T_A induced by the coefficient matrix A. With this in mind consider the following definition.

Definition 5.9.1:Particular Solution of a System of Equations

Suppose a linear system of equations can be written in the form

$$T\left(\vec{x}\right)=\vec{b}$$

If $T\left(\vec{x}_p\right)=\vec{b},$ then $\vec{x}_p$ is called a **particular solution** of the linear system.

Recall that a system is called homogeneous if every equation in the system is equal to 0. Suppose we represent a homogeneous system of equations by $T\left(\vec{x}\right)=0$. It turns out that the $\vec{x}$ for which $T\left(\vec{x}\right)=0$ are part of a special set called the **null space** of T. We may also refer to the null space as the **kernel** of T, and we write $ker\left(T\right)$.

Consider the following definition.

Definition 5.9.2:Null Space or Kernel of a Linear Transformation

Let T be a linear transformation. Define

$$\ker(T)=\left\{\vec{x}:T\left(\vec{x}\right)=\vec{0}\right\}$$

The kernel, $\ker(T)$ consists of the set of all vectors $\vec{x}$ for which $T(\vec{x})=\vec{0}$. This is also called the **null space** of T.

We may also refer to the kernel of T as the **solution space** of the equation $T\left(\vec{x}\right)=\vec{0}$.

Consider the following example.

Example 5.9.1:The Kernel of the Derivative

Let $\frac{d}{dx}$ denote the linear transformation defined on f, the functions which are defined on $\mathbb{R}$ and have a continuous derivative. Find $\ker\left(\frac{d}{dx}\right)$.

Solution

The example asks for functions f which the property that $\frac{df}{dx}=0$. As you may know from calculus, these functions are the constant functions. Thus $\ker\left(\frac{d}{dx}\right)$ is the set of constant functions.

Definition 5.9.2 states that $\ker(T)$ is the set of solutions to the equation,

$$T\left(\vec{x}\right)=\vec{0}$$

Since we can write $T(\vec{x})$ as $A\vec{x}$, you have been solving such equations for quite some time.

We have spent a lot of time finding solutions to systems of equations in general, as well as homogeneous systems. Suppose we look at a system given by $A\vec{x}=\vec{b}$, and consider the related homogeneous system. By this, we mean that we replace $\vec{b}$ by $\vec{0}$ and look at $A\vec{x}=\vec{0}$. It turns out that there is a very important relationship between the solutions of the original system and the solutions of the associated homogeneous system. In the following theorem, we use linear transformations to denote a system of equations. Remember that $T(\vec{x})=A\vec{x}$.

Theorem 5.9.1:Particular Solution and General Solution

Suppose $\vec{x}_p$ is a solution to the linear system given by ,

$$T(\vec{x})=\vec{b}$$

Then if $\vec{y}$ is any other solution to $T(\vec{x})=\vec{b}$, there exists $\vec{x}_0\in\ker(T)$ such that

$$\vec{y}=\vec{x}_p+\vec{x}_0$$

Hence, every solution to the linear system can be written as a sum of a particular solution, $\vec{x}_p$, and a solution $\vec{x}_0$ to the associated homogeneous system given by $T(\vec{x})=\vec{0}$.

Proof

Consider $\vec{y}-\vec{x}_p=\vec{y}+(-1)\vec{x}_p$. Then $T(\vec{y}-\vec{x}_p)=T(\vec{y})-T(\vec{x}_p)$. Since $\vec{y}$ and $\vec{x}_p$ are both solutions to the system, it follows that $T(\vec{y})=\vec{b}$ and $T(\vec{x}_p)=\vec{b}$.

Hence, $T(\vec{y})-T(\vec{x}_p)=\vec{b}-\vec{b}=\vec{0}$. Let $\vec{x}_0=\vec{y}-\vec{x}_p$. Then, $T(\vec{x}_0)=\vec{0}$ so $\vec{x}_0$ is a solution to the associated homogeneous system and so is in $\ker(T)$.

Sometimes people remember the above theorem in the following form. The solutions to the system $T(\vec{x})=\vec{b}$ are given by $\vec{x}_p+\ker(T)$ where $\vec{x}_p$ is a particular solution to $T(\vec{x})=\vec{b}$.

For now, we have been speaking about the kernel or null space of a linear transformation T. However, we know that every linear transformation T is determined by some matrix A. Therefore, we can also speak about the null space of a matrix. Consider the following example.

Example 5.9.2:The Null Space of a Matrix

Let

$$A=\begin{bmatrix}1&2&3&0\\2&1&1&2\\4&5&7&2\end{bmatrix}$$

Find $\text{null}(A)$. Equivalently, find the solutions to the system of equations $A\vec{x}=\vec{0}$.

Solution

We are asked to find $\left\{\vec{x}:A\vec{x}=\vec{0}\right\}$. In other words we want to solve the system, $A\vec{x}=\vec{0}$. Let $\vec{x}=\begin{bmatrix}x\\y\\z\\w\end{bmatrix}$. Then this amounts to solving

$$\begin{bmatrix}1&2&3&0\\2&1&1&2\\4&5&7&2\end{bmatrix}\begin{bmatrix}x\\y\\z\\w\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

This is the linear system

$$\begin{aligned} x+2y+3z &= 0 \\ 2x+y+z+2w &= 0 \\ 4x+5y+7z+2w &= 0 \end{aligned}$$

To solve, set up the augmented matrix and row reduce to find the reduced row-echelon form.

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 0 & 0 \\ 2 & 1 & 1 & 2 & 0 \\ 4 & 5 & 7 & 2 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|c} 1 & 0 & -\frac{1}{3} & \frac{4}{3} & 0 \\ 0 & 1 & \frac{5}{3} & -\frac{2}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

This yields $x = \frac{1}{3}z - \frac{4}{3}w$ and $y = \frac{2}{3}w - \frac{5}{3}z$. Since null (A) consists of the solutions to this system, it consists vectors of the form,

$$\begin{bmatrix} \frac{1}{3}z - \frac{4}{3}w \\ \frac{2}{3}w - \frac{5}{3}z \\ z \\ w \end{bmatrix} = z \begin{bmatrix} \frac{1}{3} \\ -\frac{5}{3} \\ 1 \\ 0 \end{bmatrix} + w \begin{bmatrix} -\frac{4}{3} \\ \frac{2}{3} \\ 0 \\ 1 \end{bmatrix}$$

Consider the following example.

Example 5.9.3:A General Solution

The **general solution** of a linear system of equations is the set of all possible solutions. Find the general solution to the linear system,

$$\begin{bmatrix} 1 & 2 & 3 & 0 \\ 2 & 1 & 1 & 2 \\ 4 & 5 & 7 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 9 \\ 7 \\ 25 \end{bmatrix}$$

given that $\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}$ is one solution.

Solution

Note the matrix of this system is the same as the matrix in Example 5.9.2. Therefore, from Theorem 5.9.1, you will obtain all solutions to the above linear system by adding a particular solution $\vec{x}_p$ to the solutions of the associated homogeneous system, $\vec{x}$. One particular solution is given above by

$$\vec{x}_p = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

Using this particular solution along with the solutions found in Example 5.9.2, we obtain the following solutions,

$$z \begin{bmatrix} \frac{1}{3} \\ -\frac{5}{3} \\ 1 \\ 0 \end{bmatrix} + w \begin{bmatrix} -\frac{4}{3} \\ \frac{2}{3} \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

Hence, any solution to the above linear system is of this form.

Below is a video on finding the solution as a sum of a particular solution and the homogeneous solution.



This page titled 5.9: The General Solution of a Linear System is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **5.9: The General Solution of a Linear System** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

5.E: Exercises

Exercise 5.E.1

Show the map T: $\mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\vec{x}) = A\vec{x}$ where A is an $m \times n$ matrix and $\vec{x}$ is an $m \times 1$ column vector is a linear transformation.

Answer

This result follows from the properties of matrix multiplication.

Exercise 5.E.2

Show that the function $T_{\vec{u}}$ defined by $T_{\vec{u}}(\vec{v}) = \vec{v} - proj_{\vec{u}}(\vec{v})$ is also a linear transformation.

Answer

$$\begin{aligned} T_{\vec{u}}(a\vec{v}+b\vec{w}) &= a\vec{v}+b\vec{w} - \frac{(a\vec{v}+b\vec{w} \bullet \vec{u})}{||\vec{u}||^2}\vec{u} \\ &= a\vec{v} - a\frac{(\vec{v} \bullet \vec{u})}{||\vec{u}||^2}\vec{u} + b\vec{w} - b\frac{(\vec{w} \bullet \vec{u})}{||\vec{u}||^2}\vec{u} \\ &= aT_{\vec{u}}(\vec{v}) + bT_{\vec{u}}(\vec{w}) \end{aligned}$$

Exercise 5.E.3

Let $\vec{u}$ be a fixed vector. The function $T_{\vec{u}}$ defined by $T_{\vec{u}}\vec{v} = \vec{u} + \vec{v}$ has the effect of translating all vectors by adding $\vec{u} \neq \vec{0}$. Show this is not a linear transformation. Explain why it is not possible to represent $T_{\vec{u}}$ in $\mathbb{R}^3$ by multiplying by a 3×3 matrix.

Answer

Linear transformations take $\vec{0}$ to $\vec{0}$ which T does not. Also $T_{\vec{a}}(\vec{u}+\vec{v}) \neq T_{\vec{a}}\vec{u} + T_{\vec{a}}\vec{v}$.

Exercise 5.E.4

Consider the following functions which map $\mathbb{R}^n$ to $\mathbb{R}^n$.

a. T multiplies the jth component of $\vec{x}$ by a nonzero number b.
b. T replaces the ith component of $\vec{x}$ with b times the jth component added to the ith component.
c. T switches the ith and jth components.

Show these functions are linear transformations and describe their matrices A such that $T(\vec{x}) = A\vec{x}$.

Answer

a. The matrix of T is the elementary matrix which multiplies the jth diagonal entry of the identity matrix by b.
b. The matrix of T is the elementary matrix which takes b times the jth row and adds to the ith row.
c. The matrix of T is the elementary matrix which switches the ith and the jth rows where the two components are in the ith and jth positions.

Exercise 5.E.5

You are given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ and you know that

$$T(A_i) = B_i$$

where $[A_1 \cdots a_n]^{-1}$ exists. Show that the matrix of T is of the form

$$[B_1 \cdots B_n][A_1 \cdots A_n]^{-1}$$

Answer

Suppose

$$\begin{bmatrix} \vec{c}_1^T \\ \vdots \\ \vec{c}_n^T \end{bmatrix} = [\,\vec{a}_1 \quad \cdots \quad \vec{a}_n\,]^{-1}$$

Thus $\vec{c}_i^T \vec{a}_j = \delta_{ij}$. Therefore

$$\begin{aligned} [\,\vec{b}_1 \quad \cdots \quad \vec{b}_n\,]\,[\,\vec{a}_1 \quad \cdots \quad \vec{a}_n\,]^{-1}\vec{a}_i &= [\,\vec{b}_1 \quad \cdots \quad \vec{b}_n\,] \begin{bmatrix} \vec{c}_1^T \\ \vdots \\ \vec{c}_n^T \end{bmatrix} \vec{a}_i \\ &= [\,\vec{b}_1 \quad \cdots \quad \vec{b}_n\,]\,\vec{e}_i \\ &= \vec{b}_i \end{aligned}$$

Thus $T\vec{a}_i = [\,\vec{b}_1 \quad \cdots \quad \vec{b}_n\,]\,[\,\vec{a}_1 \quad \cdots \quad \vec{a}_n\,]^{-1}\vec{a}_i = A\vec{a}_i$. If $\vec{x}$ is arbitrary, then since the matrix $[\,\vec{a}_1 \quad \cdots \quad \vec{a}_n\,]$ is invertible, there exists a unique $\vec{y}$ such that $[\,\vec{a}_1 \quad \cdots \quad \vec{a}_n\,]\,\vec{y} = \vec{x}$ Hence

$$T\vec{x} = T\left(\sum_{i=1}^n y_i\vec{a}_i\right) = \sum_{i=1}^n y_i T\vec{a}_i = \sum_{i=1}^n y_i A\vec{a}_i = A\left(\sum_{i=1}^n y_i\vec{a}_i\right) = A\vec{x}$$

Exercise $5.E.6$

Suppose T is a linear transformation such that

$$T\begin{bmatrix} 1 \\ 2 \\ -6 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix}$$
$$T\begin{bmatrix} -1 \\ -1 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix}$$
$$T\begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x}) = A\vec{x}$.

Answer

$$\begin{bmatrix} 5 & 1 & 5 \\ 1 & 1 & 3 \\ 3 & 5 & -2 \end{bmatrix}\begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & 1 \\ 4 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 37 & 17 & 11 \\ 17 & 7 & 5 \\ 11 & 14 & 6 \end{bmatrix}$$

Exercise $5.E.7$

Suppose T is a linear transformation such that

$$T\begin{bmatrix}1\\1\\-8\end{bmatrix}=\begin{bmatrix}1\\3\\1\end{bmatrix}$$
$$T\begin{bmatrix}-1\\0\\6\end{bmatrix}=\begin{bmatrix}2\\4\\1\end{bmatrix}$$
$$T\begin{bmatrix}0\\-1\\3\end{bmatrix}=\begin{bmatrix}6\\1\\-1\end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x})=A\vec{x}$.

Answer

$$\begin{bmatrix}1&2&6\\3&4&1\\1&1&-1\end{bmatrix}\begin{bmatrix}6&3&1\\5&3&1\\6&2&1\end{bmatrix}=\begin{bmatrix}52&21&9\\44&23&8\\5&4&1\end{bmatrix}$$

Exercise $5.E.8$

Suppose T is a linear transformation such that

$$T\begin{bmatrix}1\\3\\-7\end{bmatrix}=\begin{bmatrix}-3\\1\\3\end{bmatrix}$$
$$T\begin{bmatrix}-1\\-2\\6\end{bmatrix}=\begin{bmatrix}1\\3\\-3\end{bmatrix}$$
$$T\begin{bmatrix}0\\-1\\2\end{bmatrix}=\begin{bmatrix}5\\3\\-3\end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x})=A\vec{x}$.

Answer

$$\begin{bmatrix}-3&1&5\\1&3&3\\3&-3&-3\end{bmatrix}\begin{bmatrix}2&2&1\\1&2&1\\4&1&1\end{bmatrix}=\begin{bmatrix}15&1&3\\17&11&7\\-9&-3&-3\end{bmatrix}$$

Exercise $5.E.9$

Suppose T is a linear transformation such that

$$T\begin{bmatrix}1\\1\\-7\end{bmatrix}=\begin{bmatrix}3\\3\\3\end{bmatrix}$$
$$T\begin{bmatrix}-1\\0\\6\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}$$
$$T\begin{bmatrix}0\\-1\\2\end{bmatrix}=\begin{bmatrix}1\\3\\-1\end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x}) = A\vec{x}$.

Answer

$$\begin{bmatrix} 3 & 1 & 1 \\ 3 & 2 & 3 \\ 3 & 3 & -1 \end{bmatrix} \begin{bmatrix} 6 & 2 & 1 \\ 5 & 2 & 1 \\ 6 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 29 & 9 & 5 \\ 46 & 13 & 8 \\ 27 & 11 & 5 \end{bmatrix}$$

Exercise $5.E.10$

Suppose T is a linear transformation such that

$$T \begin{bmatrix} 1 \\ 2 \\ -18 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 5 \end{bmatrix}$$
$$T \begin{bmatrix} -1 \\ -1 \\ 15 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 5 \end{bmatrix}$$
$$T \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ -2 \end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x}) = A\vec{x}$.

Answer

$$\begin{bmatrix} 5 & 3 & 2 \\ 2 & 3 & 5 \\ 5 & 5 & -2 \end{bmatrix} \begin{bmatrix} 11 & 4 & 1 \\ 10 & 4 & 1 \\ 12 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 109 & 38 & 10 \\ 112 & 35 & 10 \\ 81 & 34 & 8 \end{bmatrix}$$

Exercise $5.E.11$

Consider the following functions $T : \mathbb{R}^3 \to \mathbb{R}^2$. Show that each is a linear transformation and determine for each the matrix A such that $T(\vec{x}) = A\vec{x}$.

a. $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x + 2y + 3z \\ 2y - 3x + z \end{bmatrix}$

b. $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7x + 2y + z \\ 3x - 11y + 2z \end{bmatrix}$

c. $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x + 2y + z \\ x + 2y + 6z \end{bmatrix}$

d. $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2y - 5x + z \\ x + y + z \end{bmatrix}$

Exercise $5.E.12$

Consider the following functions $T : \mathbb{R}^3 \to \mathbb{R}^2$. Explain why each of these functions T is not linear.

a. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+3z+1 \\ 2y-3x+z \end{bmatrix}$

b. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y^2+3z \\ 2y+3x+z \end{bmatrix}$

c. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \sin x+2y+3z \\ 2y+3x+z \end{bmatrix}$

d. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+3z \\ 2y+3x-\ln z \end{bmatrix}$

Exercise 5.E.13

Suppose

$$[A_1 \cdots A_n]^{-1}$$

exists where each $A_j \in \mathbb{R}^n$ and let vectors $\{B_1, \cdots, B_n\}$ in $\mathbb{R}^m$ be given. Show that there **always** exists a linear transformation T such that $T(A_i) = B_i$.

Exercise 5.E.14

Find the matrix for $T(\vec{w}) = proj_{\vec{v}}(\vec{w})$ where $\vec{v} = [\,1 \quad -2 \quad 3\,]^T$.

Answer

Recall that $proj_{\vec{u}}(\vec{v}) = \frac{\vec{v} \bullet \vec{u}}{||\vec{u}||^2}\vec{u}$ and so the desired matrix has ith column equal to $proj_{\vec{u}}(\vec{e}_i)$. Therefore, the matrix desired is

$$\frac{1}{14}\begin{bmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \\ 3 & -6 & 9 \end{bmatrix}$$

Exercise 5.E.15

Find the matrix for $T(\vec{w}) = proj_{\vec{v}}(\vec{w})$ where $\vec{v} = [\,1 \quad 5 \quad 3\,]^T$.

Answer

$$\frac{1}{35}\begin{bmatrix} 1 & 5 & 3 \\ 5 & 25 & 15 \\ 3 & 15 & 9 \end{bmatrix}$$

Exercise 5.E.16

Find the matrix for $T(\vec{w}) = proj_{\vec{v}}(\vec{w})$ where $\vec{v} = [\,1 \quad 0 \quad 3\,]^T$.

Answer

$$\frac{1}{10}\begin{bmatrix} 1 & 0 & 3 \\ 0 & 0 & 0 \\ 3 & 0 & 9 \end{bmatrix}$$

Exercise 5.E.17

Show that if a function $T:\mathbb{R}^n \to \mathbb{R}^m$ is linear, then it is always the case that $T(\vec{0}) = \vec{0}$.

Exercise 5.E.18

Let T be a linear transformation induced by the matrix $A = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$ and S a linear transformation induced by $B = \begin{bmatrix} 0 & -2 \\ 4 & 2 \end{bmatrix}$. Find matrix of $S \circ T$ and find $(S \circ T)(\vec{x})$ for $\vec{x} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$.

Answer

The matrix of $S \circ T$ is given by BA.

$$\begin{bmatrix} 0 & -2 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -4 \\ 10 & 8 \end{bmatrix}$$

Now, $(S \circ T)(\vec{x}) = (BA)\vec{x}$.

$$\begin{bmatrix} 2 & -4 \\ 10 & 8 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \end{bmatrix}$$

Exercise 5.E.19

Let T be a linear transformation and suppose $T\left(\begin{bmatrix} 1 \\ -4 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$. Suppose S is a linear transformation induced by the matrix $B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$. Find $(S \circ T)(\vec{x})$ for $\vec{x} = \begin{bmatrix} 1 \\ -4 \end{bmatrix}$.

Answer

To find $(S \circ T)(\vec{x})$ we compute $S(T(\vec{x}))$.

$$\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \end{bmatrix} = \begin{bmatrix} -4 \\ -11 \end{bmatrix}$$

Exercise 5.E.20

Let T be a linear transformation induced by the matrix $A = \begin{bmatrix} 2 & 3 \\ 1 & 1 \end{bmatrix}$ and S a lienar transformation induced by $B = \begin{bmatrix} -1 & 3 \\ 1 & -2 \end{bmatrix}$. Find matrix of $S \circ T$ and find $(S \circ T)(\vec{x})$ for $\vec{x} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$.

Exercise 5.E.21

Let T be a linear transformation induced by the matrix $A = \begin{bmatrix} 2 & 1 \\ 5 & 2 \end{bmatrix}$. Find the matrix of T^{-1}.

Answer

The matrix of T^{-1} is A^{-1}.

$$\begin{bmatrix} 2 & 1 \\ 5 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} -2 & 1 \\ 5 & -2 \end{bmatrix}$$

Exercise 5.E.22

Let T be a linear transformation induced by the matrix $A = \begin{bmatrix} 4 & -3 \\ 2 & -2 \end{bmatrix}$. Find the matrix of T^{-1}.

Exercise 5.E.23

Let T be a linear transformation and suppose $T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 9 \\ 8 \end{bmatrix}$, $T\left(\begin{bmatrix} 0 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} -4 \\ -3 \end{bmatrix}$. Find the matrix of T^{-1}.

Exercise 5.E.24

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $\pi/3$.

Answer

$$\begin{bmatrix} \cos\left(\frac{\pi}{3}\right) & -\sin\left(\frac{\pi}{3}\right) \\ \sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix}$$

Exercise 5.E.25

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $\pi/4$.

Answer

$$\begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & -\sin\left(\frac{\pi}{4}\right) \\ \sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix}$$

Exercise 5.E.26

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $-\pi/3$.

Answer

$$\begin{bmatrix} \cos\left(-\frac{\pi}{3}\right) & -\sin\left(-\frac{\pi}{3}\right) \\ \sin\left(-\frac{\pi}{3}\right) & \cos\left(-\frac{\pi}{3}\right) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix}$$

Exercise 5.E.27

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $2\pi/3$.

Answer

$$\begin{bmatrix} \cos\left(\frac{2\pi}{3}\right) & -\sin\left(\frac{2\pi}{3}\right) \\ \sin\left(\frac{2\pi}{3}\right) & \cos\left(\frac{2\pi}{3}\right) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}$$

Exercise 5.E.28

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $\pi/12$. Hint: Note that $\pi/12 = \pi/3 - \pi/4$.

Answer

$$\begin{bmatrix} \cos\left(\frac{\pi}{3}\right) & -\sin\left(\frac{\pi}{3}\right) \\ \sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right) \end{bmatrix} \begin{bmatrix} \cos\left(-\frac{\pi}{4}\right) & -\sin\left(-\frac{\pi}{4}\right) \\ \sin\left(-\frac{\pi}{4}\right) & \cos\left(-\frac{\pi}{4}\right) \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{4}\sqrt{2}\sqrt{3}+\frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2}-\frac{1}{4}\sqrt{2}\sqrt{3} \\ \frac{1}{4}\sqrt{2}\sqrt{3}-\frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2}\sqrt{3}+\frac{1}{4}\sqrt{2} \end{bmatrix}$$

? Exercise 5.E. 29

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $2\pi/3$ and then reflects across the x axis.

Answer

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\left(\frac{2\pi}{3}\right) & -\sin\left(\frac{2\pi}{3}\right) \\ \sin\left(\frac{2\pi}{3}\right) & \cos\left(\frac{2\pi}{3}\right) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & \frac{1}{2} \end{bmatrix}$$

? Exercise 5.E. 30

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $\pi/3$ and then reflects across the x axis.

Answer

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\left(\frac{\pi}{3}\right) & -\sin\left(\frac{\pi}{3}\right) \\ \sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \end{bmatrix}$$

? Exercise 5.E. 31

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $\pi/4$ and then reflects across the x axis.

Answer

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & -\sin\left(\frac{\pi}{4}\right) \\ \sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \end{bmatrix}$$

? Exercise 5.E. 32

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $\pi/6$ and then reflects across the x axis followed by a reflection across the y axis.

Answer

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\left(\frac{\pi}{6}\right) & -\sin\left(\frac{\pi}{6}\right) \\ \sin\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{6}\right) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\sqrt{3} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2}\sqrt{3} \end{bmatrix}$$

? Exercise 5.E. 33

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ across the x axis and then rotates every vector through an angle of $\pi/4$.

Answer

$$\begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & -\sin\left(\frac{\pi}{4}\right) \\ \sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \end{bmatrix}$$

? Exercise 5.E.34

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ across the y axis and then rotates every vector through an angle of $\pi/4$.

Answer

$$\begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & -\sin\left(\frac{\pi}{4}\right) \\ \sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix}$$

? Exercise 5.E.35

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ across the x axis and then rotates every vector through an angle of $\pi/6$.

Answer

$$\begin{bmatrix} \cos\left(\frac{\pi}{6}\right) & -\sin\left(\frac{\pi}{6}\right) \\ \sin\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{6}\right) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sqrt{3} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2}\sqrt{3} \end{bmatrix}$$

? Exercise 5.E.36

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ across the y axis and then rotates every vector through an angle of $\pi/6$.

Answer

$$\begin{bmatrix} \cos\left(\frac{\pi}{6}\right) & -\sin\left(\frac{\pi}{6}\right) \\ \sin\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{6}\right) \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}\sqrt{3} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2}\sqrt{3} \end{bmatrix}$$

? Exercise 5.E.37

Find the matrix for the linear transformation which rotates every vector in $\mathbb{R}^2$ through an angle of $5\pi/12$. Hint: Note that $5\pi/12 = 2\pi/3 - \pi/4$.

Answer

$$\begin{bmatrix} \cos\left(\frac{2\pi}{3}\right) & -\sin\left(\frac{2\pi}{3}\right) \\ \sin\left(\frac{2\pi}{3}\right) & \cos\left(\frac{2\pi}{3}\right) \end{bmatrix} \begin{bmatrix} \cos\left(-\frac{\pi}{4}\right) & -\sin\left(-\frac{\pi}{4}\right) \\ \sin\left(-\frac{\pi}{4}\right) & \cos\left(-\frac{\pi}{4}\right) \end{bmatrix} =$$

$$\begin{bmatrix} \frac{1}{4}\sqrt{2}\sqrt{3} - \frac{1}{4}\sqrt{2} & -\frac{1}{4}\sqrt{2}\sqrt{3} - \frac{1}{4}\sqrt{2} \\ \frac{1}{4}\sqrt{2}\sqrt{3} + \frac{1}{4}\sqrt{2} & \frac{1}{4}\sqrt{2}\sqrt{3} - \frac{1}{4}\sqrt{2} \end{bmatrix}$$

Note that it doesn't matter about the order in this case.

Exercise 5.E.38

Find the matrix of the linear transformation which rotates every vector in $\mathbb{R}^3$ counter clockwise about the z axis when viewed from the positive z axis through an angle of 30° and then reflects through the xy plane.

Answer

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\left(\frac{\pi}{6}\right) & -\sin\left(\frac{\pi}{6}\right) & 0 \\ \sin\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{6}\right) & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sqrt{3} & -\frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2}\sqrt{3} & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Exercise 5.E.39

Let $\vec{u} = \begin{bmatrix} a \\ b \end{bmatrix}$ be a unit vector in $\mathbb{R}^2$. Find the matrix which reflects all vectors across this vector, as shown in the following picture.

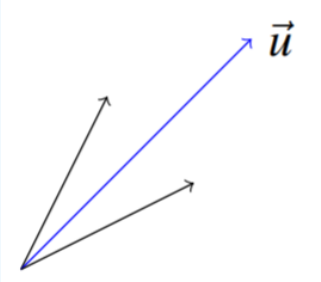


Figure 5.E.1

Hint: Notice that $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$ for some θ. First rotate through $-\theta$. Next reflect through the x axis. Finally rotate through θ.

Answer

$$\begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos(-\theta) & -\sin(\theta) \\ \sin(-\theta) & \cos(-\theta) \end{bmatrix}$$
$$= \begin{bmatrix} \cos^2\theta - \sin^2\theta & 2\cos\theta\sin\theta \\ 2\cos\theta\sin\theta & \sin^2\theta - \cos^2\theta \end{bmatrix}$$

Now to write in terms of (a,b), note that $a/\sqrt{a^2+b^2} = \cos\theta$, $b/\sqrt{a^2+b^2} = \sin\theta$. Now plug this in to the above. The result is

$$\begin{bmatrix} \frac{a^2-b^2}{a^2+b^2} & 2\frac{ab}{a^2+b^2} \\ 2\frac{ab}{a^2+b^2} & \frac{b^2-a^2}{a^2+b^2} \end{bmatrix} = \frac{1}{a^2+b^2} \begin{bmatrix} a^2-b^2 & 2ab \\ 2ab & b^2-a^2 \end{bmatrix}$$

Since this is a unit vector, $a^2+b^2 = 1$ and so you get

$$\begin{bmatrix} a^2-b^2 & 2ab \\ 2ab & b^2-a^2 \end{bmatrix}$$

Exercise 5.E.40

Let T be a linear transformation given by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Is T one to one? Is T onto?

Exercise 5.E.41

Let T be a linear transformation given by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 2 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Is T one to one? Is T onto?

Exercise 5.E.42

Let T be a linear transformation given by

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Is T one to one? Is T onto?

Exercise 5.E.43

Let T be a linear transformation given by

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 3 & -5 \\ 2 & 0 & 2 \\ 2 & 4 & -6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Is T one to one? Is T onto?

Exercise 5.E.44

Give an example of a 3×2 matrix with the property that the linear transformation determined by this matrix is one to one but not onto.

Answer

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Exercise 5.E.45

Suppose A is an $m \times n$ matrix in which $m \leq n$. Suppose also that the rank of A equals m. Show that the transformation T determined by A maps $\mathbb{R}^n$ onto $\mathbb{R}^m$. Hint: The vectors $\vec{e}_1, \cdots, \vec{e}_m$ occur as columns in the reduced row-echelon form for A.

Answer

This says that the columns of A have a subset of m vectors which are linearly independent. Therefore, this set of vectors is a basis for $\mathbb{R}^m$. It follows that the span of the columns is all of $\mathbb{R}^m$. Thus A is onto.

Exercise 5.E. 46

Suppose A is an $m \times n$ matrix in which $m \geq n$. Suppose also that the rank of A equals n. Show that A is one to one. Hint: If not, there exists a vector, $\vec{x}$ such that $A\vec{x} = 0$, and this implies at least one column of A is a linear combination of the others. Show this would require the rank to be less than n.

Answer

The columns are independent. Therefore, A is one to one.

Exercise 5.E. 47

Explain why an $n \times n$ matrix A is both one to one and onto if and only if its rank is n.

Answer

The rank is n is the same as saying the columns are independent which is the same as saying A is one to one which is the same as saying the columns are a basis. Thus the span of the columns of A is all of $\mathbb{R}^n$ and so A is onto. If A is onto, then the columns must be linearly independent since otherwise the span of these columns would have dimension less than n and so the dimension of $\mathbb{R}^n$ would be less than n.

Exercise 5.E. 48

Let V and W be subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and let $T : V \to W$ be a linear transformation. Suppose that $\{T\vec{v}_1, \cdots, T\vec{v}_r\}$ is linearly independent. Show that it must be the case that $\{T\vec{v}_1, \cdots, T\vec{v}_r\}$ is also linearly independent.

Answer

If $\sum_i^r a_i \vec{v}_r = 0$, then using linearity properties of T we get

$$0 = T(0) = T\left(\sum_i^r a_i \vec{v}_r\right) = \sum_i^r a_i T(\vec{v}_r).$$

Since we assume that $\{T\vec{v}_a, \cdots, T\vec{v}_r\}$ is linearly independent, we must have all $a_i = 0$, and therefore we conclude that $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is also linearly independent.

Exercise 5.E. 49

Let

$$V = span\left\{\begin{bmatrix}1\\1\\2\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\\1\end{bmatrix}\right\}$$

Let $T\vec{x} = A\vec{x}$ where A is the matrix

$$\begin{bmatrix}1 & 1 & 1 & 1\\0 & 1 & 1 & 0\\0 & 1 & 2 & 1\\1 & 1 & 1 & 2\end{bmatrix}$$

Give a basis for $im(T)$.

Exercise 5.E.50

Let

$$V = span\left\{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\4\\4\\1\end{bmatrix}\right\}$$

Let $T\vec{x} = A\vec{x}$ where A is the matrix

$$\begin{bmatrix}1 & 1 & 1 & 1\\0 & 1 & 1 & 0\\0 & 1 & 2 & 1\\1 & 1 & 1 & 2\end{bmatrix}$$

Find a basis for $im(T)$. In this case, the original vectors do not form an independent set.

Answer

Since the third vector is a linear combinations of the first two, then the image of the third vector will also be a linear combinations of the image of the first two. However the image of the first two vectors are linearly independent (check!), and hence form a basis of the image.

Thus a basis for $im(T)$ is:

$$V = span\left\{\begin{bmatrix}2\\0\\1\\3\end{bmatrix}, \begin{bmatrix}4\\2\\4\\5\end{bmatrix}\right\}$$

Exercise 5.E.51

If $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is linearly independent and T is a one to one linear transformation, show that $\{T\vec{v}_1, \cdots, T\vec{v}_r\}$ is also linearly independent. Give an example which shows that if T is only linear, it can happen that, although $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is linearly independent, $\{T\vec{v}_1, \cdots, T\vec{v}_r\}$ is not. In fact, show that it can happen that each of the $T\vec{v}_j$ equals 0.

Exercise 5.E.52

Let V and W be subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and let $T : V \to W$ be a linear transformation. Show that if T is onto W and if $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is a basis for V, then $span\{T\vec{v}_1, \cdots, T\vec{v}_r\} = W$.

Exercise 5.E.53

Define T: $\mathbb{R}^4 \to \mathbb{R}^3$ as follows.

$$T\vec{x} = \begin{bmatrix}3 & 2 & 1 & 8\\2 & 2 & -2 & 6\\1 & 1 & -1 & 3\end{bmatrix}\vec{x}$$

Find a basis for $im(T)$. Also find a basis for $\ker(T)$.

Exercise 5.E.54

Define $T: \mathbb{R}^3 \to \mathbb{R}^3$ as follows.

$$T\vec{x} = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \vec{x}$$

where on the right, it is just matrix multiplication of the vector $\vec{x}$ which is meant. Explain why T is an isomorphism of $\mathbb{R}^3$ to $\mathbb{R}^3$.

Exercise 5.E.55

Suppose $T: \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation given by

$$T\vec{x} = A\vec{x}$$

where A is a 3×3 matrix. Show that T is an isomorphism if and only if A is invertible.

Exercise 5.E.56

Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation given by

$$T\vec{x} = A\vec{x}$$

where A is an $m \times n$ matrix. Show that T is never an ismorphism if $m \neq n$. In particular, show that if $m > n$, T cannot be onto and if $m < n$, then T cannot be one to one.

Exercise 5.E.57

Define $T: \mathbb{R}^2 \to \mathbb{R}^3$ as follows.

$$T\vec{x} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \vec{x}$$

where on the right, it is just matrix multiplication of the vector $\vec{x}$ which is meant. Show that T is one to one. Next let $W = im(T)$. Show that T is an isomorphism of $\mathbb{R}^2$ and $im(T)$.

Exercise 5.E.58

In the above problem, find a 2×3 matrix A such that the restriction of A to $im(T)$ gives the same result as T^{-1} on $im(T)$. Hint: You might let A be such that

$$A \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; A \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

now find another vector $\vec{v} \in \mathbb{R}^3$ such that

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \vec{v} \right\}$$

is a basis. You could pick

$$\vec{v} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

for example. Explain why this one works or one of your choice works. Then you could define $A\vec{v}$ to equal some vector in $\mathbb{R}^2$. Explain why there will be more than one such matrix A which will deliver the inverse isomorphism T^{-1} on $im(T)$.

? Exercise $5.E.59$

Now let V equan $span \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ and let T: $V \to W$ be a linear transformation where

$$W = span \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

and

$$T \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, T \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Explain why T is an isomorphism. Determine a matrix A which, when multiplied on the left gives the same result as T on V and a matrix B which delivers T^{-1} on W. Hint: You need to have

$$A \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Now enlarge $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ to obtain a basis for $\mathbb{R}^3$. You could add in $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ for example, and then pick another vector in $\mathbb{R}^4$ and let $A \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ equal this other vector. Then you would have

$$A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

This would involve picking for the new vector in $\mathbb{R}^4$ the vector $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$. Then you could find A. You can do something similar to find a matrix for T^{-1} denoted as B.

? Exercise $5.E.60$

Let $V = \mathbb{R}^3$ and let

$$W = span(S), \text{ where } S = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} \right\}$$

Find a basis of W consisting of vectors in S.

Answer

In this case $\dim(W) = 1$ and a basis for W consisting of vectors in S can be obtained by taking any (nonzero) vector from S.

Exercise 5.E.61

Let T be a linear transformation given by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Find a basis for $\ker(T)$ and $im(T)$.

Answer

A basis for $\ker(T)$ is $\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$ and a basis for $im(T)$ is $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$. There are many other possibilities for the specific bases, but in this case $\dim(\ker(T)) = 1$ and $\dim(im(T)) = 1$.

Exercise 5.E.62

Let T be a linear transformation given by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Find a basis for $\ker(T)$ and $im(T)$.

Answer

In this case $\ker(T) = \{0\}$ and $im(T) = \mathbb{R}^2$ (pick any basis of $\mathbb{R}^2$).

Exercise 5.E.63

Let $V = \mathbb{R}^3$ and let

$$W = span\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}\right\}$$

Extend this basis of W to a basis of V.

Answer

There are many possible such extensions, one is (how do we know?):

$$\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$$

Exercise 5.E.64

Let T be a linear transformation given by

$$T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}1&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}$$

What is $\dim(\ker(T))$?

Answer

We can easily see that $\dim(im(T))=1$, and thus $\dim(\ker(T))=3-\dim(im(T))=3-1=2$.

Exercise 5.E.65

Let $B=\left\{\begin{bmatrix}2\\-1\end{bmatrix},\begin{bmatrix}3\\2\end{bmatrix}\right\}$ be a basis of $\mathbb{R}^2$ and let $\vec{x}=\begin{bmatrix}5\\-7\end{bmatrix}$ be a vector in $\mathbb{R}^2$. Find $C_B(\vec{x})$.

Exercise 5.E.66

Let $B=\left\{\begin{bmatrix}1\\-1\\2\end{bmatrix},\begin{bmatrix}2\\1\\2\end{bmatrix},\begin{bmatrix}-1\\0\\2\end{bmatrix}\right\}$ be a basis of $\mathbb{R}^3$ and let $\vec{x}=\begin{bmatrix}5\\-1\\4\end{bmatrix}$ be a vector in $\mathbb{R}^2$. Find $C_B(\vec{x})$.

Answer

$$C_B(\vec{x})=\begin{bmatrix}2\\1\\-1\end{bmatrix}$$

Exercise 5.E.67

Let T: $\mathbb{R}^2\to\mathbb{R}^2$ be a linear transformation defined by $T\left(\begin{bmatrix}a\\b\end{bmatrix}\right)=\begin{bmatrix}a+b\\a-b\end{bmatrix}$.

Consider the two bases

$$B_1=\{\vec{v}_1,\vec{v}_2\}=\left\{\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}-1\\1\end{bmatrix}\right\}$$

and

$$B_2=\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\-1\end{bmatrix}\right\}$$

Find the matrix M_{B_2,B_1} of T with respect to the bases B_1 and B_2.

Answer

$$M_{B_2B_1}=\begin{bmatrix}1&0\\-1&1\end{bmatrix}$$

Exercise 5.E.68

Write the solution set of the following system as a linear combination of vectors

$$\begin{bmatrix}1&-1&2\\1&-2&1\\3&-4&5\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} -3\hat{t} \\ -\hat{t} \\ \hat{t} \end{bmatrix}$, $\hat{t}_3 \in \mathbb{R}$. A basis for the solution space is $\begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix}$

Exercise 5.E.69

Using Exercise 5.E.68 find the general solution to the following linear system.

$$\begin{bmatrix} 1 & -1 & 2 \\ 1 & -2 & 1 \\ 3 & -4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Answer

Note that this has the same matrix as the above problem. Solution is: $\begin{bmatrix} -3\hat{t}_3 \\ -\hat{t}_3 \\ \hat{t}_3 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$, $\hat{t}_3 \in \mathbb{R}$

Exercise 5.E.70

Write the solution set of the following system as a linear combination of vectors

$$\begin{bmatrix} 0 & -1 & 2 \\ 1 & -2 & 1 \\ 1 & -4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} 3\hat{t} \\ 2\hat{t} \\ \hat{t} \end{bmatrix}$, A basis is $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

Exercise 5.E.71

Using Exercise 5.E.70 find the general solution to the following linear system.

$$\begin{bmatrix} 0 & -1 & 2 \\ 1 & -2 & 1 \\ 1 & -4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} 3\hat{t} \\ 2\hat{t} \\ \hat{t} \end{bmatrix} + \begin{bmatrix} -3 \\ -1 \\ 0 \end{bmatrix}$, $\hat{t} \in \mathbb{R}$

Exercise 5.E.72

Write the solution set of the following system as a linear combination of vectors.

$$\begin{bmatrix} 1 & -1 & 2 \\ 1 & -2 & 0 \\ 3 & -4 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} -4\hat{t} \\ -2\hat{t} \\ \hat{t} \end{bmatrix}$. A basis is $\begin{bmatrix} -4 \\ -2 \\ 1 \end{bmatrix}$

? Exercise 5.E.73

Using Exercise 5.E.72 find the general solution to the following linear system.

$$\begin{bmatrix} 1 & -1 & 2 \\ 1 & -2 & 0 \\ 3 & -4 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} -4\hat{t} \\ -2\hat{t} \\ \hat{t} \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \ \hat{t} \in \mathbb{R}$

? Exercise 5.E.74

Write the solution set of the following system as a linear combination of vectors

$$\begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & 1 \\ 1 & -2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} -\hat{t} \\ 2\hat{t} \\ \hat{t} \end{bmatrix}, \ \hat{t} \in \mathbb{R}$

? Exercise 5.E.75

Using Exercise 5.E.74 find the general solution to the following linear system.

$$\begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & 1 \\ 1 & -2 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} -\hat{t} \\ 2\hat{t} \\ \hat{t} \end{bmatrix} + \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$

? Exercise 5.E.76

Write the solution set of the following system as a linear combination of vectors

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & -1 & 1 & 0 \\ 3 & -1 & 3 & 2 \\ 3 & 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} 0 \\ -\hat{t} \\ -\hat{t} \\ \hat{t} \end{bmatrix}$, $\hat{t} \in \mathbb{R}$

? Exercise 5.E.77

Using Exercise 5.E.76 find the general solution to the following linear system.

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 1 & -1 & 1 & 0 \\ 3 & -1 & 3 & 2 \\ 3 & 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 3 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} 0 \\ -\hat{t} \\ -\hat{t} \\ \hat{t} \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \\ -1 \\ 0 \end{bmatrix}$

? Exercise 5.E.78

Write the solution set of the following system as a linear combination of vectors

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} -s-t \\ s \\ s \\ t \end{bmatrix}$, $s, t \in \mathbb{R}$. A basis is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

? Exercise 5.E.79

Using Exercise 5.E.78 find the general solution to the following linear system.

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -3 \\ 0 \end{bmatrix}$$

Answer

Solution is:

$$\begin{bmatrix} -\hat{t} \\ \hat{t} \\ \hat{t} \\ 0 \end{bmatrix} + \begin{bmatrix} -8 \\ 5 \\ 0 \\ 5 \end{bmatrix}$$

? Exercise $5.E.80$

Write the solution set of the following system as a linear combination of vectors

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 0 \\ 3 & 1 & 1 & 2 \\ 3 & 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Answer

Solution is:

$$\begin{bmatrix} -\frac{1}{2}s - \frac{1}{2}t \\ \frac{1}{2}s - \frac{1}{2}t \\ s \\ t \end{bmatrix}$$

for $s, t \in \mathbb{R}$. A basis is

$$\left\{ \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

? Exercise $5.E.81$

Using Exercise $5.E.80$ find the general solution to the following linear system.

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 0 \\ 3 & 1 & 1 & 2 \\ 3 & 3 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ 3 \end{bmatrix}$$

Answer

Solution is:

$$\begin{bmatrix} \frac{3}{2} \\ -\frac{1}{2} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -\frac{1}{2}s - \frac{1}{2}t \\ \frac{1}{2}s - \frac{1}{2}t \\ s \\ t \end{bmatrix}$$

Exercise 5.E.82

Write the solution set of the following system as a linear combination of vectors

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} -\hat{t} \\ \hat{t} \\ \hat{t} \\ 0 \end{bmatrix}$, a basis is $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$.

Exercise 5.E.83

Using Exercise 5.E.82 find the general solution to the following linear system.

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 2 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -3 \\ 1 \end{bmatrix}$$

Answer

Solution is: $\begin{bmatrix} -\hat{t} \\ \hat{t} \\ \hat{t} \\ 0 \end{bmatrix} + \begin{bmatrix} -9 \\ 5 \\ 0 \\ 6 \end{bmatrix}, t \in \mathbb{R}$.

Exercise 5.E.84

Suppose $A\vec{x} = \vec{b}$ has a solution. Explain why the solution is unique precisely when $A\vec{x} = \vec{0}$ has only the trivial solution.

Answer

If not, then there would be a infintely many solutions to $A\vec{x} = \vec{0}$ and each of these added to a solution to $A\vec{x} = \vec{b}$ would be a solution to $A\vec{x} = \vec{b}$.

5.E: Exercises is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

6: Complex Numbers

6.1: Complex Numbers
6.2: Polar Form
6.3: Roots of Complex Numbers
6.4: The Quadratic Formula
6.E: Exercises

This page titled 6: Complex Numbers is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

6.1: Complex Numbers

Outcomes

A. Understand the geometric significance of a complex number as a point in the plane.
B. Prove algebraic properties of addition and multiplication of complex numbers, and apply these properties. Understand the action of taking the conjugate of a complex number.
C. Understand the absolute value of a complex number and how to find it as well as its geometric significance.

Although very powerful, the real numbers are inadequate to solve equations such as $x^2+1=0$, and this is where complex numbers come in. We define the number i as the imaginary number such that $i^2=-1$, and define complex numbers as those of the form $z=a+bi$ where a and b are real numbers. We call this the standard form, or Cartesian form, of the complex number z. Then, we refer to a as the *real* part of z, and b as the *imaginary* part of z. It turns out that such numbers not only solve the above equation, but in fact also solve any polynomial of degree at least 1 with complex coefficients. This property, called the Fundamental Theorem of Algebra, is sometimes referred to by saying $\mathbb{C}$ is algebraically closed. Gauss is usually credited with giving a proof of this theorem in 1797 but many others worked on it and the first completely correct proof was due to Argand in 1806.

Just as a real number can be considered as a point on the line, a complex number $z=a+bi$ can be considered as a point (a,b) in the plane whose x coordinate is a and whose y coordinate is b. For example, in the following picture, the point $z=3+2i$ can be represented as the point in the plane with coordinates $(3,2)$.

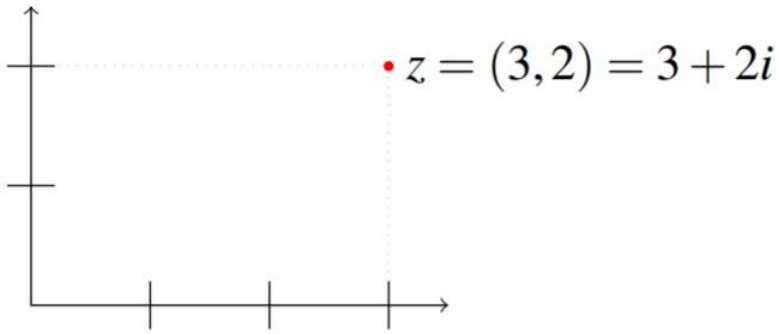


Figure 6.1.1

Addition of complex numbers is defined as follows.

$$(a+bi)+(c+di)=(a+c)+(b+d)\,i$$

This addition obeys all the usual properties as the following theorem indicates.

Theorem 6.1.1: Properties of Addition of Complex Numbers

Let z, w, and v be complex numbers. Then the following properties hold.

- Commutative Law for Addition

$$z+w=w+z$$

- Additive Identity

$$z+0=z$$

- Existence of Additive Inverse

$$\text{For each } z\in\mathbb{C}, \text{there exists } -z\in\mathbb{C} \text{ such that } z+(-z)=0$$
$$\text{In fact if } z=a+bi, \text{ then } -z=-a-bi.$$

- Associative Law for Addition

$$(z+w)+v=z+(w+v)$$

Proof

The proof of this theorem is left as an exercise for the reader.

Now, multiplication of complex numbers is defined the way you would expect, recalling that $i^2=-1$.

$$\begin{aligned}(a+bi)(c+di) &= ac+adi+bci+i^2bd\\ &= (ac-bd)+(ad+bc)i\end{aligned}$$

Consider the following examples.

Example 6.1.1: Multiplication of Complex Numbers

- $(2-3i)(-3+4i)=6+17i$
- $(4-7i)(6-2i)=10-50i$
- $(-3+6i)(5-i)=-9+33i$

The following are important properties of multiplication of complex numbers.

Theorem 6.1.2: Properties of Multiplication of Complex Numbers

Let z, w and v be complex numbers. Then, the following properties of multiplication hold.

- Commutative Law for Multiplication

$$zw=wz$$

- Associative Law for Multiplication

$$(zw)v=z(wv)$$

- Multiplicative Identity

$$1z=z$$

- Existence of Multiplicative Inverse

$$\text{For each } z\neq 0, \text{there exists } z^{-1} \text{ such that } zz^{-1}=1$$

- Distributive Law

$$z(w+v)=zw+zv$$

You may wish to verify some of these statements. The real numbers also satisfy the above axioms, and in general any mathematical structure which satisfies these axioms is called a field. There are many other fields, in particular even finite ones particularly useful for cryptography, and the reason for specifying these axioms is that linear algebra is all about fields and we can do just about anything in this subject using any field. Although here, the fields of most interest will be the familiar field of real numbers, denoted as $\mathbb{R}$, and the field of complex numbers, denoted as $\mathbb{C}$.

An important construction regarding complex numbers is the complex conjugate denoted by a horizontal line above the number, $\overline{z}$. It is defined as follows.

Definition 6.1.1: Conjugate of a Complex Number

Let $z=a+bi$ be a complex number. Then the **conjugate** of z, written $\overline{z}$ is given by

$$\overline{a+bi}=a-bi$$

Geometrically, the action of the conjugate is to reflect a given complex number across the x axis. Algebraically, it changes the sign on the imaginary part of the complex number. Therefore, for a real number a, $\overline{a} = a$.

Example 6.1.2: Conjugate of a Complex Number

- If $z = 3+4i$, then $\overline{z} = 3-4i$, i.e., $\overline{3+4i} = 3-4i$.
- $\overline{-2+5i} = -2-5i$.
- $\overline{i} = -i$.
- $\overline{7} = 7$.

Consider the following computation.

$$\begin{aligned}\left(\overline{a+bi}\right)(a+bi) &= (a-bi)(a+bi)\\ &= a^2+b^2-(ab-ab)i = a^2+b^2\end{aligned}$$

Notice that there is no imaginary part in the product, thus multiplying a complex number by its conjugate results in a real number.

Theorem 6.1.3: Properties of the Conjugate

Let z and w be complex numbers. Then, the following properties of the conjugate hold.

- $\overline{z \pm w} = \overline{z} \pm \overline{w}$.
- $\overline{(zw)} = \overline{z}\,\overline{w}$.
- $\overline{(\overline{z})} = z$.
- $\overline{\left(\frac{z}{w}\right)} = \frac{\overline{z}}{\overline{w}}$.
- z is real if and only if $\overline{z} = z$.

Division of complex numbers is defined as follows. Let $z = a+bi$ and $w = c+di$ be complex numbers such that c, d are not both zero. Then the quotient z divided by w is

$$\begin{aligned}\frac{z}{w} &= \frac{a+bi}{c+di}\\ &= \frac{a+bi}{c+di}\times\frac{c-di}{c-di}\\ &= \frac{(ac+bd)+(bc-ad)i}{c^2+d^2}\\ &= \frac{ac+bd}{c^2+d^2}+\frac{bc-ad}{c^2+d^2}i.\end{aligned}$$

In other words, the quotient $\frac{z}{w}$ is obtained by multiplying both top and bottom of $\frac{z}{w}$ by $\overline{w}$ and then simplifying the expression.

Example 6.1.3: Division of Complex Numbers

$$\frac{1}{i} = \frac{1}{i}\times\frac{-i}{-i} = \frac{-i}{-i^2} = -i$$

$$\frac{2-i}{3+4i} = \frac{2-i}{3+4i}\times\frac{3-4i}{3-4i} = \frac{(6-4)+(-3-8)i}{3^2+4^2} = \frac{2-11i}{25} = \frac{2}{25}-\frac{11}{25}i$$

$$\frac{1-2i}{-2+5i} = \frac{1-2i}{-2+5i}\times\frac{-2-5i}{-2-5i} = \frac{(-2-10)+(4-5)i}{2^2+5^2} = -\frac{12}{29}-\frac{1}{29}i$$

Interestingly every nonzero complex number $a+bi$ has a unique multiplicative inverse. In other words, for a nonzero complex number z, there exists a number z^{-1} (or $\frac{1}{z}$) so that $zz^{-1} = 1$. Note that $z = a+bi$ is nonzero exactly when $a^2+b^2 \neq 0$, and its

inverse can be written in standard form as defined now.

Definition 6.1.2: Inverse of a Complex Number

Let $z = a + bi$ be a complex number. Then the multiplicative inverse of z, written z^{-1} exists if and only if $a^2 + b^2 \neq 0$ and is given by

$$z^{-1} = \frac{1}{a+bi} = \frac{1}{a+bi} \times \frac{a-bi}{a-bi} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} - i\frac{b}{a^2+b^2}$$

Note that we may write z^{-1} as $\frac{1}{z}$. Both notations represent the multiplicative inverse of the complex number z. Consider now an example.

Example 6.1.4: Inverse of a Complex Number

Consider the complex number $z = 2 + 6i$. Then z^{-1} is defined, and

$$\begin{aligned} \frac{1}{z} &= \frac{1}{2+6i} \\ &= \frac{1}{2+6i} \times \frac{2-6i}{2-6i} \\ &= \frac{2-6i}{2^2+6^2} \\ &= \frac{2-6i}{40} \\ &= \frac{1}{20} - \frac{3}{20}i \end{aligned}$$

You can always check your answer by computing zz^{-1}.

Another important construction of complex numbers is that of the absolute value, also called the modulus. Consider the following definition.

Definition 6.1.3: Absolute Value

The absolute value, or modulus, of a complex number, denoted $|z|$ is defined as follows.

$$|a+bi| = \sqrt{a^2+b^2}$$

Thus, if z is the complex number $z = a + bi$, it follows that

$$|z| = (z\bar{z})^{1/2}$$

Also from the definition, if $z = a + bi$ and $w = c + di$ are two complex numbers, then $|zw| = |z|\,|w|$. Take a moment to verify this.

The triangle inequality is an important property of the absolute value of complex numbers. There are two useful versions which we present here, although the first one is officially called the triangle inequality.

Proposition 6.1.1: Triangle Inequality

Let z, w be complex numbers.

The following two inequalities hold for any complex numbers z, w:

$$\begin{aligned} |z+w| &\leq |z| + |w| \\ ||z| - |w|| &\leq |z-w| \end{aligned}$$

The first one is called the *Triangle Inequality*.

Proof

Let $z = a + bi$ and $w = c + di$. First note that

$$z\overline{w} = (a + bi)(c - di) = ac + bd + (bc - ad)i$$

and so $|ac + bd| \leq |z\overline{w}| = |z||w|$.

Then,

$$|z + w|^2 = (a + c + i(b + d))(a + c - i(b + d))$$

$$= (a + c)^2 + (b + d)^2 = a^2 + c^2 + 2ac + 2bd + b^2 + d^2$$

$$\leq |z|^2 + |w|^2 + 2|z||w| = (|z| + |w|)^2$$

Taking the square root, we have that

$$|z + w| \leq |z| + |w|$$

so this verifies the triangle inequality.

To get the second inequality, write

$$z = z - w + w,\ w = w - z + z$$

and so by the first form of the inequality we get both:

$$|z| \leq |z - w| + |w|,\ |w| \leq |z - w| + |z|$$

Hence, both $|z| - |w|$ and $|w| - |z|$ are no larger than $|z - w|$. This proves the second version because $||z| - |w||$ is one of $|z| - |w|$ or $|w| - |z|$.

With this definition, it is important to note the following. You may wish to take the time to verify this remark.

Let $z = a + bi$ and $w = c + di$. Then

$$|z - w| = \sqrt{(a - c)^2 + (b - d)^2}.$$

Thus the distance between the point in the plane determined by the ordered pair (a, b) and the ordered pair (c, d) equals $|z - w|$ where z and w are as just described.

For example, consider the distance between $(2, 5)$ and $(1, 8)$. Letting $z = 2 + 5i$ and $w = 1 + 8i$, $z - w = 1 - 3i$, $(z - w)\left(\overline{z - w}\right) = (1 - 3i)(1 + 3i) = 10$ so $|z - w| = \sqrt{10}$.

Recall that we refer to $z = a + bi$ as the standard form of the complex number. In the next section, we examine another form in which we can express the complex number.

This page titled 6.1: Complex Numbers is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx).

- **6.1: Complex Numbers** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

6.2: Polar Form

Outcomes

A. Convert a complex number from standard form to polar form, and from polar form to standard form.

In the previous section, we identified a complex number $z = a + bi$ with a point (a, b) in the coordinate plane. There is another form in which we can express the same number, called the *polar form*. The polar form is the focus of this section. It will turn out to be very useful if not crucial for certain calculations as we shall soon see.

Suppose $z = a + bi$ is a complex number, and let $r = \sqrt{a^2 + b^2} = |z|$. Recall that r is the **modulus** of z. Note first that

$$\left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2 = \frac{a^2 + b^2}{r^2} = 1$$

and so $\left(\frac{a}{r}, \frac{b}{r}\right)$ is a point on the unit circle. Therefore, there exists an angle θ (in radians) such that

$$\cos\theta = \frac{a}{r}, \ \sin\theta = \frac{b}{r}$$

In other words θ is an angle such that $a = r\cos\theta$ and $b = r\sin\theta$, that is $\theta = \cos^{-1}(a/r)$ and $\theta = \sin^{-1}(b/r)$. We call this angle θ the **argument** of z.

We often speak of the **principal argument** of z. This is the unique angle $\theta \in (-\pi, \pi]$ such that

$$\cos\theta = \frac{a}{r}, \ \sin\theta = \frac{b}{r}$$

The polar form of the complex number $z = a + bi = r(\cos\theta + i\sin\theta)$ is for convenience written as:

$$z = re^{i\theta}$$

where θ is the argument of z.

Definition 6.2.1: Polar Form of a Complex Number

Let $z = a + bi$ be a complex number. Then the **polar form** of z is written as

$$z = re^{i\theta}$$

where $r = \sqrt{a^2 + b^2}$ and θ is the argument of z.

When given $z = re^{i\theta}$, the identity $e^{i\theta} = \cos\theta + i\sin\theta$ will convert z back to standard form. Here we think of $e^{i\theta}$ as a short cut for $\cos\theta + i\sin\theta$. This is all we will need in this course, but in reality $e^{i\theta}$ can be considered as the complex equivalent of the exponential function where this turns out to be a true equality.

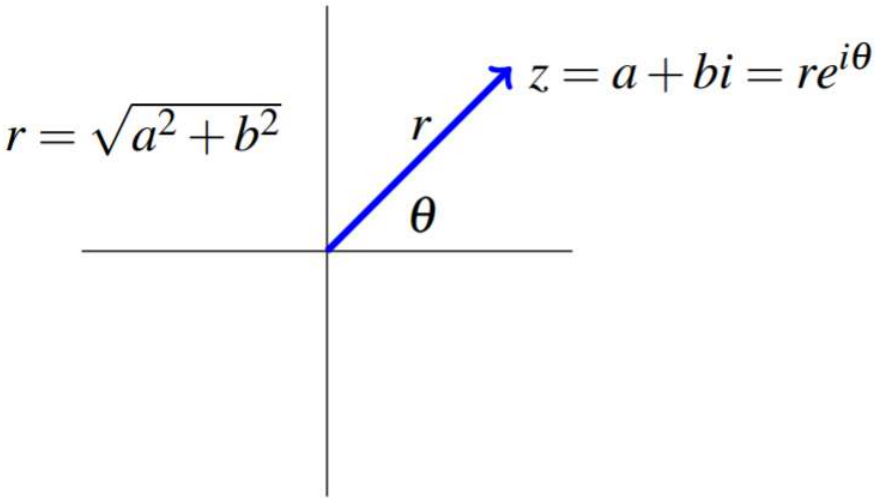


Figure 6.2.1

Thus we can convert any complex number in the standard (Cartesian) form $z = a + bi$ into its polar form. Consider the following example.

Example 6.2.1: Standard to Polar Form

Let $z = 2 + 2i$ be a complex number. Write z in the polar form

$$z = re^{i\theta}$$

Solution

First, find r. By the above discussion, $r = \sqrt{a^2 + b^2} = |z|$. Therefore,

$$r = \sqrt{2^2 + 2^2} = \sqrt{8} = 2\sqrt{2}$$

Now, to find θ, we plot the point $(2, 2)$ and find the angle from the positive x axis to the line between this point and the origin. In this case, $\theta = 45^\circ = \frac{\pi}{4}$. That is we found the unique angle θ such that $\theta = \cos^{-1}(1/\sqrt{2})$ and $\theta = \sin^{-1}(1/\sqrt{2})$.

Note that in polar form, we always express angles in radians, not degrees.

Hence, we can write z as

$$z = 2\sqrt{2}e^{i\frac{\pi}{4}}$$

Notice that the standard and polar forms are completely equivalent. That is not only can we transform a complex number from standard form to its polar form, we can also take a complex number in polar form and convert it back to standard form.

Example 6.2.3: Polar to Standard Form

Let $z = 2e^{2\pi i/3}$. Write z in the standard form

$$z = a + bi$$

Solution

Let $z = 2e^{2\pi i/3}$ be the polar form of a complex number. Recall that $e^{i\theta} = \cos\theta + i\sin\theta$. Therefore using standard values of sin and cos we get:

$$\begin{aligned} z = 2e^{i2\pi/3} &= 2(\cos(2\pi/3) + i\sin(2\pi/3)) \\ &= 2\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right) \\ &= -1 + \sqrt{3}i \end{aligned}$$

which is the standard form of this complex number.

You can always verify your answer by converting it back to polar form and ensuring you reach the original answer.

This page titled 6.2: Polar Form is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **6.2: Polar Form** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

6.3: Roots of Complex Numbers

Outcomes

A. Understand De Moivre's theorem and be able to use it to find the roots of a complex number.

A fundamental identity is the formula of De Moivre with which we begin this section.

Theorem 6.3.1: De Moivre's Theorem

For any positive integer n, we have

$$\left(e^{i\theta}\right)^n = e^{in\theta}$$

Thus for any real number $r > 0$ and any positive integer n, we have:

$$\left(r\left(\cos\theta + i\sin\theta\right)\right)^n = r^n\left(\cos n\theta + i\sin n\theta\right)$$

Proof

The proof is by induction on n. It is clear the formula holds if $n = 1$. Suppose it is true for n. Then, consider $n+1$.

$$\left(r\left(\cos\theta + i\sin\theta\right)\right)^{n+1} = \left(r\left(\cos\theta + i\sin\theta\right)\right)^n\left(r\left(\cos\theta + i\sin\theta\right)\right)$$

which by induction equals

$$\begin{aligned} &= r^{n+1}\left(\cos n\theta + i\sin n\theta\right)\left(\cos\theta + i\sin\theta\right) \\ &= r^{n+1}\left(\left(\cos n\theta\cos\theta - \sin n\theta\sin\theta\right) + i\left(\sin n\theta\cos\theta + \cos n\theta\sin\theta\right)\right) \\ &= r^{n+1}\left(\cos(n+1)\theta + i\sin(n+1)\theta\right) \end{aligned}$$

by the formulas for the cosine and sine of the sum of two angles.

The process used in the previous proof, called *mathematical induction* is very powerful in Mathematics and Computer Science and explored in more detail in the Appendix.

Now, consider a corollary of Theorem 6.3.1.

Corollary 6.3.1: Roots of Complex Numbers

Let z be a non zero complex number. Then there are always exactly k many k^{th} roots of z in $\mathbb{C}$.

Proof

Let $z = a + bi$ and let $z = |z|\left(\cos\theta + i\sin\theta\right)$ be the polar form of the complex number. By De Moivre's theorem, a complex number

$$w = re^{i\alpha} = r\left(\cos\alpha + i\sin\alpha\right)$$

is a kth root of z if and only if

$$w^k = \left(re^{i\alpha}\right)^k = r^k e^{ik\alpha} = r^k\left(\cos k\alpha + i\sin k\alpha\right) = |z|\left(\cos\theta + i\sin\theta\right)$$

This requires $r^k = |z|$ and so $r = |z|^{1/k}$. Also, both $\cos(k\alpha) = \cos\theta$ and $\sin(k\alpha) = \sin\theta$. This can only happen if

$$k\alpha = \theta + 2\ell\pi$$

for ℓ an integer. Thus

$$\alpha = \frac{\theta + 2\ell\pi}{k},\ \ell = 0, 1, 2, \cdots, k-1$$

and so the kth roots of z are of the form

$$|z|^{1/k}\left(\cos\left(\frac{\theta+2\ell\pi}{k}\right)+i\sin\left(\frac{\theta+2\ell\pi}{k}\right)\right),\ \ell=0,1,2,\cdots,k-1$$

Since the cosine and sine are periodic of period 2π, there are exactly k distinct numbers which result from this formula.

The procedure for finding the kth roots of $z\in\mathbb{C}$ is as follows.

Procedure 6.3.1: Finding Roots of a Complex Number

Let w be a complex number. We wish to find the nth roots of w, that is all z such that $z^n=w$.

There are n distinct n^{th} roots and they can be found as follows:.

1. Express both z and w in polar form $z=re^{i\theta}, w=se^{i\phi}$. Then $z^n=w$ becomes:
$$(re^{i\theta})^n=r^ne^{in\theta}=se^{i\phi}$$
We need to solve for r and θ.
2. Solve the following two equations:
$$\begin{aligned} r^n&=s\\ e^{in\theta}&=e^{i\phi}\end{aligned}\tag{6.3.1}$$
3. The solutions to $r^n=s$ are given by $r=\sqrt[n]{s}$.
4. The solutions to $e^{in\theta}=e^{i\phi}$ are given by:
$$n\theta=\phi+2\pi\ell,\ \text{for } \ell=0,1,2,\cdots,n-1$$
or
$$\theta=\frac{\phi}{n}+\frac{2}{n}\pi\ell,\ \text{for } \ell=0,1,2,\cdots,n-1$$
5. Using the solutions r,θ to the equations given in (6.3.1) construct the nth roots of the form $z=re^{i\theta}$.

Notice that once the roots are obtained in the final step, they can then be converted to standard form if necessary. Let's consider an example of this concept. Note that according to Corollary 6.3.1, there are exactly 3 cube roots of a complex number.

Example 6.3.1: Finding Cube Roots

Find the three cube roots of i. In other words find all z such that $z^3=i$.

Solution

First, convert each number to polar form: $z=re^{i\theta}$ and $i=1e^{i\pi/2}$. The equation now becomes

$$(re^{i\theta})^3=r^3e^{3i\theta}=1e^{i\pi/2}$$

Therefore, the two equations that we need to solve are $r^3=1$ and $3i\theta=i\pi/2$. Given that $r\in\mathbb{R}$ and $r^3=1$ it follows that $r=1$.

Solving the second equation is as follows. First divide by i. Then, since the argument of i is not unique we write $3\theta=\pi/2+2\pi\ell$ for $\ell=0,1,2$.

$$\begin{aligned} 3\theta&=\pi/2+2\pi\ell \text{ for } \ell=0,1,2\\ \theta&=\pi/6+\frac{2}{3}\pi\ell \text{ for } \ell=0,1,2\end{aligned}$$

For $\ell=0$:

$$\theta=\pi/6+\frac{2}{3}\pi(0)=\pi/6$$

For $\ell = 1$:

$$\theta = \pi/6 + \frac{2}{3}\pi(1) = \frac{5}{6}\pi$$

For $\ell = 2$:

$$\theta = \pi/6 + \frac{2}{3}\pi(2) = \frac{3}{2}\pi$$

Therefore, the three roots are given by

$$1e^{i\pi/6}, 1e^{i\frac{5}{6}\pi}, 1e^{i\frac{3}{2}\pi}$$

Written in standard form, these roots are, respectively,

$$\frac{\sqrt{3}}{2} + i\frac{1}{2}, -\frac{\sqrt{3}}{2} + i\frac{1}{2}, -i$$

The ability to find k^{th} roots can also be used to factor some polynomials.

Example 6.3.2: Solving a Polynomial Equation

Factor the polynomial $x^3 - 27$.

Solution

First find the cube roots of 27. By the above procedure , these cube roots are

$$3, 3\left(\frac{-1}{2} + i\frac{\sqrt{3}}{2}\right),$$

and

$$3\left(\frac{-1}{2} - i\frac{\sqrt{3}}{2}\right).$$

You may wish to verify this using the above steps.

Therefore,

$$x^3 - 27 = (x-3)\left(x - 3\left(\frac{-1}{2} + i\frac{\sqrt{3}}{2}\right)\right)\left(x - 3\left(\frac{-1}{2} - i\frac{\sqrt{3}}{2}\right)\right)$$

Note also

$$\left(x - 3\left(\frac{-1}{2} + i\frac{\sqrt{3}}{2}\right)\right)\left(x - 3\left(\frac{-1}{2} - i\frac{\sqrt{3}}{2}\right)\right) = x^2 + 3x + 9$$

and so

$$x^3 - 27 = (x-3)\left(x^2 + 3x + 9\right)$$

where the quadratic polynomial $x^2 + 3x + 9$ cannot be factored without using complex numbers.

Note that even though the polynomial $x^3 - 27$ has all real coefficients, it has some complex zeros, $3\left(\frac{-1}{2} + i\frac{\sqrt{3}}{2}\right)$, and $3\left(\frac{-1}{2} - i\frac{\sqrt{3}}{2}\right)$. These zeros are complex conjugates of each other. It is always the case that if a polynomial has real coefficients and a complex root, it will also have a root equal to the complex conjugate.

This page titled 6.3: Roots of Complex Numbers is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **6.3: Roots of Complex Numbers** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

6.4: The Quadratic Formula

Outcomes

A. Use the Quadratic Formula to find the complex roots of a quadratic equation.

The roots (or solutions) of a quadratic equation $ax^2+bx+c=0$ where a, b, c are real numbers are obtained by solving the familiar quadratic formula given by

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

When working with real numbers, we cannot solve this formula if $b^2-4ac<0$. However, complex numbers allow us to find square roots of negative numbers, and the quadratic formula remains valid for finding roots of the corresponding quadratic equation. In this case there are exactly two distinct (complex) square roots of b^2-4ac, which are $i\sqrt{4ac-b^2}$ and $-i\sqrt{4ac-b^2}$.

Here is an example.

Example 6.4.1: Solutions to Quadratic Equation

Find the solutions to $x^2+2x+5=0$.

Solution

In terms of the quadratic equation above, $a=1$, $b=2$, and $c=5$. Therefore, we can use the quadratic formula with these values, which becomes

$$x=\frac{-b\pm\sqrt{b^2-4ac}}{2a}=\frac{-2\pm\sqrt{(2)^2-4(1)(5)}}{2(1)}$$

Solving this equation, we see that the solutions are given by

$$x=\frac{-2i\pm\sqrt{4-20}}{2}=\frac{-2\pm4i}{2}=-1\pm2i$$

We can verify that these are solutions of the original equation. We will show $x=-1+2i$ and leave $x=-1-2i$ as an exercise.

$$\begin{aligned}x^2+2x+5 &=(-1+2i)^2+2(-1+2i)+5\\ &=1-4i-4-2+4i+5\\ &=0\end{aligned}$$

Hence $x=-1+2i$ is a solution.

What if the coefficients of the quadratic equation are actually complex numbers? Does the formula hold even in this case? The answer is yes. This is a hint on how to do Exercise 6.E.26 below, a special case of the fundamental theorem of algebra, and an ingredient in the proof of some versions of this theorem.

Consider the following example.

Example 6.4.2: Solutions to Quadratic Equation

Find the solutions to $x^2-2ix-5=0$.

Solution

In terms of the quadratic equation above, $a=1$, $b=-2i$, and $c=-5$. Therefore, we can use the quadratic formula with these values, which becomes

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{2i \pm \sqrt{(-2i)^2 - 4(1)(-5)}}{2(1)}$$

Solving this equation, we see that the solutions are given by

$$x = \frac{2i \pm \sqrt{-4 + 20}}{2} = \frac{2i \pm 4}{2} = i \pm 2$$

We can verify that these are solutions of the original equation. We will show $x = i + 2$ and leave $x = i - 2$ as an exercise.

$$\begin{aligned} x^2 - 2ix - 5 &= (i+2)^2 - 2i(i+2) - 5 \\ &= -1 + 4i + 4 + 2 - 4i - 5 \\ &= 0 \end{aligned}$$

Hence $x = i + 2$ is a solution.

We conclude this section by stating an essential theorem.

Theorem 6.4.1: The Fundamental Theorem of Algebra

Any polynomial of degree at least 1 with complex coefficients has a root which is a complex number.

This page titled 6.4: The Quadratic Formula is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **6.4: The Quadratic Formula** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

6.E: Exercises

Exercise 6.E.1

Let $z = 2 + 7i$ and let $w = 3 - 8i$. Compute the following.

a. $z + w$
b. $z - 2w$
c. zw
d. $\frac{w}{z}$

Answer

a. $z + w = 5 - i$
b. $z - 2w = -4 + 23i$
c. $zw = 62 + 5i$
d. $\frac{w}{z} = -\frac{50}{53} - \frac{37}{53}i$

Exercise 6.E.2

Let $z = 1 - 4i$. Compute the following.

a. $\overline{z}$
b. z^{-1}
c. $|z|$

Exercise 6.E.3

Let $z = 3 + 5i$ and $w = 2 - i$. Compute the following.

a. $\overline{zw}$
b. $|zw|$
c. $z^{-1}w$

Exercise 6.E.4

If z is a complex number, show there exists a complex number w with $|w| = 1$ and $wz = |z|$.

Answer

If $z = 0$, let $w = 1$. If $z \neq 0$, let $w = \frac{\overline{z}}{|z|}$

Exercise 6.E.5

If z, w are complex numbers prove $\overline{zw} = \overline{z}\,\overline{w}$ and then show by induction that $\overline{z_1 \cdots z_m} = \overline{z_1} \cdots \overline{z_m}$. Also verify that $\overline{\sum_{k=1}^{m} z_k} = \sum_{k=1}^{m} \overline{z_k}$. In words this says the conjugate of a product equals the product of the conjugates and the conjugate of a sum equals the sum of the conjugates.

Answer

$$\overline{(a+bi)(c+di)} = \overline{ac - bd + (ad+bc)i} = (ac-bd) - (ad+bc)i(a-bi)(c-di) = ac - bd - (ad+bc)i$$

which is the same thing. Thus it holds for a product of two complex numbers. Now suppose you have that it is true for the product of n complex numbers. Then

$$\overline{z_1 \cdots z_{n+1}} = \overline{z_1 \cdots z_n}\ \overline{z_{n+1}}$$

and now, by induction this equals

$$\overline{z_1} \cdots \overline{z_n}\ \overline{z_{n+1}}$$

As to sums, this is even easier.

$$\overline{\sum_{j=1}^{n}(x_j+iy_j)} = \overline{\sum_{j=1}^{n} x_j + i\sum_{j=1}^{n} y_j}$$

$$= \sum_{j=1}^{n} x_j - i\sum_{j=1}^{n} y_j = \sum_{j=1}^{n} x_j - iy_j = \sum_{j=1}^{n} \overline{(x_j+iy_j)}.$$

Exercise 6.E.6

Suppose $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ where all the a_k are real numbers. Suppose also that $p(z) = 0$ for some $z \in \mathbb{C}$. Show it follows that $p(\overline{z}) = 0$ also.

Answer

If $p(z) = 0$, then you have

$$\begin{aligned}\overline{p(z)} = 0 &= \overline{a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0} \\ &= \overline{a_n z^n} + \overline{a_{n-1}z^{n-1}} + \cdots + \overline{a_1 z} + \overline{a_0} \\ &= \overline{a_n}\,\overline{z}^n + \overline{a_{n-1}}\,\overline{z}^{n-1} + \cdots + \overline{a_1}\,\overline{z} + \overline{a_0} \\ &= a_n\overline{z}^n + a_{n-1}\overline{z}^{n-1} + \cdots + a_1\overline{z} + a_0 \\ &= p(\overline{z})\end{aligned}$$

Exercise 6.E.7

I claim that $1 = -1$. Here is why.

$$-1 = i^2 = \sqrt{-1}\sqrt{-1} = \sqrt{(-1)^2} = \sqrt{1} = 1$$

This is clearly a remarkable result but is there something wrong with it? If so, what is wrong?

Answer

The problem is that there is no single $\sqrt{-1}$.

Exercise 6.E.8

Let $z = 3 + 3i$ be a complex number written in standard form. Convert z to polar form, and write it in the form $z = re^{i\theta}$.

Exercise 6.E.9

Let $z = 2i$ be a complex number written in standard form. Convert z to polar form, and write it in the form $z = re^{i\theta}$.

Exercise 6.E.10

Let $z = 4e^{\frac{2\pi}{3}i}$ be a complex number written in polar form. Convert z to standard form, and write it in the form $z = a + bi$.

Exercise 6.E.11

Let $z = -1e^{\frac{\pi}{6}i}$ be a complex number written in polar form. Convert z to standard form, and write it in the form $z = a + bi$.

Exercise 6.E.12

If z and w are two complex numbers and the polar form of z involves the angle θ while the polar form of w involves the angle φ, show that in the polar form for zw the angle involved is $\theta + \varphi$.

Answer

You have $z = |z|(\cos\theta + i\sin\theta)$ and $w = |w|(\cos\varphi + i\sin\varphi)$. Then when you multiply these, you get

$$\begin{aligned} & |z|\,|w|(\cos\theta + i\sin\theta)(\cos\varphi + i\sin\varphi) \\ = & |z|\,|w|(\cos\theta\cos\varphi - \sin\theta\sin\varphi + i(\cos\theta\sin\varphi + \cos\varphi\sin\theta)) \\ = & |z|\,|w|(\cos(\theta+\varphi) + i\sin(\theta+\varphi)) \end{aligned}$$

Exercise 6.E.13

Give the complete solution to $x^4 + 16 = 0$.

Answer

Solution is:

$$(1-i)\sqrt{2},\ -(1+i)\sqrt{2},\ -(1-i)\sqrt{2},\ (1+i)\sqrt{2}$$

Exercise 6.E.14

Find the complex cube roots of 8.

Answer

The cube roots are the solutions to $z^3 - 8 = 0$, Solution is: $-1 + i\sqrt{3},\ -1 + i\sqrt{3},\ 2$

Exercise 6.E.15

Find the four fourth roots of -16.

Answer

The fourth roots are the solutions to $z^4 + 16 = 0$, Solution is:

$$(1-i)\sqrt{2},\ -(1+i)\sqrt{2},\ -(1-i)\sqrt{2},\ (1+i)\sqrt{2}$$

Exercise 6.E.16

De Moivre's theorem says $[r(\cos t + i\sin t)]^n = r^n(\cos nt + i\sin nt)$ for n a positive integer. Does this formula continue to hold for all integers n, even negative integers? Explain.

Answer

Yes, it holds for all integers. First of all, it clearly holds if $n = 0$. Suppose now that n is a negative integer. Then $-n > 0$ and so

$$[r(\cos t + i\sin t)]^n = \frac{1}{[r(\cos t + i\sin t)]^{-n}} = \frac{1}{r^{-n}(\cos(-nt) + i\sin(-nt))}$$

$$= \frac{r^n}{(\cos(nt) - i\sin(nt))} = \frac{r^n(\cos(nt) + i\sin(nt))}{(\cos(nt) - i\sin(nt))(\cos(nt) + i\sin(nt))}$$
$$= r^n(\cos(nt) + i\sin(nt))$$

because $(\cos(nt) - i\sin(nt))(\cos(nt) + i\sin(nt)) = 1$.

Exercise 6.E.17

Factor $x^3 + 8$ as a product of linear factors. Hint: Use the result of 6.E.14.

Answer

Solution is: $i\sqrt{3}+1,\ 1-i\sqrt{3},\ -2$ and so this polynomial equals

$$(x+2)\left(x-(i\sqrt{3}+1)\right)\left(x-(1-i\sqrt{3})\right)$$

Exercise 6.E.18

Write $x^3 + 27$ in the form $(x+3)(x^2+ax+b)$ where x^2+ax+b cannot be factored any more using only real numbers.

Answer

$x^3 + 27 = (x+3)(x^2-3x+9)$

Exercise 6.E.19

Completely factor $x^4 + 16$ as a product of linear factors. Hint: Use the result of 6.E.15.

Answer

Solution is:

$$(1-i)\sqrt{2},\ -(1+i)\sqrt{2},\ -(1-i)\sqrt{2},\ (1+i)\sqrt{2}.$$

These are just the fourth roots of -16. Then to factor, you get

$$\left(x-\left((1-i)\sqrt{2}\right)\right)\left(x-\left(-(1+i)\sqrt{2}\right)\right).$$
$$\left(x-\left(-(1-i)\sqrt{2}\right)\right)\left(x-\left((1+i)\sqrt{2}\right)\right)$$

Exercise 6.E.20

Factor $x^4 + 16$ as the product of two quadratic polynomials each of which cannot be factored further without using complex numbers.

Answer

$x^4 + 16 = \left(x^2 - 2\sqrt{2}x + 4\right)\left(x^2 + 2\sqrt{2}x + 4\right)$. You can use the information in the preceding problem. Note that $(x-z)(x-\bar{z})$ has real coefficients.

Exercise 6.E.21

If n is an integer, is it always true that $(\cos\theta - i\sin\theta)^n = \cos(n\theta) - i\sin(n\theta)$? Explain.

Answer

Yes, this is true.

$$\begin{aligned}(\cos\theta - i\sin\theta)^n &= (\cos(-\theta) + i\sin(-\theta))^n \\ &= \cos(-n\theta) + i\sin(-n\theta) \\ &= \cos(n\theta) - i\sin(n\theta)\end{aligned}$$

Exercise 6.E. 22

Suppose $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ is a polynomial and it has n zeros,

$$z_1,\ z_2, \cdots, z_n$$

listed according to multiplicity. (z is a root of multiplicity m if the polynomial $f(x) = (x-z)^m$ divides $p(x)$ but $(x-z)f(x)$ does not.) Show that

$$p(x) = a_n(x - z_1)(x - z_2)\cdots(x - z_n)$$

Answer

$p(x) = (x - z_1)q(x) + r(x)$ where $r(x)$ is a nonzero constant or equal to 0. However, $r(z_1) = 0$ and so $r(x) = 0$. Now do to $q(x)$ what was done to $p(x)$ and continue until the degree of the resulting $q(x)$ equals 0. Then you have the above factorization.

Exercise 6.E. 23

Show that $1 + i,\ 2 + i$ are the only two roots to

$$p(x) = x^2 - (3 + 2i)x + (1 + 3i)$$

Hence complex zeros do not necessarily come in conjugate pairs if the coefficients of the equation are not real.

Answer

$$(x - (1 + i))(x - (2 + i)) = x^2 - (3 + 2i)x + 1 + 3i$$

Exercise 6.E. 24

Give the solutions to the following quadratic equations having real coefficients.

a. $x^2 - 2x + 2 = 0$
b. $3x^2 + x + 3 = 0$
c. $x^2 - 6x + 13 = 0$
d. $x^2 + 4x + 9 = 0$
e. $4x^2 + 4x + 5 = 0$

Answer

a. Solution is: $1 + i,\ 1 - i$
b. Solution is: $\frac{1}{6}i\sqrt{35} - \frac{1}{6},\ -\frac{1}{6}i\sqrt{35} - \frac{1}{6}$
c. Solution is: $3 + 2i,\ 3 - 2i$
d. Solution is: $i\sqrt{5} - 2,\ -i\sqrt{5} - 2$
e. Solution is: $-\frac{1}{2} + i,\ -\frac{1}{2} - i$

Exercise 6.E. 25

Give the solutions to the following quadratic equations having complex coefficients.

a. $x^2 + 2x + 1 + i = 0$
b. $4x^2 + 4ix - 5 = 0$

c. $4x^2+(4+4i)x+1+2i=0$
d. $x^2-4ix-5=0$
e. $3x^2+(1-i)x+3i=0$

Answer

a. Solution is: $x=-1+\frac{1}{2}\sqrt{2}-\frac{1}{2}i\sqrt{2},\ x=-1-\frac{1}{2}\sqrt{2}+\frac{1}{2}i\sqrt{2}$
b. Solution is: $x=1-\frac{1}{2}i,\ x=-1-\frac{1}{2}i$
c. Solution is: $x=-\frac{1}{2},\ x=-\frac{1}{2}-i$
d. Solution is: $x=-1+2i,\ x=1+2i$
e. Solution is: $x=-\frac{1}{6}+\frac{1}{6}\sqrt{19}+\left(\frac{1}{6}-\frac{1}{6}\sqrt{19}\right)i,\ x=-\frac{1}{6}-\frac{1}{6}\sqrt{19}+\left(\frac{1}{6}+\frac{1}{6}\sqrt{19}\right)i$

? Exercise $6.E.26$

Prove the fundamental theorem of algebra for quadratic polynomials having coefficients in $\mathbb{C}$. That is, show that an equation of the form $ax^2+bx+c=0$ where $a,\ b,\ c$ are complex numbers, $a\neq 0$ has a complex solution. Hint: Consider the fact, noted earlier that the expressions given from the quadratic formula do in fact serve as solutions.

This page titled 6.E: Exercises is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

CHAPTER OVERVIEW

7: Spectral Theory

7.1: Eigenvalues and Eigenvectors of a Matrix
7.2: Diagonalization
7.3: Applications of Spectral Theory
7.4: Orthogonality
7.E: Exercises

This page titled 7: Spectral Theory is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

7.1: Eigenvalues and Eigenvectors of a Matrix

Outcomes

A. Describe eigenvalues geometrically and algebraically.
B. Find eigenvalues and eigenvectors for a square matrix.

Spectral Theory refers to the study of eigenvalues and eigenvectors of a matrix. It is of fundamental importance in many areas and is the subject of our study for this chapter.

Definition of Eigenvectors and Eigenvalues

In this section, we will work with the entire set of complex numbers, denoted by $\mathbb{C}$. Recall that the real numbers, $\mathbb{R}$ are contained in the complex numbers, so the discussions in this section apply to both real and complex numbers.

To illustrate the idea behind what will be discussed, consider the following example.

Example 7.1.1:Eigenvectors and Eigenvalues

Let

$$A = \begin{bmatrix} 0 & 5 & -10 \\ 0 & 22 & 16 \\ 0 & -9 & -2 \end{bmatrix}$$

Compute the product AX for

$$X = \begin{bmatrix} -5 \\ -4 \\ 3 \end{bmatrix}, X = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

What do you notice about AX in each of these products?

Solution

First, compute AX for

$$X = \begin{bmatrix} -5 \\ -4 \\ 3 \end{bmatrix}$$

This product is given by

$$AX = \begin{bmatrix} 0 & 5 & -10 \\ 0 & 22 & 16 \\ 0 & -9 & -2 \end{bmatrix} \begin{bmatrix} -5 \\ -4 \\ 3 \end{bmatrix} = \begin{bmatrix} -50 \\ -40 \\ 30 \end{bmatrix} = 10 \begin{bmatrix} -5 \\ -4 \\ 3 \end{bmatrix}$$

In this case, the product AX resulted in a vector which is equal to 10 times the vector X. In other words, $AX = 10X$.

Let's see what happens in the next product. Compute AX for the vector

$$X = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

This product is given by

$$AX = \begin{bmatrix} 0 & 5 & -10 \\ 0 & 22 & 16 \\ 0 & -9 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

In this case, the product AX resulted in a vector equal to 0 times the vector X, $AX = 0X$.

Perhaps this matrix is such that AX results in kX, for every vector X. However, consider

$$\begin{bmatrix} 0 & 5 & -10 \\ 0 & 22 & 16 \\ 0 & -9 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ 38 \\ -11 \end{bmatrix}$$

In this case, AX did not result in a vector of the form kX for some scalar k.

There is something special about the first two products calculated in Example 7.1.1. Notice that for each, $AX = kX$ where k is some scalar. When this equation holds for some X and k, we call the scalar k an **eigenvalue** of A. We often use the special symbol λ instead of k when referring to eigenvalues. In Example 7.1.1, the values 10 and 0 are eigenvalues for the matrix A and we can label these as $\lambda_1 = 10$ and $\lambda_2 = 0$.

When $AX = \lambda X$ for some $X \neq 0$, we call such an X an **eigenvector** of the matrix A. The eigenvectors of A are associated to an eigenvalue. Hence, if λ_1 is an eigenvalue of A and $AX = \lambda_1 X$, we can label this eigenvector as X_1. Note again that in order to be an eigenvector, X must be nonzero.

There is also a geometric significance to eigenvectors. When you have a **nonzero** vector which, when multiplied by a matrix results in another vector which is parallel to the first or equal to **0**, this vector is called an eigenvector of the matrix. This is the meaning when the vectors are in $\mathbb{R}^n$.

The formal definition of eigenvalues and eigenvectors is as follows.

Definition 7.1.1:Eigenvalues and Eigenvectors

Let A be an $n \times n$ matrix and let $X \in \mathbb{C}^n$ be a **nonzero vector** for which

$$AX = \lambda X \tag{7.1.1}$$

for some scalar λ. Then λ is called an **eigenvalue** of the matrix A and X is called an **eigenvector** of A associated with λ, or a λ-eigenvector of A.

The set of all eigenvalues of an $n \times n$ matrix A is denoted by $\sigma(A)$ and is referred to as the **spectrum** of A.

The eigenvectors of a matrix A are those vectors X for which multiplication by A results in a vector in the same direction or opposite direction to X. Since the zero vector 0 has no direction this would make no sense for the zero vector. As noted above, 0 is never allowed to be an eigenvector.

Let's look at eigenvectors in more detail. Suppose X satisfies (7.1.1). Then

$$AX - \lambda X = 0$$
$$\text{or}$$
$$(A - \lambda I) X = 0$$

for some $X \neq 0$. Equivalently you could write $(\lambda I - A) X = 0$, which is more commonly used. Hence, when we are looking for eigenvectors, we are looking for nontrivial solutions to this homogeneous system of equations!

Recall that the solutions to a homogeneous system of equations consist of basic solutions, and the linear combinations of those basic solutions. In this context, we call the basic solutions of the equation $(\lambda I - A) X = 0$ **basic eigenvectors**. It follows that any (nonzero) linear combination of basic eigenvectors is again an eigenvector.

Suppose the matrix $(\lambda I - A)$ is invertible, so that $(\lambda I - A)^{-1}$ exists. Then the following equation would be true.

$$\begin{aligned} X &= IX \\ &= \left((\lambda I - A)^{-1} (\lambda I - A) \right) X \\ &= (\lambda I - A)^{-1} ((\lambda I - A) X) \\ &= (\lambda I - A)^{-1} 0 \\ &= 0 \end{aligned}$$

This claims that $X = 0$. However, we have required that $X \neq 0$. Therefore $(\lambda I - A)$ cannot have an inverse!

Recall that if a matrix is not invertible, then its determinant is equal to 0. Therefore we can conclude that

$$\det(\lambda I - A) = 0 \tag{7.1.2}$$

Note that this is equivalent to $\det(A - \lambda I) = 0$.

The expression $\det(\lambda I - A)$ is a polynomial (in the variable x) called the **characteristic polynomial of A**, and $\det(\lambda I - A) = 0$ is called the **characteristic equation**. For this reason we may also refer to the eigenvalues of A as **characteristic values**, but the former is often used for historical reasons.

Below is a video on eigenvalues and eigenvectors.



The following theorem claims that the roots of the characteristic polynomial are the eigenvalues of A. Thus when [eigen2] holds, A has a nonzero eigenvector.

Theorem 7.1.1: The Existence of an Eigenvector

Let A be an $n \times n$ matrix and suppose $\det(\lambda I - A) = 0$ for some $\lambda \in \mathbb{C}$.
Then λ is an eigenvalue of A and thus there exists a nonzero vector $X \in \mathbb{C}^n$ such that $AX = \lambda X$.

Proof

For A an $n \times n$ matrix, the method of Laplace Expansion demonstrates that $\det(\lambda I - A)$ is a polynomial of degree n. As such, the equation (7.1.2) has a solution $\lambda \in \mathbb{C}$ by the Fundamental Theorem of Algebra. The fact that λ is an eigenvalue is left as an exercise.

Finding Eigenvectors and Eigenvalues

Now that eigenvalues and eigenvectors have been defined, we will study how to find them for a matrix A.

First, consider the following definition.

Definition 7.1.2:Multiplicity of an Eigenvalue

Let A be an $n \times n$ matrix with characteristic polynomial given by $\det(\lambda I - A)$. Then, the multiplicity of an eigenvalue λ of A is the number of times λ occurs as a root of that characteristic polynomial.

For example, suppose the characteristic polynomial of A is given by $(\lambda - 2)^2$. Solving for the roots of this polynomial, we set $(\lambda - 2)^2 = 0$ and solve for λ. We find that $\lambda = 2$ is a root that occurs twice. Hence, in this case, $\lambda = 2$ is an eigenvalue of A of multiplicity equal to 2.

Below is a video on finding λ to make $A - \lambda I$ singular.



Below is a video on finding an eigenvalue from Ax=b.



We will now look at how to find the eigenvalues and eigenvectors for a matrix A in detail. The steps used are summarized in the following procedure.

Procedure 7.1.1:Finding Eigenvalues and Eigenvectors

Let A be an $n \times n$ matrix.

1. First, find the eigenvalues λ of A by solving the equation $\det(\lambda I - A) = 0$.
2. For each λ, find the basic eigenvectors $X \neq 0$ by finding the basic solutions to $(\lambda I - A) X = 0$.

To verify your work, make sure that $AX = \lambda X$ for each λ and associated eigenvector X.

We will explore these steps further in the following example.

Example 7.1.2:Find the Eigenvalues and Eigenvectors

Let $A = \begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix}$. Find its eigenvalues and eigenvectors.

Solution

We will use Procedure 7.1.1. First we find the eigenvalues of A by solving the equation

$$\det(\lambda I - A) = 0$$

This gives

$$\det\left(\lambda\begin{bmatrix}1&0\\0&1\end{bmatrix}-\begin{bmatrix}-5&2\\-7&4\end{bmatrix}\right)=0$$

$$\det\begin{bmatrix}\lambda+5&-2\\7&\lambda-4\end{bmatrix}=0$$

Computing the determinant as usual, the result is

$$\lambda^2+\lambda-6=0$$

Solving this equation, we find that $\lambda_1=2$ and $\lambda_2=-3$.

Now we need to find the basic eigenvectors for each λ. First we will find the eigenvectors for $\lambda_1=2$. We wish to find all vectors $X\neq 0$ such that $AX=2X$. These are the solutions to $(2I-A)X=0$.

$$\left(2\begin{bmatrix}1&0\\0&1\end{bmatrix}-\begin{bmatrix}-5&2\\-7&4\end{bmatrix}\right)\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

$$\begin{bmatrix}7&-2\\7&-2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

The augmented matrix for this system and corresponding reduced row-echelon form are given by

$$\left[\begin{array}{cc|c}7&-2&0\\7&-2&0\end{array}\right]\to\cdots\to\left[\begin{array}{cc|c}1&-\frac{2}{7}&0\\0&0&0\end{array}\right]$$

The solution is any vector of the form

$$\begin{bmatrix}\frac{2}{7}s\\s\end{bmatrix}=s\begin{bmatrix}\frac{2}{7}\\1\end{bmatrix}$$

Multiplying this vector by 7 we obtain a simpler description for the solution to this system, given by

$$t\begin{bmatrix}2\\7\end{bmatrix}$$

This gives the basic eigenvector for $\lambda_1=2$ as

$$\begin{bmatrix}2\\7\end{bmatrix}$$

To check, we verify that $AX=2X$ for this basic eigenvector.

$$\begin{bmatrix}-5&2\\-7&4\end{bmatrix}\begin{bmatrix}2\\7\end{bmatrix}=\begin{bmatrix}4\\14\end{bmatrix}=2\begin{bmatrix}2\\7\end{bmatrix}$$

This is what we wanted, so we know this basic eigenvector is correct.

Next we will repeat this process to find the basic eigenvector for $\lambda_2=-3$. We wish to find all vectors $X\neq 0$ such that $AX=-3X$. These are the solutions to $((-3)I-A)X=0$.

$$\left((-3)\begin{bmatrix}1&0\\0&1\end{bmatrix}-\begin{bmatrix}-5&2\\-7&4\end{bmatrix}\right)\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

$$\begin{bmatrix}2&-2\\7&-7\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

The augmented matrix for this system and corresponding reduced row-echelon form are given by

$$\left[\begin{array}{cc|c}2&-2&0\\7&-7&0\end{array}\right]\to\cdots\to\left[\begin{array}{cc|c}1&-1&0\\0&0&0\end{array}\right]$$

The solution is any vector of the form

$$\begin{bmatrix} s \\ s \end{bmatrix} = s \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

This gives the basic eigenvector for $\lambda_2 = -3$ as

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

To check, we verify that $AX = -3X$ for this basic eigenvector.

$$\begin{bmatrix} -5 & 2 \\ -7 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \end{bmatrix} = -3 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

This is what we wanted, so we know this basic eigenvector is correct.

Below is a video on finding eigenvalues of a matrix.



Below is a video on finding eigenvectors given eigenvalues.



Below is a video on finding an eigenvalue given matrix A and an eigenvector.



The following is an example using Procedure 7.1.1 for a 3×3 matrix.

Example 7.1.3:Find the Eigenvalues and Eigencectors

Find the eigenvalues and eigenvectors for the matrix

$$A = \begin{bmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{bmatrix}$$

Solution

We will use Procedure 7.1.1. First we need to find the eigenvalues of A. Recall that they are the solutions of the equation

$$\det(\lambda I - A) = 0$$

In this case the equation is

$$\det\left(\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{bmatrix}\right) = 0$$

which becomes

$$\det\begin{bmatrix} \lambda - 5 & 10 & 5 \\ -2 & \lambda - 14 & -2 \\ 4 & 8 & \lambda - 6 \end{bmatrix} = 0$$

Using Laplace Expansion, compute this determinant and simplify. The result is the following equation.

$$(\lambda - 5)\left(\lambda^2 - 20\lambda + 100\right) = 0$$

Solving this equation, we find that the eigenvalues are $\lambda_1 = 5, \lambda_2 = 10$ and $\lambda_3 = 10$. Notice that 10 is a root of multiplicity two due to

$$\lambda^2 - 20\lambda + 100 = (\lambda - 10)^2$$

Therefore, $\lambda_2 = 10$ is an eigenvalue of multiplicity two.

Now that we have found the eigenvalues for A, we can compute the eigenvectors.

First we will find the basic eigenvectors for $\lambda_1 = 5$. In other words, we want to find all non-zero vectors X so that $AX = 5X$. This requires that we solve the equation $(5I - A)X = 0$ for X as follows.

$$\left(5 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{bmatrix}\right) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

That is you need to find the solution to

$$\begin{bmatrix} 0 & 10 & 5 \\ -2 & -9 & -2 \\ 4 & 8 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

By now this is a familiar problem. You set up the augmented matrix and row reduce to get the solution. Thus the matrix you must row reduce is

$$\left[\begin{array}{ccc|c} 0 & 10 & 5 & 0 \\ -2 & -9 & -2 & 0 \\ 4 & 8 & -1 & 0 \end{array}\right]$$

The reduced row-echelon form is

$$\left[\begin{array}{ccc|c} 1 & 0 & -\frac{5}{4} & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and so the solution is any vector of the form

$$\begin{bmatrix} \frac{5}{4}s \\ -\frac{1}{2}s \\ s \end{bmatrix} = s \begin{bmatrix} \frac{5}{4} \\ -\frac{1}{2} \\ 1 \end{bmatrix}$$

where $s \in \mathbb{R}$. If we multiply this vector by 4, we obtain a simpler description for the solution to this system, as given by

$$t \begin{bmatrix} 5 \\ -2 \\ 4 \end{bmatrix} \tag{7.1.3}$$

where $t \in \mathbb{R}$. Here, the basic eigenvector is given by

$$X_1 = \begin{bmatrix} 5 \\ -2 \\ 4 \end{bmatrix}$$

Notice that we cannot let $t = 0$ here, because this would result in the zero vector and eigenvectors are never equal to 0! Other than this value, every other choice of t in (7.1.3) results in an eigenvector.

It is a good idea to check your work! To do so, we will take the original matrix and multiply by the basic eigenvector X_1. We check to see if we get $5X_1$.

$$\begin{bmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{bmatrix} \begin{bmatrix} 5 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 25 \\ -10 \\ 20 \end{bmatrix} = 5 \begin{bmatrix} 5 \\ -2 \\ 4 \end{bmatrix}$$

This is what we wanted, so we know that our calculations were correct.

Next we will find the basic eigenvectors for $\lambda_2, \lambda_3 = 10$. These vectors are the basic solutions to the equation,

$$\left(10 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{bmatrix} \right) \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

That is you must find the solutions to

$$\begin{bmatrix} 5 & 10 & 5 \\ -2 & -4 & -2 \\ 4 & 8 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Consider the augmented matrix

$$\left[\begin{array}{rrr|r} 5 & 10 & 5 & 0 \\ -2 & -4 & -2 & 0 \\ 4 & 8 & 4 & 0 \end{array}\right]$$

The reduced row-echelon form for this matrix is

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and so the eigenvectors are of the form

$$\begin{bmatrix} -2s-t \\ s \\ t \end{bmatrix} = s \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Note that you can't pick t and s both equal to zero because this would result in the zero vector and eigenvectors are never equal to zero.

Here, there are two basic eigenvectors, given by

$$X_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, X_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Taking any (nonzero) linear combination of X_2 and X_3 will also result in an eigenvector for the eigenvalue $\lambda = 10$. As in the case for $\lambda = 5$, always check your work! For the first basic eigenvector, we can check $AX_2 = 10X_2$ as follows.

$$\begin{bmatrix} 5 & -10 & -5 \\ 2 & 14 & 2 \\ -4 & -8 & 6 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -10 \\ 0 \\ 10 \end{bmatrix} = 10 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

This is what we wanted. Checking the second basic eigenvector, X_3, is left as an exercise.

Below is a video on finding the eigenvalues of a 3x3 matrix.



Below is a video on finding the eigenvalues of a 4x4 matrix.



It is important to remember that for any eigenvector X, $X \neq 0$. However, it is possible to have eigenvalues equal to zero. This is illustrated in the following example.

Example 7.1.4:A Zero Eigenvalue

Let

$$A = \begin{bmatrix} 2 & 2 & -2 \\ 1 & 3 & -1 \\ -1 & 1 & 1 \end{bmatrix}$$

Find the eigenvalues and eigenvectors of A.

Solution

First we find the eigenvalues of A. We will do so using Definition 7.1.1.

In order to find the eigenvalues of A, we solve the following equation.

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & -2 & 2 \\ -1 & \lambda - 3 & 1 \\ 1 & -1 & \lambda - 1 \end{bmatrix} = 0$$

This reduces to $\lambda^3 - 6\lambda^2 + 8\lambda = 0$. You can verify that the solutions are $\lambda_1 = 0, \lambda_2 = 2, \lambda_3 = 4$. Notice that while eigenvectors can never equal 0, it is possible to have an eigenvalue equal to 0.

Now we will find the basic eigenvectors. For $\lambda_1 = 0$, we need to solve the equation $(0I - A)X = 0$. This equation becomes $-AX = 0$, and so the augmented matrix for finding the solutions is given by

$$\left[\begin{array}{ccc|c} -2 & -2 & 2 & 0 \\ -1 & -3 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{array}\right]$$

The reduced row-echelon form is

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore, the eigenvectors are of the form $t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ where $t \neq 0$ and the basic eigenvector is given by

$$X_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

We can verify that this eigenvector is correct by checking that the equation $AX_1 = 0X_1$ holds. The product AX_1 is given by

$$AX_1 = \begin{bmatrix} 2 & 2 & -2 \\ 1 & 3 & -1 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This clearly equals $0X_1$, so the equation holds. Hence, $AX_1 = 0X_1$ and so 0 is an eigenvalue of A.

Computing the other basic eigenvectors is left as an exercise.

In the following sections, we examine ways to simplify this process of finding eigenvalues and eigenvectors by using properties of special types of matrices.

Eigenvalues and Eigenvectors for Special Types of Matrices

There are three special kinds of matrices which we can use to simplify the process of finding eigenvalues and eigenvectors. Throughout this section, we will discuss similar matrices, elementary matrices, as well as triangular matrices.

We begin with a definition.

Definition 7.1.3:Similar Matrices

Let A and B be $n \times n$ matrices. Suppose there exists an invertible matrix P such that

$$A = P^{-1}BP$$

Then A and B are called **similar matrices**.

It turns out that we can use the concept of similar matrices to help us find the eigenvalues of matrices. Consider the following lemma.

Lemma 7.1.1:Similar Matrices and Eigenvalues

Let A and B be similar matrices, so that $A = P^{-1}BP$ where A, B are $n \times n$ matrices and P is invertible. Then A, B have the same eigenvalues.

Proof

We need to show two things. First, we need to show that if $A = P^{-1}BP$, then A and B have the same eigenvalues. Secondly, we show that if A and B have the same eigenvalues, then $A = P^{-1}BP$.

Here is the proof of the first statement. Suppose $A = P^{-1}BP$ and λ is an eigenvalue of A, that is $AX = \lambda X$ for some $X \neq 0$. Then

$$P^{-1}BPX = \lambda X$$

and so

$$BPX = \lambda PX$$

Since P is one to one and $X \neq 0$, it follows that $PX \neq 0$. Here, PX plays the role of the eigenvector in this equation. Thus λ is also an eigenvalue of B. One can similarly verify that any eigenvalue of B is also an eigenvalue of A, and thus both matrices have the same eigenvalues as desired.

Proving the second statement is similar and is left as an exercise.

Note that this proof also demonstrates that the eigenvectors of A and B will (generally) be *different*. We see in the proof that $AX = \lambda X$, while $B(PX) = \lambda(PX)$. Therefore, for an eigenvalue λ, A will have the eigenvector X while B will have the eigenvector PX.

The second special type of matrices we discuss in this section is elementary matrices. Recall from Definition 2.8.1 that an elementary matrix E is obtained by applying one row operation to the identity matrix.

It is possible to use elementary matrices to simplify a matrix before searching for its eigenvalues and eigenvectors. This is illustrated in the following example.

Example 7.1.5: Simplify Using Elementary Matrices

Find the eigenvalues for the matrix

$$A = \begin{bmatrix} 33 & 105 & 105 \\ 10 & 28 & 30 \\ -20 & -60 & -62 \end{bmatrix}$$

Solution

This matrix has big numbers and therefore we would like to simplify as much as possible before computing the eigenvalues.

We will do so using row operations. First, add 2 times the second row to the third row. To do so, left multiply A by $E(2,2)$. Then right multiply A by the inverse of $E(2,2)$ as illustrated.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 33 & 105 & 105 \\ 10 & 28 & 30 \\ -20 & -60 & -62 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 33 & -105 & 105 \\ 10 & -32 & 30 \\ 0 & 0 & -2 \end{bmatrix}$$

By Lemma 7.1.1, the resulting matrix has the same eigenvalues as A where here, the matrix $E(2,2)$ plays the role of P.

We do this step again, as follows. In this step, we use the elementary matrix obtained by adding -3 times the second row to the first row.

$$\begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 33 & -105 & 105 \\ 10 & -32 & 30 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 15 \\ 10 & -2 & 30 \\ 0 & 0 & -2 \end{bmatrix} \tag{7.1.4}$$

Again by Lemma 7.1.1, this resulting matrix has the same eigenvalues as A. At this point, we can easily find the eigenvalues. Let

$$B = \begin{bmatrix} 3 & 0 & 15 \\ 10 & -2 & 30 \\ 0 & 0 & -2 \end{bmatrix}$$

Then, we find the eigenvalues of B (and therefore of A) by solving the equation $\det(\lambda I - B) = 0$. You should verify that this equation becomes

$$(\lambda + 2)(\lambda + 2)(\lambda - 3) = 0$$

Solving this equation results in eigenvalues of $\lambda_1 = -2, \lambda_2 = -2$, and $\lambda_3 = 3$. Therefore, these are also the eigenvalues of A.

Through using elementary matrices, we were able to create a matrix for which finding the eigenvalues was easier than for A. At this point, you could go back to the original matrix A and solve $(\lambda I - A)X = 0$ to obtain the eigenvectors of A.

Notice that when you multiply on the right by an elementary matrix, you are doing the column operation defined by the elementary matrix. In (7.1.4) multiplication by the elementary matrix on the right merely involves taking three times the first column and adding to the second. Thus, without referring to the elementary matrices, the transition to the new matrix in (7.1.4) can be illustrated by

$$\begin{bmatrix} 33 & -105 & 105 \\ 10 & -32 & 30 \\ 0 & 0 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & -9 & 15 \\ 10 & -32 & 30 \\ 0 & 0 & -2 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 0 & 15 \\ 10 & -2 & 30 \\ 0 & 0 & -2 \end{bmatrix}$$

The third special type of matrix we will consider in this section is the triangular matrix. Recall Definition 3.1.6 which states that an upper (lower) triangular matrix contains all zeros below (above) the main diagonal. Remember that finding the determinant of a triangular matrix is a simple procedure of taking the product of the entries on the main diagonal.. It turns out that there is also a simple way to find the eigenvalues of a triangular matrix.

In the next example we will demonstrate that the eigenvalues of a triangular matrix are the entries on the main diagonal.

Example 7.1.6:Eigenvalues for a Triangular Matrix

Let $A = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 4 & 7 \\ 0 & 0 & 6 \end{bmatrix}$. Find the eigenvalues of A.

Solution

We need to solve the equation $\det(\lambda I - A) = 0$ as follows

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 1 & -2 & -4 \\ 0 & \lambda - 4 & -7 \\ 0 & 0 & \lambda - 6 \end{bmatrix} = (\lambda - 1)(\lambda - 4)(\lambda - 6) = 0$$

Solving the equation $(\lambda - 1)(\lambda - 4)(\lambda - 6) = 0$ for λ results in the eigenvalues $\lambda_1 = 1, \lambda_2 = 4$ and $\lambda_3 = 6$. Thus the eigenvalues are the entries on the main diagonal of the original matrix.

The same result is true for lower triangular matrices. For any triangular matrix, the eigenvalues are equal to the entries on the main diagonal. To find the eigenvectors of a triangular matrix, we use the usual procedure.

In the next section, we explore an important process involving the eigenvalues and eigenvectors of a matrix.

This page titled 7.1: Eigenvalues and Eigenvectors of a Matrix is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **7.1: Eigenvalues and Eigenvectors of a Matrix** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

7.2: Diagonalization

Outcomes

A. Determine when it is possible to diagonalize a matrix.
B. When possible, diagonalize a matrix.

Similarity and Diagonalization

We begin this section by recalling the definition of similar matrices. Recall that if A, B are two $n \times n$ matrices, then they are **similar** if and only if there exists an invertible matrix P such that

$$A = P^{-1}BP$$

In this case we write $A \sim B$. The concept of similarity is an example of an **equivalence relation**.

Lemma 7.2.1:Similarity is an Equivalence Relation

Similarity is an equivalence relation, i.e. for $n \times n$ matrices A, B, and C,

1. $A \sim A$ (reflexive)
2. If $A \sim B$, then $B \sim A$ (symmetric)
3. If $A \sim B$ and $B \sim C$, then $A \sim C$ (transitive)

Proof

It is clear that $A \sim A$, taking $P = I$.

Now, if $A \sim B$, then for some P invertible,

$$A = P^{-1}BP$$

and so

$$PAP^{-1} = B$$

But then

$$\left(P^{-1}\right)^{-1} AP^{-1} = B$$

which shows that $B \sim A$.

Now suppose $A \sim B$ and $B \sim C$. Then there exist invertible matrices P, Q such that

$$A = P^{-1}BP, \ B = Q^{-1}CQ$$

Then,

$$A = P^{-1}\left(Q^{-1}CQ\right)P = (QP)^{-1}C(QP)$$

showing that A is similar to C.

Another important concept necessary to this section is the trace of a matrix. Consider the definition.

Definition 7.2.1:Trace of a Matrix

If $A = [a_{ij}]$ is an $n \times n$ matrix, then the trace of A is

$$\text{trace}(A) = \sum_{i=1}^{n} a_{ii}.$$

In words, the trace of a matrix is the sum of the entries on the main diagonal.

Lemma 7.2.2:Properties of Trace

For $n \times n$ matrices A and B, and any $k \in \mathbb{R}$,

1. $\text{trace}(A+B) = \text{trace}(A) + \text{trace}(B)$
2. $\text{trace}(kA) = k \cdot \text{trace}(A)$
3. $\text{trace}(AB) = \text{trace}(BA)$

The following theorem includes a reference to the characteristic polynomial of a matrix. Recall that for any $n \times n$ matrix A, the characteristic polynomial of A is $c_A(x) = \det(xI - A)$.

Theorem 7.2.1:Properties of Similar Matrices

If A and B are $n \times n$ matrices and $A \sim B$, then

1. $\det(A) = \det(B)$
2. $\text{rank}(A) = \text{rank}(B)$
3. $\text{trace}(A) = \text{trace}(B)$
4. $c_A(x) = c_B(x)$
5. A and B have the same eigenvalues

We now proceed to the main concept of this section. When a matrix is similar to a diagonal matrix, the matrix is said to be diagonalizable. We define a diagonal matrix D as a matrix containing a zero in every entry except those on the main diagonal. More precisely, if d_{ij} is the ij^{th} entry of a diagonal matrix D, then $d_{ij} = 0$ unless $i = j$. Such matrices look like the following.

$$D = \begin{bmatrix} * & & 0 \\ & \ddots & \\ 0 & & * \end{bmatrix}$$

where $*$ is a number which might not be zero.

The following is the formal definition of a diagonalizable matrix.

Definition 7.2.2:Diagonalizable

Let A be an $n \times n$ matrix. Then A is said to be **diagonalizable** if there exists an invertible matrix P such that

$$P^{-1}AP = D$$

where D is a diagonal matrix.

Notice that the above equation can be rearranged as $A = PDP^{-1}$. Suppose we wanted to compute A^{100}. By diagonalizing A first it suffices to then compute $\left(PDP^{-1}\right)^{100}$, which reduces to $PD^{100}P^{-1}$. This last computation is much simpler than A^{100}. While this process is described in detail later, it provides motivation for diagonalization.

Diagonalizing a Matrix

The most important theorem about diagonalizability is the following major result.

Theorem 7.2.2:Eigenvectors and Diagonalizable Matrices

An $n \times n$ matrix A is diagonalizable if and only if there is an invertible matrix P given by

$$P = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}$$

where the X_k are eigenvectors of A.

Moreover if A is diagonalizable, the corresponding eigenvalues of A are the diagonal entries of the diagonal matrix D.

Proof

Suppose P is given as above as an invertible matrix whose columns are eigenvectors of A. Then P^{-1} is of the form

$$P^{-1} = \begin{bmatrix} W_1^T \\ W_2^T \\ \vdots \\ W_n^T \end{bmatrix}$$

where $W_k^T X_j = \delta_{kj},$ which is the Kronecker's symbol defined by

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

Then

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} W_1^T \\ W_2^T \\ \vdots \\ W_n^T \end{bmatrix} \begin{bmatrix} AX_1 & AX_2 & \cdots & AX_n \end{bmatrix} \\ &= \begin{bmatrix} W_1^T \\ W_2^T \\ \vdots \\ W_n^T \end{bmatrix} \begin{bmatrix} \lambda_1 X_1 & \lambda_2 X_2 & \cdots & \lambda_n X_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \end{aligned}$$

Conversely, suppose A is diagonalizable so that $P^{-1}AP = D$. Let

$$P = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix}$$

where the columns are the X_k and

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

Then

$$AP = PD = \begin{bmatrix} X_1 & X_2 & \cdots & X_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

and so

$$\begin{bmatrix} AX_1 & AX_2 & \cdots & AX_n \end{bmatrix} = \begin{bmatrix} \lambda_1 X_1 & \lambda_2 X_2 & \cdots & \lambda_n X_n \end{bmatrix}$$

showing the X_k are eigenvectors of A and the λ_k are eigenvectors.

Below is a video on the diagonalization of matrices



Notice that because the matrix P defined above is invertible it follows that the set of eigenvectors of A, $\{X_1, X_2, \cdots, X_n\}$, form a basis of $\mathbb{R}^n$.

We demonstrate the concept given in the above theorem in the next example. Note that not only are the columns of the matrix P formed by eigenvectors, but P must be invertible so must consist of a wide variety of eigenvectors. We achieve this by using basic eigenvectors for the columns of P.

Example 7.2.1:Diagonalize a Matrix

Let

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & -1 \\ -2 & -4 & 4 \end{bmatrix}$$

Find an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$.

Solution

By Theorem 7.2.2 we use the eigenvectors of A as the columns of P, and the corresponding eigenvalues of A as the diagonal entries of D.

First, we will find the eigenvalues of A. To do so, we solve $\det(\lambda I - A) = 0$ as follows.

$$\det\left(\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & -1 \\ -2 & -4 & 4 \end{bmatrix}\right) = 0$$

This computation is left as an exercise, and you should verify that the eigenvalues are $\lambda_1 = 2, \lambda_2 = 2$, and $\lambda_3 = 6$.

Next, we need to find the eigenvectors. We first find the eigenvectors for $\lambda_1, \lambda_2 = 2$. Solving $(2I - A)X = 0$ to find the eigenvectors, we find that the eigenvectors are

$$t \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

where t, s are scalars. Hence there are two basic eigenvectors which are given by

$$X_1 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, X_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

You can verify that the basic eigenvector for $\lambda_3 = 6$ is $X_3 = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}$

Then, we construct the matrix P as follows.

$$P = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix}$$

That is, the columns of P are the basic eigenvectors of A. Then, you can verify that

$$P^{-1} = \begin{bmatrix} -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{bmatrix}$$

Thus,

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 1 & 4 & -1 \\ -2 & -4 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix} \end{aligned}$$

You can see that the result here is a diagonal matrix where the entries on the main diagonal are the eigenvalues of A. We expected this based on Theorem 7.2.2. Notice that eigenvalues on the main diagonal *must* be in the same order as the corresponding eigenvectors in P.

Below is a video on the diagonalization of matrices given eigenvalues and eigenvectors of a 3x3 matrix



Below is a video on the diagonalization of a 2x2 matrix.



Below is a video on the diagonalization of a 3x3 matrix.



Consider the next important theorem.

Theorem 7.2.3:Linearly Independent Eigenvectors

Let A be an $n \times n$ matrix, and suppose that A has distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$. For each i, let X_i be a λ_i-eigenvector of A. Then $\{X_1, X_2, \ldots, X_m\}$ is linearly independent.

The corollary that follows from this theorem gives a useful tool in determining if A is diagonalizable.

Corollary 7.2.1:Distinct Eigenvalues

Let A be an $n \times n$ matrix and suppose it has n distinct eigenvalues. Then it follows that A is diagonalizable.

It is possible that a matrix A cannot be diagonalized. In other words, we cannot find an invertible matrix P so that $P^{-1}AP = D$.

Consider the following example.

Example 7.2.2:A Matrix which cannot be Diagonalized

Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

If possible, find an invertible matrix P and diagonal matrix D so that $P^{-1}AP = D$.

Solution

Through the usual procedure, we find that the eigenvalues of A are $\lambda_1 = 1, \lambda_2 = 1$. To find the eigenvectors, we solve the equation $(\lambda I - A)\, X = 0$. The matrix $(\lambda I - A)$ is given by

$$\begin{bmatrix} \lambda - 1 & -1 \\ 0 & \lambda - 1 \end{bmatrix}$$

Substituting in $\lambda = 1$, we have the matrix

$$\begin{bmatrix} 1-1 & -1 \\ 0 & 1-1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$$

Then, solving the equation $(\lambda I - A)\, X = 0$ involves carrying the following augmented matrix to its reduced row-echelon form.

$$\left[\begin{array}{cc|c} 0 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 0 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Then the eigenvectors are of the form

$$t \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and the basic eigenvector is

$$X_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

In this case, the matrix A has one eigenvalue of multiplicity two, but only one basic eigenvector. In order to diagonalize A, we need to construct an invertible 2×2 matrix P. However, because A only has one basic eigenvector, we cannot construct this P. Notice that if we were to use X_1 as both columns of P, P would not be invertible. For this reason, we cannot repeat eigenvectors in P.

Hence this matrix cannot be diagonalized.

The idea that a matrix may not be diagonalizable suggests that conditions exist to determine when it is possible to diagonalize a matrix. We saw earlier in Corollary 7.2.1 that an $n \times n$ matrix with n distinct eigenvalues is diagonalizable. It turns out that there are other useful diagonalizability tests.

First we need the following definition.

Definition 7.2.3:Eigenspace

Let A be an $n \times n$ matrix and $\lambda \in \mathbb{R}$. The eigenspace of A corresponding to λ, written $E_\lambda(A)$ is the set of all eigenvectors corresponding to λ.

In other words, the eigenspace $E_\lambda(A)$ is all X such that $AX = \lambda X$. Notice that this set can be written $E_\lambda(A) = \text{null}(\lambda I - A)$, showing that $E_\lambda(A)$ is a subspace of $\mathbb{R}^n$.

Recall that the multiplicity of an eigenvalue λ is the number of times that it occurs as a root of the characteristic polynomial.

Consider now the following lemma.

Lemma 7.2.3:Dimension of the Eigenspace

If A is an $n \times n$ matrix, then

$$\dim(E_\lambda(A)) \leq m$$

where λ is an eigenvalue of A of multiplicity m.

This result tells us that if λ is an eigenvalue of A, then the number of linearly independent λ-eigenvectors is never more than the multiplicity of λ. We now use this fact to provide a useful diagonalizability condition.

Theorem 7.2.4:Diagonalizability Condition

Let A be an $n \times n$ matrix A. Then A is diagonalizable if and only if for each eigenvalue λ of A, $\dim(E_\lambda(A))$ is equal to the multiplicity of λ.

Complex Eigenvalues

In some applications, a matrix may have eigenvalues which are complex numbers. For example, this often occurs in differential equations. These questions are approached in the same way as above.

Consider the following example.

Example 7.2.3:A Real Matrix with Complex Eigenvalues

Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{bmatrix}$$

Find the eigenvalues and eigenvectors of A.

Solution

We will first find the eigenvalues as usual by solving the following equation.

$$\det\left(\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{bmatrix}\right) = 0$$

This reduces to $(\lambda - 1)\left(\lambda^2 - 4\lambda + 5\right) = 0$. The solutions are $\lambda_1 = 1, \lambda_2 = 2 + i$ and $\lambda_3 = 2 - i$.

There is nothing new about finding the eigenvectors for $\lambda_1 = 1$ so this is left as an exercise.

Consider now the eigenvalue $\lambda_2 = 2 + i$. As usual, we solve the equation $(\lambda I - A)\, X = 0$ as given by

$$\left((2+i) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{bmatrix}\right) X = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

In other words, we need to solve the system represented by the augmented matrix

$$\left[\begin{array}{ccc|c} 1+i & 0 & 0 & 0 \\ 0 & i & 1 & 0 \\ 0 & -1 & i & 0 \end{array}\right]$$

We now use our row operations to solve the system. Divide the first row by $(1+i)$ and then take $-i$ times the second row and add to the third row. This yields

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & i & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Now multiply the second row by $-i$ to obtain the reduced row-echelon form, given by

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & -i & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore, the eigenvectors are of the form

$$t\begin{bmatrix}0\\i\\1\end{bmatrix}$$

and the basic eigenvector is given by

$$X_2=\begin{bmatrix}0\\i\\1\end{bmatrix}$$

As an exercise, verify that the eigenvectors for $\lambda_3=2-i$ are of the form

$$t\begin{bmatrix}0\\-i\\1\end{bmatrix}$$

Hence, the basic eigenvector is given by

$$X_3=\begin{bmatrix}0\\-i\\1\end{bmatrix}$$

As usual, be sure to check your answers! To verify, we check that $AX_3=(2-i)\,X_3$ as follows.

$$\begin{bmatrix}1&0&0\\0&2&-1\\0&1&2\end{bmatrix}\begin{bmatrix}0\\-i\\1\end{bmatrix}=\begin{bmatrix}0\\-1-2i\\2-i\end{bmatrix}=(2-i)\begin{bmatrix}0\\-i\\1\end{bmatrix}$$

Therefore, we know that this eigenvector and eigenvalue are correct.

Notice that in Example 7.2.3, two of the eigenvalues were given by $\lambda_2=2+i$ and $\lambda_3=2-i$. You may recall that these two complex numbers are **conjugates**. It turns out that whenever a matrix containing real entries has a complex eigenvalue λ, it also has an eigenvalue equal to $\overline{\lambda}$, the conjugate of λ.

Below is a video on determining Ax using a diagonalized matrix.



This page titled 7.2: Diagonalization is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **7.2: Diagonalization** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

7.3: Applications of Spectral Theory

Outcomes

A. Use diagonalization to find a high power of a matrix.
B. Use diagonalization to solve dynamical systems.

Raising a Matrix to a High Power

Suppose we have a matrix A and we want to find A^{50}. One could try to multiply A with itself 50 times, but this is computationally extremely intensive (try it!). However diagonalization allows us to compute high powers of a matrix relatively easily. Suppose A is diagonalizable, so that $P^{-1}AP = D$. We can rearrange this equation to write $A = PDP^{-1}$.

Now, consider A^2. Since $A = PDP^{-1}$, it follows that

$$A^2 = \left(PDP^{-1}\right)^2 = PDP^{-1}PDP^{-1} = PD^2P^{-1}$$

Similarly,

$$A^3 = \left(PDP^{-1}\right)^3 = PDP^{-1}PDP^{-1}PDP^{-1} = PD^3P^{-1}$$

In general,

$$A^n = \left(PDP^{-1}\right)^n = PD^nP^{-1}$$

Therefore, we have reduced the problem to finding D^n. In order to compute D^n, then because D is diagonal we only need to raise every entry on the main diagonal of D to the power of n.

Below is a video on powers of a diagonal matrix.



Below is a video on diagonalizing a 2x2 matrix to determine a power.



Through this method, we can compute large powers of matrices. Consider the following example.

Example 7.3.1: Raising a Matrix to a High Power

Let $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$. Find A^{50}.

Solution

We will first diagonalize A. The steps are left as an exercise and you may wish to verify that the eigenvalues of A are $\lambda_1 = 1, \lambda_2 = 1$, and $\lambda_3 = 2$.

The basic eigenvectors corresponding to $\lambda_1, \lambda_2 = 1$ are

$$X_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, X_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

The basic eigenvector corresponding to $\lambda_3 = 2$ is

$$X_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Now we construct P by using the basic eigenvectors of A as the columns of P. Thus

$$P = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} = \begin{bmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Then also

$$P^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix}$$

which you may wish to verify.

Then,

$$\begin{aligned} P^{-1}AP &= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \\ &= D \end{aligned}$$

Now it follows by rearranging the equation that

$$A = PDP^{-1} = \begin{bmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix}$$

Therefore,

$$\begin{aligned} A^{50} &= PD^{50}P^{-1} \\ &= \begin{bmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}^{50} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix} \end{aligned}$$

By our discussion above, D^{50} is found as follows.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}^{50} = \begin{bmatrix} 1^{50} & 0 & 0 \\ 0 & 1^{50} & 0 \\ 0 & 0 & 2^{50} \end{bmatrix}$$

It follows that

$$\begin{aligned} A^{50} &= \begin{bmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1^{50} & 0 & 0 \\ 0 & 1^{50} & 0 \\ 0 & 0 & 2^{50} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ -1 & -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2^{50} & -1+2^{50} & 0 \\ 0 & 1 & 0 \\ 1-2^{50} & 1-2^{50} & 1 \end{bmatrix} \end{aligned}$$

Below is a video on diagonalizing a 2x2 matrix to find the matrix raised to a high power.



Through diagonalization, we can efficiently compute a high power of A. Without this, we would be forced to multiply this by hand!

The next section explores another interesting application of diagonalization.

Raising a Symmetric Matrix to a High Power

We already have seen how to use matrix diagonalization to compute powers of matrices. This requires computing eigenvalues of the matrix A, and finding an invertible matrix of eigenvectors P such that $P^{-1}AP$ is diagonal. In this section we will see that if the matrix A is symmetric (see Definition 2.5.2), then we can actually find such a matrix P that is an orthogonal matrix of eigenvectors. Thus P^{-1} is simply its transpose P^T, and P^TAP is diagonal. When this happens we say that A is **orthogonally diagonalizable**

In fact this happens if and only if A is a symmetric matrix as shown in the following important theorem.

Theorem 7.3.1:Principal Axis Theorem

The following conditions are equivalent for an $n \times n$ matrix A:

1. A is symmetric.
2. A has an orthonormal set of eigenvectors.
3. A is orthogonally diagonalizable.

Proof

The complete proof is beyond this course, but to give an idea assume that A has an orthonormal set of eigenvectors, and let P consist of these eigenvectors as columns. Then $P^{-1} = P^T$, and $P^TAP = D$ a diagonal matrix. But then $A = PDP^T$, and

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

so A is symmetric.

Now given a symmetric matrix A, one shows that eigenvectors corresponding to different eigenvalues are always orthogonal. So it suffices to apply the Gram-Schmidt process on the set of basic eigenvectors of each eigenvalue to obtain an orthonormal set of eigenvectors.

We demonstrate this in the following example.

Example 7.3.2:Orthogonal Diagonalization of a Symmetric Matrix

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{3}{2} \end{bmatrix}$. Find an orthogonal matrix P such that P^TAP is a diagonal matrix.

Solution

In this case, verify that the eigenvalues are 2 and 1. First we will find an eigenvector for the eigenvalue 2. This involves row reducing the following augmented matrix.

$$\left[\begin{array}{ccc|c} 2-1 & 0 & 0 & 0 \\ 0 & 2-\frac{3}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 2-\frac{3}{2} & 0 \end{array}\right]$$

The reduced row-echelon form is

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and so an eigenvector is

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Finally to obtain an eigenvector of length one (unit eigenvector) we simply divide this vector by its length to yield:

$$\begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Next consider the case of the eigenvalue 1. To obtain basic eigenvectors, the matrix which needs to be row reduced in this case is

$$\left[\begin{array}{ccc|c} 1-1 & 0 & 0 & 0 \\ 0 & 1-\frac{3}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1-\frac{3}{2} & 0 \end{array}\right]$$

The reduced row-echelon form is

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore, the eigenvectors are of the form

$$\begin{bmatrix} s \\ -t \\ t \end{bmatrix}$$

Note that all these vectors are automatically orthogonal to eigenvectors corresponding to the first eigenvalue. This follows from the fact that A is symmetric, as mentioned earlier.

We obtain basic eigenvectors

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

Since they are themselves orthogonal (by luck here) we do not need to use the Gram-Schmidt process and instead simply normalize these vectors to obtain

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

An orthogonal matrix P to orthogonally diagonalize A is then obtained by letting these basic vectors be the columns.

$$P = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

We verify this works. $P^T AP$ is of the form

$$\begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

which is the desired diagonal matrix.

We can now apply this technique to efficiently compute high powers of a symmetric matrix.

Example 7.3.3:Powers of a Symmetric Matrix

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{3}{2} \end{bmatrix}$. Compute A^7.

Solution

We found in Example 7.3.2 that $P^T AP = D$ is diagonal, where

$$P = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Thus $A = PDP^T$ and $A^7 = PDP^T\ PDP^T\ \cdots PDP^T = PD^7P^T$ which gives:

$$\begin{aligned} A^7 &= \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}^7 \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2^7 \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{2^7}{\sqrt{2}} & \frac{2^7}{\sqrt{2}} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{2^7+1}{2} & \frac{2^7-1}{2} \\ 0 & \frac{2^7-1}{2} & \frac{2^7+1}{2} \end{bmatrix} \end{aligned}$$

Markov Matrices

There are applications of great importance which feature a special type of matrix. Matrices whose columns consist of non-negative numbers that sum to one are called **Markov matrices**. An important application of Markov matrices is in population migration, as illustrated in the following definition.

Definition 7.3.1:Migration Matrices

Let m locations be denoted by the numbers $1, 2, \cdots, m$. Suppose it is the case that each year the proportion of residents in location j which move to location i is a_{ij}. Also suppose no one escapes or emigrates from without these m locations. This last assumption requires $\sum_i a_{ij} = 1$, and means that the matrix A, such that $A = [a_{ij}]$, is a Markov matrix. In this context, A is also called a **migration matrix**.

Consider the following example which demonstrates this situation.

Example 7.3.4:Migration Matrix

Let A be a Markov matrix given by

$$A = \begin{bmatrix} .4 & .2 \\ .6 & .8 \end{bmatrix}$$

Verify that A is a Markov matrix and describe the entries of A in terms of population migration.

Solution

The columns of A are comprised of non-negative numbers which sum to 1. Hence, A is a Markov matrix.

Now, consider the entries a_{ij} of A in terms of population. The entry $a_{11} = .4$ is the proportion of residents in location one which stay in location one in a given time period. Entry $a_{21} = .6$ is the proportion of residents in location 1 which move to location 2 in the same time period. Entry $a_{12} = .2$ is the proportion of residents in location 2 which move to location 1. Finally, entry $a_{22} = .8$ is the proportion of residents in location 2 which stay in location 2 in this time period.

Considered as a Markov matrix, these numbers are usually identified with probabilities. Hence, we can say that the probability that a resident of location one will stay in location one in the time period is .4.

Observe that in Example 7.3.4 if there was initially say 15 thousand people in location 1 and 10 thousands in location 2, then after one year there would be $.4 \times 15 + .2 \times 10 = 8$ thousands people in location 1 the following year, and similarly there would be $.6 \times 15 + .8 \times 10 = 17$ thousands people in location 2 the following year.

More generally let $X_n = [x_{1n} \cdots x_{mn}]^T$ where x_{in} is the population of location i at time period n. We call X_n the **state vector at period n**. In particular, we call X_0 the initial state vector. Letting A be the migration matrix, we compute the population in each location i one time period later by AX_n. In order to find the population of location i after k years, we compute the ith component of $A^k X$. This discussion is summarized in the following theorem.

Theorem 7.3.2:State Vector

Let A be the migration matrix of a population and let X_n be the vector whose entries give the population of each location at time period n. Then X_n is the state vector at period n and it follows that

$$X_{n+1} = AX_n$$

The sum of the entries of X_n will equal the sum of the entries of the initial vector X_0. Since the columns of A sum to 1, this sum is preserved for every multiplication by A as demonstrated below.

$$\sum_i \sum_j a_{ij} x_j = \sum_j x_j \left(\sum_i a_{ij} \right) = \sum_j x_j$$

Consider the following example.

Example 7.3.5:Using a Migration Matrix

Consider the migration matrix

$$A = \begin{bmatrix} .6 & 0 & .1 \\ .2 & .8 & 0 \\ .2 & .2 & .9 \end{bmatrix}$$

for locations $1, 2$, and 3. Suppose initially there are 100 residents in location 1, 200 in location 2 and 400 in location 3. Find the population in the three locations after $1, 2$, and 10 units of time.

Solution

Using Theorem 7.3.2 we can find the population in each location using the equation $X_{n+1} = AX_n$. For the population after 1 unit, we calculate $X_1 = AX_0$ as follows.

$$X_1 = AX_0$$
$$\begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \end{bmatrix} = \begin{bmatrix} .6 & 0 & .1 \\ .2 & .8 & 0 \\ .2 & .2 & .9 \end{bmatrix} \begin{bmatrix} 100 \\ 200 \\ 400 \end{bmatrix} = \begin{bmatrix} 100 \\ 180 \\ 420 \end{bmatrix}$$

Therefore after one time period, location 1 has 100 residents, location 2 has 180, and location 3 has 420. Notice that the **total** population is unchanged, it simply migrates within the given locations. We find the locations after two time periods in the same way.

$$X_2 = AX_1$$
$$\begin{bmatrix} x_{12} \\ x_{22} \\ x_{32} \end{bmatrix} = \begin{bmatrix} .6 & 0 & .1 \\ .2 & .8 & 0 \\ .2 & .2 & .9 \end{bmatrix} \begin{bmatrix} 100 \\ 180 \\ 420 \end{bmatrix} = \begin{bmatrix} 102 \\ 164 \\ 434 \end{bmatrix}$$

We could progress in this manner to find the populations after 10 time periods. However from our above discussion, we can simply calculate $(A^n X_0)_i$, where n denotes the number of time periods which have passed. Therefore, we compute the populations in each location after 10 units of time as follows.

$$X_{10} = A^{10} X_0$$
$$\begin{bmatrix} x_{110} \\ x_{210} \\ x_{310} \end{bmatrix} = \begin{bmatrix} .6 & 0 & .1 \\ .2 & .8 & 0 \\ .2 & .2 & .9 \end{bmatrix}^{10} \begin{bmatrix} 100 \\ 200 \\ 400 \end{bmatrix} = \begin{bmatrix} 115.085\,829\,22 \\ 120.130\,672\,44 \\ 464.783\,498\,34 \end{bmatrix}$$

Since we are speaking about populations, we would need to round these numbers to provide a logical answer. Therefore, we can say that after 10 units of time, there will be 115 residents in location one, 120 in location two, and 465 in location three.

A second important application of Markov matrices is the concept of random walks. Suppose a walker has m locations to choose from, denoted $1, 2, \cdots, m$. Let a_{ij} refer to the probability that the person will travel **to** location i **from** location j. Again, this requires that

$$\sum_{i=1}^{k} a_{ij} = 1$$

In this context, the vector $X_n = [x_{1n} \cdots x_{mn}]^T$ contains the probabilities x_{in} the walker ends up in location $i, 1 \leq i \leq m$ at time n.

Example 7.3.6:Random Walks

Suppose three locations exist, referred to as locations $1, 2$ and 3. The Markov matrix of probabilities $A = [a_{ij}]$ is given by

$$\begin{bmatrix} 0.4 & 0.1 & 0.5 \\ 0.4 & 0.6 & 0.1 \\ 0.2 & 0.3 & 0.4 \end{bmatrix}$$

If the walker starts in location 1, calculate the probability that he ends up in location 3 at time $n = 2$.

Solution

Since the walker begins in location 1, we have

$$X_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

The goal is to calculate x_{32}. To do this we calculate X_2, using $X_{n+1} = AX_n$.

$$\begin{aligned} X_1 &= AX_0 \\ &= \begin{bmatrix} 0.4 & 0.1 & 0.5 \\ 0.4 & 0.6 & 0.1 \\ 0.2 & 0.3 & 0.4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0.4 \\ 0.4 \\ 0.2 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} X_2 &= AX_1 \\ &= \begin{bmatrix} 0.4 & 0.1 & 0.5 \\ 0.4 & 0.6 & 0.1 \\ 0.2 & 0.3 & 0.4 \end{bmatrix} \begin{bmatrix} 0.4 \\ 0.4 \\ 0.2 \end{bmatrix} \\ &= \begin{bmatrix} 0.3 \\ 0.42 \\ 0.28 \end{bmatrix} \end{aligned}$$

This gives the probabilities that our walker ends up in locations 1, 2, and 3. For this example we are interested in location 3, with a probability on 0.28.

Returning to the context of migration, suppose we wish to know how many residents will be in a certain location after a very long time. It turns out that if some power of the migration matrix has all positive entries, then there is a vector X_s such that $A^n X_0$ approaches X_s as n becomes very large. Hence as more time passes and n increases, $A^n X_0$ will become closer to the vector X_s.

Consider Theorem 7.3.2. Let n increase so that X_n approaches X_s. As X_n becomes closer to X_s, so too does X_{n+1}. For sufficiently large n, the statement $X_{n+1} = AX_n$ can be written as $X_s = AX_s$.

This discussion motivates the following theorem.

Theorem 7.3.3:Steady State Vector

Let A be a migration matrix. Then there exists a **steady state vector** written X_s such that

$$X_s = AX_s$$

where X_s has positive entries which have the same sum as the entries of X_0.

As n increases, the state vectors X_n will approach X_s.

Note that the condition in Theorem 7.3.3 can be written as $(I - A)X_s = 0$, representing a homogeneous system of equations.

Consider the following example. Notice that it is the same example as the Example 7.3.5 but here it will involve a longer time frame.

Example 7.3.7:Populations over the Long Run

Consider the migration matrix

$$A = \begin{bmatrix} .6 & 0 & .1 \\ .2 & .8 & 0 \\ .2 & .2 & .9 \end{bmatrix}$$

for locations $1, 2$, and 3. Suppose initially there are 100 residents in location 1, 200 in location 2 and 400 in location 4. Find the population in the three locations after a long time.

Solution

By Theorem 7.3.3 the steady state vector X_s can be found by solving the system $(I - A)X_s = 0$.

Thus we need to find a solution to

$$\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} .6 & 0 & .1 \\ .2 & .8 & 0 \\ .2 & .2 & .9 \end{bmatrix} \right) \begin{bmatrix} x_{1s} \\ x_{2s} \\ x_{3s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The augmented matrix and the resulting reduced row-echelon form are given by

$$\left[\begin{array}{rrr|c} 0.4 & 0 & -0.1 & 0 \\ -0.2 & 0.2 & 0 & 0 \\ -0.2 & -0.2 & 0.1 & 0 \end{array} \right] \rightarrow \cdots \rightarrow \left[\begin{array}{rrr|c} 1 & 0 & -0.25 & 0 \\ 0 & 1 & -0.25 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

Therefore, the eigenvectors are

$$t \begin{bmatrix} 0.25 \\ 0.25 \\ 1 \end{bmatrix}$$

The initial vector X_0 is given by

$$\begin{bmatrix} 100 \\ 200 \\ 400 \end{bmatrix}$$

Now all that remains is to choose the value of t such that

$$0.25t + 0.25t + t = 100 + 200 + 400$$

Solving this equation for t yields $t = \frac{1400}{3}$. Therefore the population in the long run is given by

$$\frac{1400}{3} \begin{bmatrix} 0.25 \\ 0.25 \\ 1 \end{bmatrix} = \begin{bmatrix} 116.666\,666\,666\,666\,7 \\ 116.666\,666\,666\,666\,7 \\ 466.666\,666\,666\,666\,7 \end{bmatrix}$$

Again, because we are working with populations, these values need to be rounded. The steady state vector X_s is given by

$$\begin{bmatrix} 117 \\ 117 \\ 466 \end{bmatrix}$$

We can see that the numbers we calculated in Example 7.3.5 for the populations after the 10th unit of time are not far from the long term values.

Consider another example.

Example 7.3.8:Populations After a Long Time

Suppose a migration matrix is given by

$$A = \begin{bmatrix} \frac{1}{5} & \frac{1}{2} & \frac{1}{5} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{11}{20} & \frac{1}{4} & \frac{3}{10} \end{bmatrix}$$

Find the comparison between the populations in the three locations after a long time.

Solution

In order to compare the populations in the long term, we want to find the steady state vector X_s. Solve

$$\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{5} & \frac{1}{2} & \frac{1}{5} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{11}{20} & \frac{1}{4} & \frac{3}{10} \end{bmatrix} \right) \begin{bmatrix} x_{1s} \\ x_{2s} \\ x_{3s} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The augmented matrix and the resulting reduced row-echelon form are given by

$$\left[\begin{array}{ccc|c} \frac{4}{5} & -\frac{1}{2} & -\frac{1}{5} & 0 \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{2} & 0 \\ -\frac{11}{20} & -\frac{1}{4} & \frac{7}{10} & 0 \end{array} \right] \rightarrow \cdots \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -\frac{16}{19} & 0 \\ 0 & 1 & -\frac{18}{19} & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

and so an eigenvector is

$$\begin{bmatrix} 16 \\ 18 \\ 19 \end{bmatrix}$$

Therefore, the proportion of population in location 2 to location 1 is given by $\frac{18}{16}$. The proportion of population 3 to location 2 is given by $\frac{19}{18}$.

Eigenvalues of Markov Matrices

The following is an important proposition.

Proposition 7.3.1:Eigenvalues of a Migration Matrix

Let $A = [a_{ij}]$ be a migration matrix. Then 1 is always an eigenvalue for A.

Proof

Remember that the determinant of a matrix always equals that of its transpose. Therefore,

$$\det(\lambda I - A) = \det\left((\lambda I - A)^T \right) = \det\left(\lambda I - A^T \right)$$

because $I^T = I$. Thus the characteristic equation for A is the same as the characteristic equation for A^T. Consequently, A and A^T have the same eigenvalues. We will show that 1 is an eigenvalue for A^T and then it will follow that 1 is an eigenvalue for A.

Remember that for a migration matrix, $\sum_i a_{ij} = 1$. Therefore, if $A^T = [b_{ij}]$ with $b_{ij} = a_{ji}$, it follows that

$$\sum_j b_{ij} = \sum_j a_{ji} = 1$$

Therefore, from matrix multiplication,

$$A^T \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} \sum_j b_{ij} \\ \vdots \\ \sum_j b_{ij} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Notice that this shows that $\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$ is an eigenvector for A^T corresponding to the eigenvalue, $\lambda = 1$. As explained above, this shows that $\lambda = 1$ is an eigenvalue for A because A and A^T have the same eigenvalues.

Below is a video on powers of matrices and Markov matrices.



Dynamical Systems

The migration matrices discussed above give an example of a discrete dynamical system. We call them discrete because they involve discrete values taken at a sequence of points rather than on a continuous interval of time.

An example of a situation which can be studied in this way is a predator prey model. Consider the following model where x is the number of prey and y the number of predators in a certain area at a certain time. These are functions of $n \in \mathbb{N}$ where $n = 1, 2, \cdots$ are the ends of intervals of time which may be of interest in the problem. In other words, $x\left(n\right)$ is the number of prey at the end of the nth interval of time. An example of this situation may be modeled by the following equation

$$\begin{bmatrix} x\left(n+1\right) \\ y\left(n+1\right) \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} x\left(n\right) \\ y\left(n\right) \end{bmatrix}$$

This says that from time period n to $n+1$, x increases if there are more x and decreases as there are more y. In the context of this example, this means that as the number of predators increases, the number of prey decreases. As for y, it increases if there are more y and also if there are more x.

This is an example of a matrix recurrence which we define now.

Definition 7.3.2:Matrix Recurrence

Suppose a dynamical system is given by

$$\begin{aligned} x_{n+1} &= ax_n + by_n \\ y_{n+1} &= cx_n + dy_n \end{aligned}$$

This system can be expressed as $V_{n+1} = AV_n$ where $V_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix}$ and $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

In this section, we will examine how to find solutions to a dynamical system given certain initial conditions. This process involves several concepts previously studied, including matrix diagonalization and Markov matrices. The procedure is given as follows. Recall that when diagonalized, we can write $A^n = PD^nP^{-1}$.

Procedure 7.3.1:Solving a Dynamical System

Suppose a dynamical system is given by

$$\begin{aligned} x_{n+1} &= ax_n + by_n \\ y_{n+1} &= cx_n + dy_n \end{aligned}$$

Given initial conditions x_0 and y_0, the solutions to the system are found as follows:

1. Express the dynamical system in the form $V_{n+1} = AV_n$.
2. Diagonalize A to be written as $A = PDP^{-1}$.
3. Then $V_n = PD^nP^{-1}V_0$ where V_0 is the vector containing the initial conditions.

4. If given specific values for n, substitute into this equation. Otherwise, find a general solution for n.

We will now consider an example in detail.

Example 7.3.9: Solutions of a Discrete Dynamical System

Suppose a dynamical system is given by

$$\begin{aligned} x_{n+1} &= 1.5x_n - 0.5y_n \\ y_{n+1} &= 1.0x_n \end{aligned}$$

Express this system as a matrix recurrence and find solutions to the dynamical system for initial conditions $x_0 = 20, y_0 = 10$.

Solution

First, we express the system as a matrix recurrence.

$$\begin{aligned} V_{n+1} &= AV_n \\ \begin{bmatrix} x(n+1) \\ y(n+1) \end{bmatrix} &= \begin{bmatrix} 1.5 & -0.5 \\ 1.0 & 0 \end{bmatrix} \begin{bmatrix} x(n) \\ y(n) \end{bmatrix} \end{aligned}$$

Then

$$A = \begin{bmatrix} 1.5 & -0.5 \\ 1.0 & 0 \end{bmatrix}$$

You can verify that the eigenvalues of A are 1 and .5. By diagonalizing, we can write A in the form

$$P^{-1}DP = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & .5 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

Now given an initial condition

$$V_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$$

the solution to the dynamical system is given by

$$\begin{aligned} V_n &= PD^nP^{-1}V_0 \\ \begin{bmatrix} x(n) \\ y(n) \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & .5 \end{bmatrix}^n \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & (.5)^n \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \\ &= \begin{bmatrix} y_0((.5)^n - 1) - x_0((.5)^n - 2) \\ y_0(2(.5)^n - 1) - x_0(2(.5)^n - 2) \end{bmatrix} \end{aligned}$$

If we let n become arbitrarily large, this vector approaches

$$\begin{bmatrix} 2x_0 - y_0 \\ 2x_0 - y_0 \end{bmatrix}$$

Thus for large n,

$$\begin{bmatrix} x(n) \\ y(n) \end{bmatrix} \approx \begin{bmatrix} 2x_0 - y_0 \\ 2x_0 - y_0 \end{bmatrix}$$

Now suppose the initial condition is given by

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 20 \\ 10 \end{bmatrix}$$

Then, we can find solutions for various values of n. Here are the solutions for values of n between 1 and 5

$$n=1:\begin{bmatrix}25.0\\20.0\end{bmatrix},n=2:\begin{bmatrix}27.5\\25.0\end{bmatrix},n=3:\begin{bmatrix}28.75\\27.5\end{bmatrix}$$

$$n=4:\begin{bmatrix}29.375\\28.75\end{bmatrix},n=5:\begin{bmatrix}29.688\\29.375\end{bmatrix}$$

Notice that as n increases, we approach the vector given by

$$\begin{bmatrix}2x_0-y_0\\2x_0-y_0\end{bmatrix}=\begin{bmatrix}2\left(20\right)-10\\2\left(20\right)-10\end{bmatrix}=\begin{bmatrix}30\\30\end{bmatrix}$$

These solutions are graphed in the following figure.

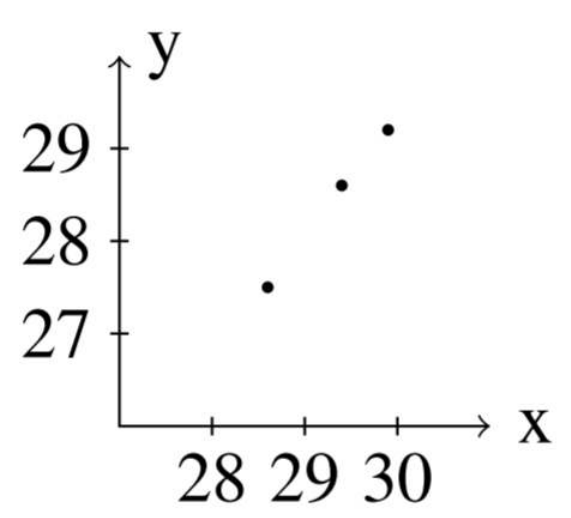


Figure 7.3.1

The following example demonstrates another system which exhibits some interesting behavior. When we graph the solutions, it is possible for the ordered pairs to spiral around the origin.

Example 7.3.10:Finding Solutions to a Dynamical System

Suppose a dynamical system is of the form

$$\begin{bmatrix}x\left(n+1\right)\\y\left(n+1\right)\end{bmatrix}=\begin{bmatrix}0.7&0.7\\-0.7&0.7\end{bmatrix}\begin{bmatrix}x\left(n\right)\\y\left(n\right)\end{bmatrix}$$

Find solutions to the dynamical system for given initial conditions.

Solution

Let

$$A=\begin{bmatrix}0.7&0.7\\-0.7&0.7\end{bmatrix}$$

To find solutions, we must diagonalize A. You can verify that the eigenvalues of A are complex and are given by $\lambda_1=.7+.7i$ and $\lambda_2=.7-.7i$. The eigenvector for $\lambda_1=.7+.7i$ is

$$\begin{bmatrix}1\\i\end{bmatrix}$$

and that the eigenvector for $\lambda_2=.7-.7i$ is

$$\begin{bmatrix}1\\-i\end{bmatrix}$$

Thus the matrix A can be written in the form

$$\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} .7+.7i & 0 \\ 0 & .7-.7i \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}i \\ \frac{1}{2} & \frac{1}{2}i \end{bmatrix}$$

and so,

$$\begin{aligned} V_n &= PD^nP^{-1}V_0 \\ \begin{bmatrix} x(n) \\ y(n) \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} (.7+.7i)^n & 0 \\ 0 & (.7-.7i)^n \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2}i \\ \frac{1}{2} & \frac{1}{2}i \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \end{aligned}$$

The explicit solution is given by

$$\begin{bmatrix} x_0\left(\frac{1}{2}(0.7-0.7i)^n+\frac{1}{2}(0.7+0.7i)^n\right)+y_0\left(\frac{1}{2}i(0.7-0.7i)^n-\frac{1}{2}i(0.7+0.7i)^n\right) \\ y_0\left(\frac{1}{2}(0.7-0.7i)^n+\frac{1}{2}(0.7+0.7i)^n\right)-x_0\left(\frac{1}{2}i(0.7-0.7i)^n-\frac{1}{2}i(0.7+0.7i)^n\right) \end{bmatrix}$$

Suppose the initial condition is

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$$

Then one obtains the following sequence of values which are graphed below by letting $n = 1, 2, \cdots, 20$

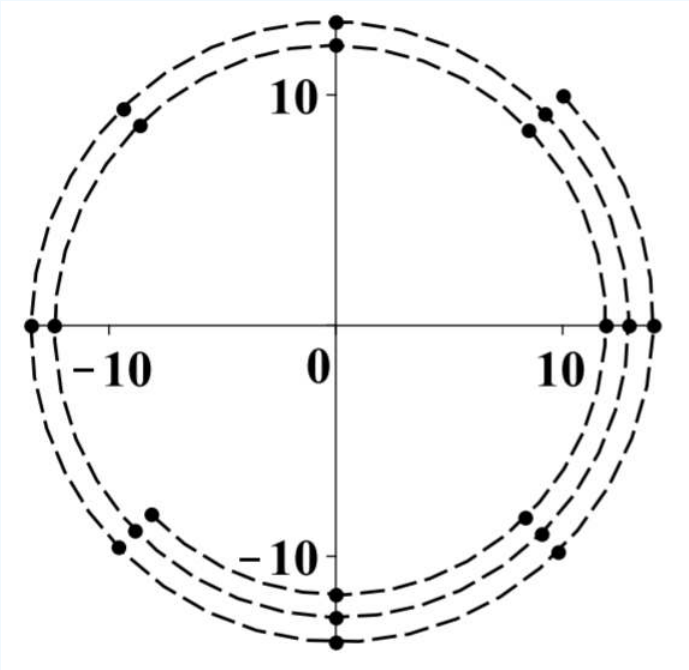


Figure 7.3.2

In this picture, the dots are the values and the dashed line is to help to picture what is happening.

These points are getting gradually closer to the origin, but they are circling the origin in the clockwise direction as they do so. As n increases, the vector $\begin{bmatrix} x(n) \\ y(n) \end{bmatrix}$ approaches $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

This type of behavior along with complex eigenvalues is typical of the deviations from an equilibrium point in the Lotka Volterra system of differential equations which is a famous model for predator-prey interactions. These differential equations are given by

$$\begin{aligned} x' &= x(a-by) \\ y' &= -y(c-dx) \end{aligned}$$

where a, b, c, d are positive constants. For example, you might have X be the population of moose and Y the population of wolves on an island.

Note that these equations make logical sense. The top says that the rate at which the moose population increases would be aX if there were no predators Y. However, this is modified by multiplying instead by $(a - bY)$ because if there are predators, these will militate against the population of moose. The more predators there are, the more pronounced is this effect. As to the predator equation, you can see that the equations predict that if there are many prey around, then the rate of growth of the predators would seem to be high. However, this is modified by the term $-cY$ because if there are many predators, there would be competition for the available food supply and this would tend to decrease Y'.

The behavior near an equilibrium point, which is a point where the right side of the differential equations equals zero, is of great interest. In this case, the equilibrium point is

$$x = \frac{c}{d}, y = \frac{a}{b}$$

Then one defines new variables according to the formula

$$x + \frac{c}{d} = x,\ y = y + \frac{a}{b}$$

In terms of these new variables, the differential equations become

$$\begin{aligned} x' &= \left(x + \frac{c}{d}\right)\left(a - b\left(y + \frac{a}{b}\right)\right) \\ y' &= -\left(y + \frac{a}{b}\right)\left(c - d\left(x + \frac{c}{d}\right)\right) \end{aligned}$$

Multiplying out the right sides yields

$$\begin{aligned} x' &= -bxy - b\frac{c}{d}y \\ y' &= dxy + \frac{a}{b}dx \end{aligned}$$

The interest is for x, y small and so these equations are essentially equal to

$$x' = -b\frac{c}{d}y,\ y' = \frac{a}{b}dx$$

Replace x' with the difference quotient $\frac{x(t+h)-x(t)}{h}$ where h is a small positive number and y' with a similar difference quotient. For example one could have h correspond to one day or even one hour. Thus, for h small enough, the following would seem to be a good approximation to the differential equations.

$$\begin{aligned} x(t+h) &= x(t) - hb\frac{c}{d}y \\ y(t+h) &= y(t) + h\frac{a}{b}dx \end{aligned}$$

Let $1, 2, 3, \cdots$ denote the ends of discrete intervals of time having length h chosen above. Then the above equations take the form

$$\begin{bmatrix} x(n+1) \\ y(n+1) \end{bmatrix} = \begin{bmatrix} 1 & -\frac{hbc}{d} \\ \frac{had}{b} & 1 \end{bmatrix} \begin{bmatrix} x(n) \\ y(n) \end{bmatrix}$$

Note that the eigenvalues of this matrix are always complex.

We are not interested in time intervals of length h for h very small. Instead, we are interested in much longer lengths of time. Thus, replacing the time interval with mh,

$$\begin{bmatrix} x(n+m) \\ y(n+m) \end{bmatrix} = \begin{bmatrix} 1 & -\frac{hbc}{d} \\ \frac{had}{b} & 1 \end{bmatrix}^m \begin{bmatrix} x(n) \\ y(n) \end{bmatrix}$$

For example, if $m = 2$, you would have

$$\begin{bmatrix} x(n+2) \\ y(n+2) \end{bmatrix} = \begin{bmatrix} 1-ach^2 & -2b\frac{c}{d}h \\ 2\frac{a}{b}dh & 1-ach^2 \end{bmatrix} \begin{bmatrix} x(n) \\ y(n) \end{bmatrix}$$

Note that most of the time, the eigenvalues of the new matrix will be complex.

You can also notice that the upper right corner will be negative by considering higher powers of the matrix. Thus letting $1, 2, 3, \cdots$ denote the ends of discrete intervals of time, the desired discrete dynamical system is of the form

$$\begin{bmatrix} x(n+1) \\ y(n+1) \end{bmatrix} = \begin{bmatrix} a & -b \\ c & d \end{bmatrix} \begin{bmatrix} x(n) \\ y(n) \end{bmatrix}$$

where a, b, c, d are positive constants and the matrix will likely have complex eigenvalues because it is a power of a matrix which has complex eigenvalues.

You can see from the above discussion that if the eigenvalues of the matrix used to define the dynamical system are less than 1 in absolute value, then the origin is stable in the sense that as $n \to \infty$, the solution converges to the origin. If either eigenvalue is larger than 1 in absolute value, then the solutions to the dynamical system will usually be unbounded, unless the initial condition is chosen very carefully. The next example exhibits the case where one eigenvalue is larger than 1 and the other is smaller than 1.

The following example demonstrates a familiar concept as a dynamical system.

Example 7.3.11:The Fibonacci Sequence

The Fibonacci sequence is the sequence given by

$$1, 1, 2, 3, 5, \cdots$$

which is defined recursively in the form

$$x(0) = 1 = x(1),\ x(n+2) = x(n+1) + x(n)$$

Show how the Fibonacci Sequence can be considered a dynamical system.

Solution

This sequence is extremely important in the study of reproducing rabbits. It can be considered as a dynamical system as follows. Let $y(n) = x(n+1)$. Then the above recurrence relation can be written as

$$\begin{bmatrix} x(n+1) \\ y(n+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x(n) \\ y(n) \end{bmatrix}, \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Let

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

The eigenvalues of the matrix A are $\lambda_1 = \frac{1}{2} - \frac{1}{2}\sqrt{5}$ and $\lambda_2 = \frac{1}{2}\sqrt{5} + \frac{1}{2}$. The corresponding eigenvectors are, respectively,

$$X_1 = \begin{bmatrix} -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{bmatrix}, X_2 = \begin{bmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 \end{bmatrix}$$

You can see from a short computation that one of the eigenvalues is smaller than 1 in absolute value while the other is larger than 1 in absolute value. Now, diagonalizing A gives us

$$\begin{bmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} & -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2}\sqrt{5} - \frac{1}{2} & -\frac{1}{2}\sqrt{5} - \frac{1}{2} \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{2}\sqrt{5} + \frac{1}{2} & 0 \\ 0 & \frac{1}{2} - \frac{1}{2}\sqrt{5} \end{bmatrix}$$

Then it follows that for a given initial condition, the solution to this dynamical system is of the form

$$\begin{bmatrix} x(n) \\ y(n) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sqrt{5}-\frac{1}{2} & -\frac{1}{2}\sqrt{5}-\frac{1}{2} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \left(\frac{1}{2}\sqrt{5}+\frac{1}{2}\right)^n & 0 \\ 0 & \left(\frac{1}{2}-\frac{1}{2}\sqrt{5}\right)^n \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{5}\sqrt{5} & \frac{1}{10}\sqrt{5}+\frac{1}{2} \\ -\frac{1}{5}\sqrt{5} & \frac{1}{5}\sqrt{5}\left(\frac{1}{2}\sqrt{5}-\frac{1}{2}\right) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

It follows that

$$x(n) = \left(\frac{1}{2}\sqrt{5}+\frac{1}{2}\right)^n \left(\frac{1}{10}\sqrt{5}+\frac{1}{2}\right) + \left(\frac{1}{2}-\frac{1}{2}\sqrt{5}\right)^n \left(\frac{1}{2}-\frac{1}{10}\sqrt{5}\right)$$

Here is a picture of the ordered pairs $(x(n), y(n))$ for $n = 0, 1, \cdots, n$.

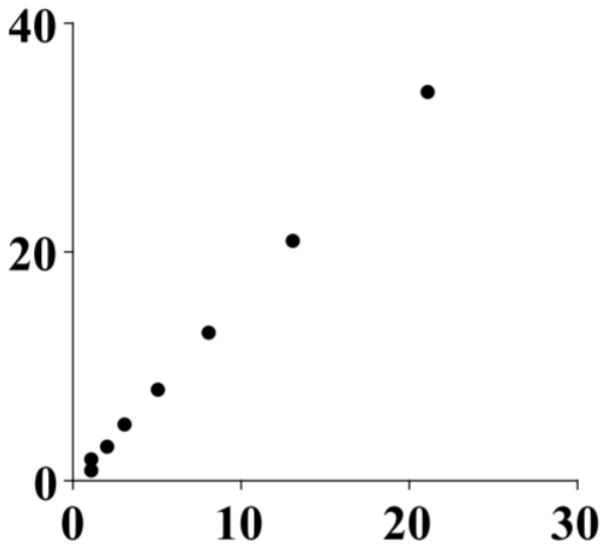


Figure 7.3.3

There is so much more that can be said about dynamical systems. It is a major topic of study in differential equations and what is given above is just an introduction.

Below is a video Fibonacci numbers using linear algebra.



The Matrix Exponential

The goal of this section is to use the concept of the matrix exponential to solve first order linear differential equations. We begin by proving the matrix exponential.

Suppose A is a diagonalizable matrix. Then the **matrix exponential**, written e^A, can be easily defined. Recall that if D is a diagonal matrix, then

$$P^{-1}AP = D$$

D is of the form

$$\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \tag{7.3.1}$$

and it follows that

$$D^m = \begin{bmatrix} \lambda_1^m & & 0 \\ & \ddots & \\ 0 & & \lambda_n^m \end{bmatrix}$$

Since A is diagonalizable,

$$A = PDP^{-1}$$

and

$$A^m = PD^mP^{-1}$$

Recall why this is true.

$$A = PDP^{-1}$$

and so

$$\begin{aligned} A^m &= \overbrace{PDP^{-1}PDP^{-1}PDP^{-1}\cdots PDP^{-1}}^{\text{m times}} \\ &= PD^mP^{-1} \end{aligned}$$

We now will examine what is meant by the matrix exponental e^A. Begin by formally writing the following power series for e^A:

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} \frac{PD^kP^{-1}}{k!} = P\left(\sum_{k=0}^{\infty} \frac{D^k}{k!}\right)P^{-1}$$

If D is given above in (7.3.1), the above sum is of the form

$$P\left(\sum_{k=0}^{\infty} \begin{bmatrix} \frac{1}{k!}\lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \frac{1}{k!}\lambda_n^k \end{bmatrix}\right)P^{-1}$$

This can be rearranged as follows:

$$\begin{aligned} e^A &= P\begin{bmatrix} \sum_{k=0}^{\infty}\frac{1}{k!}\lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \sum_{k=0}^{\infty}\frac{1}{k!}\lambda_n^k \end{bmatrix}P^{-1} \\ &= P\begin{bmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{bmatrix}P^{-1} \end{aligned}$$

This justifies the following theorem.

Theorem 7.3.4:The Matrix Exponential

Let A be a diagonalizable matrix, with eigenvalues $\lambda_1, \ldots, \lambda_n$and corresponding matrix of eigenvectors P. Then the matrix exponential, e^A, is given by

$$e^A = P \begin{bmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{bmatrix} P^{-1}$$

Example 7.3.12: Compute e^A for a Matrix A

Let

$$A = \begin{bmatrix} 2 & -1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

Find e^A.

Solution

The eigenvalues work out to be $1, 2, 3$ and eigenvectors associated with these eigenvalues are

$$\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \leftrightarrow 1, \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \leftrightarrow 2, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \leftrightarrow 3$$

Then let

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, P = \begin{bmatrix} 0 & -1 & -1 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

and so

$$P^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & -1 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

Then the matrix exponential is

$$e^{At} = \begin{bmatrix} 0 & -1 & -1 \\ -1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} e^1 & 0 & 0 \\ 0 & e^2 & 0 \\ 0 & 0 & e^3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -1 & -1 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} e^2 & e^2 - e^3 & e^2 - e^3 \\ e^2 - e & e^2 & e^2 - e \\ -e^2 + e & -e^2 + e^3 & -e^2 + e + e^3 \end{bmatrix}$$

The matrix exponential is a useful tool to solve autonomous systems of first order linear differential equations. These are equations which are of the form

$$X' = AX, X(0) = C$$

where A is a diagonalizable $n \times n$ matrix and C is a constant vector. X is a vector of functions in one variable, t:

$$X = X(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

Then X' refers to the first derivative of X and is given by

$$X' = X'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}, \; x_i'(t) = \text{the derivative of } x_i(t)$$

Then it turns out that the solution to the above system of equations is $X(t) = e^{At}C$. To see this, suppose A is diagonalizable so that

$$A = P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} P^{-1}$$

Then

$$e^{At} = P \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} P^{-1}$$

$$e^{At}C = P \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} P^{-1}C$$

Differentiating $e^{At}C$ yields

$$X' = \left(e^{At}C\right)' = P \begin{bmatrix} \lambda_1 e^{\lambda_1 t} & & \\ & \ddots & \\ & & \lambda_n e^{\lambda_n t} \end{bmatrix} P^{-1}C$$

$$= P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} P^{-1}C$$

$$= P \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} P^{-1} P \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} P^{-1}C$$

$$= A\left(e^{At}C\right) = AX$$

Therefore $X = X(t) = e^{At}C$ is a solution to $X' = AX$.

To prove that $X(0) = C$ if $X(t) = e^{At}C$:

$$X(0) = e^{A0}C = P \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} P^{-1}C = C$$

Example 7.3.13: Solving an Initial Value Problem

Solve the initial value problem

$$\begin{bmatrix} x \\ y \end{bmatrix}' = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Solution

The matrix is diagonalizable and can be written as

$$\begin{aligned} A &= PDP^{-1} \\ \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -1 & -2 \end{bmatrix} \end{aligned}$$

Therefore, the matrix exponential is of the form

$$e^{At} = \begin{bmatrix} 1 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -1 & -2 \end{bmatrix}$$

The solution to the initial value problem is

$$\begin{aligned} X(t) &= e^{At}C \\ \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ -\frac{1}{2} & -1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{2t} \end{bmatrix} \begin{bmatrix} 2 & 2 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 4e^t - 3e^{2t} \\ 3e^{2t} - 2e^t \end{bmatrix} \end{aligned}$$

We can check that this works:

$$\begin{aligned} \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} &= \begin{bmatrix} 4e^0 - 3e^{2(0)} \\ 3e^{2(0)} - 2e^0 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}$$

Lastly,

$$X' = \begin{bmatrix} 4e^t - 3e^{2t} \\ 3e^{2t} - 2e^t \end{bmatrix}' = \begin{bmatrix} 4e^t - 6e^{2t} \\ 6e^{2t} - 2e^t \end{bmatrix}$$

and

$$AX = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 4e^t - 3e^{2t} \\ 3e^{2t} - 2e^t \end{bmatrix} = \begin{bmatrix} 4e^t - 6e^{2t} \\ 6e^{2t} - 2e^t \end{bmatrix}$$

which is the same thing. Thus this is the solution to the initial value problem.

Below is a video on the using linear algebra to solve systems of differential equations.



This page titled 7.3: Applications of Spectral Theory is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **7.3: Applications of Spectral Theory** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

7.4: Orthogonality

Orthogonal Diagonalization

We begin this section by recalling some important definitions. Recall from Definition 4.11.4 that non-zero vectors are called orthogonal if their dot product equals 0. A set is orthonormal if it is orthogonal and each vector is a unit vector.

An orthogonal matrix U, from Definition 4.11.7, is one in which $UU^T = I$. In other words, the transpose of an orthogonal matrix is equal to its inverse. A key characteristic of orthogonal matrices, which will be essential in this section, is that the columns of an orthogonal matrix form an orthonormal set.

We now recall another important definition.

Definition 7.4.1: Symmetric and Skew Symmetric Matrices

A real $n \times n$ matrix A, is **symmetric** if $A^T = A$. If $A = -A^T$, then A is called **skew symmetric.**

Before proving an essential theorem, we first examine the following lemma which will be used below.

Lemma 7.4.1: The Dot Product

Let $A = (a_{ij})$ be a real symmetric $n \times n$ matrix, and let $\vec{x}, \vec{y} \in \mathbb{R}^n$. Then

$$A\vec{x} \cdot \vec{y} = \vec{x} \cdot A\vec{y}$$

Proof

This result follows from the definition of the dot product together with properties of matrix multiplication, as follows:

$$\begin{aligned} A\vec{x} \cdot \vec{y} &= \sum_{k,l} a_{kl} x_l y_k \\ &= \sum_{k,l} (a_{lk})^T x_l y_k \\ &= \vec{x} \cdot A^T \vec{y} \\ &= \vec{x} \cdot A\vec{y} \end{aligned}$$

The last step follows from $A^T = A$, since A is symmetric.

We can now prove that the eigenvalues of a real symmetric matrix are real numbers. Consider the following important theorem.

Theorem 7.4.1: Orthogonal Eigenvectors

Let A be a real symmetric matrix. Then the eigenvalues of A are real numbers and eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof

Recall that for a complex number $a+ib$, the complex conjugate, denoted by $\overline{a+ib}$ is given by $\overline{a+ib} = a-ib$. The notation, $\overline{\vec{x}}$ will denote the vector which has every entry replaced by its complex conjugate.

Suppose A is a real symmetric matrix and $A\vec{x} = \lambda\vec{x}$. Then

$$\overline{\lambda}\overline{\vec{x}}^T\vec{x} = \left(\overline{A\vec{x}}\right)^T\vec{x} = \overline{\vec{x}}^T A^T\vec{x} = \overline{\vec{x}}^T A\vec{x} = \lambda\overline{\vec{x}}^T\vec{x}$$

Dividing by $\overline{\vec{x}}^T\vec{x}$ on both sides yields $\overline{\lambda} = \lambda$ which says λ is real. To do this, we need to ensure that $\overline{\vec{x}}^T\vec{x} \neq 0$. Notice that $\overline{\vec{x}}^T\vec{x} = 0$ if and only if $\vec{x} = \vec{0}$. Since we chose $\vec{x}$ such that $A\vec{x} = \lambda\vec{x}$, $\vec{x}$ is an eigenvector and therefore must be nonzero.

Now suppose A is real symmetric and $A\vec{x} = \lambda\vec{x}$, $A\vec{y} = \mu\vec{y}$ where $\mu \neq \lambda$. Then since A is symmetric, it follows from Lemma 7.4.1about the dot product that

$$\lambda\vec{x} \cdot \vec{y} = A\vec{x} \cdot \vec{y} = \vec{x} \cdot A\vec{y} = \vec{x} \cdot \mu\vec{y} = \mu\vec{x} \cdot \vec{y}$$

Hence $(\lambda - \mu)\vec{x} \cdot \vec{y} = 0$. It follows that, since $\lambda - \mu \neq 0$, it must be that $\vec{x} \cdot \vec{y} = 0$. Therefore the eigenvectors form an orthogonal set.

The following theorem is proved in a similar manner.

Theorem 7.4.2: Eigenvalues of Skew Symmetric Matrix

The eigenvalues of a real skew symmetric matrix are either equal to 0 or are pure imaginary numbers.

Proof

First, note that if $A = 0$ is the zero matrix, then A is skew symmetric and has eigenvalues equal to 0.

Suppose $A = -A^T$ so A is skew symmetric and $A\vec{x} = \lambda\vec{x}$. Then

$$\overline{\lambda\vec{x}}^T\vec{x} = \left(\overline{A\vec{x}}\right)^T\vec{x} = \overline{\vec{x}}^T A^T\vec{x} = -\overline{\vec{x}}^T A\vec{x} = -\lambda\overline{\vec{x}}^T\vec{x}$$

and so, dividing by $\overline{\vec{x}}^T\vec{x}$ as before, $\overline{\lambda} = -\lambda$. Letting $\lambda = a + ib$, this means $a - ib = -a - ib$ and so $a = 0$. Thus λ is pure imaginary.

Consider the following example.

Example 7.4.1: Eigenvalues of a Skew Symmetric Matrix

Let $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Find its eigenvalues.

Solution

First notice that A is skew symmetric. By Theorem 7.4.2, the eigenvalues will either equal 0 or be pure imaginary. The eigenvalues of A are obtained by solving the usual equation

$$\det(\lambda I - A) = \det\begin{bmatrix} \lambda & 1 \\ -1 & \lambda \end{bmatrix} = \lambda^2 + 1 = 0$$

Hence the eigenvalues are $\pm i$, pure imaginary.

Consider the following example.

Example 7.4.2: Eigenvalues of a Symmetric Matrix

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$. Find its eigenvalues.

Solution

First, notice that A is symmetric. By Theorem 7.4.1, the eigenvalues will all be real. The eigenvalues of A are obtained by solving the usual equation

$$\det(\lambda I - A) = \det\begin{bmatrix} \lambda - 1 & -2 \\ -2 & \lambda - 3 \end{bmatrix} = \lambda^2 - 4\lambda - 1 = 0$$

The eigenvalues are given by $\lambda_1 = 2 + \sqrt{5}$ and $\lambda_2 = 2 - \sqrt{5}$ which are both real.

Recall that a diagonal matrix $D = (d_{ij})$ is one in which $d_{ij} = 0$ whenever $i \neq j$. In other words, all numbers not on the main diagonal are equal to zero.

Consider the following important theorem.

Theorem 7.4.3: Orthogonal Diagonalization

Let A be a real symmetric matrix. Then there exists an orthogonal matrix U such that

$$U^T AU = D$$

where D is a diagonal matrix. Moreover, the diagonal entries of D are the eigenvalues of A.

We can use this theorem to diagonalize a symmetric matrix, using orthogonal matrices. Consider the following corollary.

Corollary 7.4.1: Orthonormal Set of Eigenvectors

If A is a real $n \times n$ symmetric matrix, then there exists an orthonormal set of eigenvectors, $\{\vec{u}_1, \cdots, \vec{u}_n\}$.

Proof

Since A is symmetric, then by Theorem 7.4.3, there exists an orthogonal matrix U such that $U^T AU = D$, a diagonal matrix whose diagonal entries are the eigenvalues of A. Therefore, since A is symmetric and all the matrices are real,

$$\overline{D} = \overline{D^T} = \overline{U^T A^T U} = U^T A^T U = U^T AU = D$$

showing D is real because each entry of D equals its complex conjugate.

Now let

$$U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \end{bmatrix}$$

where the $\vec{u}_i$ denote the columns of U and

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

The equation, $U^T AU = D$ implies $AU = UD$ and

$$\begin{aligned} AU &= \begin{bmatrix} A\vec{u}_1 & A\vec{u}_2 & \cdots & A\vec{u}_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 \vec{u}_1 & \lambda_2 \vec{u}_2 & \cdots & \lambda_n \vec{u}_n \end{bmatrix} \\ &= UD \end{aligned}$$

where the entries denote the columns of AU and UD respectively. Therefore, $A\vec{u}_i = \lambda_i \vec{u}_i$. Since the matrix U is orthogonal, the ij^{th} entry of $U^T U$ equals δ_{ij} and so

$$\delta_{ij} = \vec{u}_i^T \vec{u}_j = \vec{u}_i \cdot \vec{u}_j$$

This proves the corollary because it shows the vectors $\{\vec{u}_i\}$ form an orthonormal set.

Definition 7.4.2: Principal Axes

Let A be an $n \times n$ matrix. Then the principal axes of A is a set of orthonormal eigenvectors of A.

In the next example, we examine how to find such a set of orthonormal eigenvectors.

Example 7.4.3: Find an Orthonormal Set of Eigenvectors

Find an orthonormal set of eigenvectors for the symmetric matrix

$$A = \begin{bmatrix} 17 & -2 & -2 \\ -2 & 6 & 4 \\ -2 & 4 & 6 \end{bmatrix}$$

Solution

Recall Procedure 7.1.1 for finding the eigenvalues and eigenvectors of a matrix. You can verify that the eigenvalues are $18, 9, 2$. First find the eigenvector for 18 by solving the equation $(18I - A)X = 0$. The appropriate augmented matrix is given by

$$\left[\begin{array}{ccc|c} 18-17 & 2 & 2 & 0 \\ 2 & 18-6 & -4 & 0 \\ 2 & -4 & 18-6 & 0 \end{array}\right]$$

The reduced row-echelon form is

$$\left[\begin{array}{ccc|c} 1 & 0 & 4 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore an eigenvector is

$$\begin{bmatrix} -4 \\ 1 \\ 1 \end{bmatrix}$$

Next find the eigenvector for $\lambda = 9$. The augmented matrix and resulting reduced row-echelon form are

$$\left[\begin{array}{ccc|c} 9-17 & 2 & 2 & 0 \\ 2 & 9-6 & -4 & 0 \\ 2 & -4 & 9-6 & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus an eigenvector for $\lambda = 9$ is

$$\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Finally find an eigenvector for $\lambda = 2$. The appropriate augmented matrix and reduced row-echelon form are

$$\left[\begin{array}{ccc|c} 2-17 & 2 & 2 & 0 \\ 2 & 2-6 & -4 & 0 \\ 2 & -4 & 2-6 & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus an eigenvector for $\lambda = 2$ is

$$\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$$

The set of eigenvectors for A is given by

$$\left\{ \begin{bmatrix} -4 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$$

You can verify that these eigenvectors form an orthogonal set. By dividing each eigenvector by its magnitude, we obtain an orthonormal set:

$$\left\{ \frac{1}{\sqrt{18}} \begin{bmatrix} -4 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$$

Consider the following example.

Example 7.4.4: Repeated Eigenvalues

Find an orthonormal set of three eigenvectors for the matrix

$$A = \begin{bmatrix} 10 & 2 & 2 \\ 2 & 13 & 4 \\ 2 & 4 & 13 \end{bmatrix}$$

Solution

You can verify that the eigenvalues of A are 9 (with multiplicity two) and 18 (with multiplicity one). Consider the eigenvectors corresponding to $\lambda = 9$. The appropriate augmented matrix and reduced row-echelon form are given by

$$\left[\begin{array}{ccc|c} 9-10 & -2 & -2 & 0 \\ -2 & 9-13 & -4 & 0 \\ -2 & -4 & 9-13 & 0 \end{array} \right] \rightarrow \cdots \rightarrow \left[\begin{array}{ccc|c} 1 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

and so eigenvectors are of the form

$$\begin{bmatrix} -2y-2z \\ y \\ z \end{bmatrix}$$

We need to find two of these which are orthogonal. Let one be given by setting $z = 0$ and $y = 1$, giving $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$.

In order to find an eigenvector orthogonal to this one, we need to satisfy

$$\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} -2y-2z \\ y \\ z \end{bmatrix} = 5y + 4z = 0$$

The values $y = -4$ and $z = 5$ satisfy this equation, giving another eigenvector corresponding to $\lambda = 9$ as

$$\begin{bmatrix} -2(-4) - 2(5) \\ (-4) \\ 5 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 5 \end{bmatrix}$$

Next find the eigenvector for $\lambda = 18$. The augmented matrix and the resulting reduced row-echelon form are given by

$$\left[\begin{array}{ccc|c} 18-10 & -2 & -2 & 0 \\ -2 & 18-13 & -4 & 0 \\ -2 & -4 & 18-13 & 0 \end{array} \right] \rightarrow \cdots \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

and so an eigenvector is

$$\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$$

Dividing each eigenvector by its length, the orthonormal set is

$$\left\{ \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \frac{\sqrt{5}}{15} \begin{bmatrix} -2 \\ -4 \\ 5 \end{bmatrix}, \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \right\}$$

In the above solution, the repeated eigenvalue implies that there would have been many other orthonormal bases which could have been obtained. While we chose to take $z = 0, y = 1$, we could just as easily have taken $y = 0$ or even $y = z = 1$. Any such change would have resulted in a different orthonormal set.

Recall the following definition.

Definition 7.4.3: Diagonalizable

An $n \times n$ matrix A is said to be **non defective** or **diagonalizable** if there exists an invertible matrix P such that $P^{-1}AP = D$ where D is a diagonal matrix.

As indicated in Theorem 7.4.3 if A is a real symmetric matrix, there exists an orthogonal matrix U such that $U^T AU = D$ where D is a diagonal matrix. Therefore, every symmetric matrix is diagonalizable because if U is an orthogonal matrix, it is invertible and its inverse is U^T. In this case, we say that A is **orthogonally diagonalizable**. Therefore every symmetric matrix is in fact orthogonally diagonalizable. The next theorem provides another way to determine if a matrix is orthogonally diagonalizable.

Theorem 7.4.4: Orthogonally Diagonalizable

Let A be an $n \times n$ matrix. Then A is orthogonally diagonalizable if and only if A has an orthonormal set of eigenvectors.

Recall from Corollary 7.4.1 that every symmetric matrix has an orthonormal set of eigenvectors. In fact these three conditions are equivalent.

In the following example, the orthogonal matrix U will be found to orthogonally diagonalize a matrix.

Example 7.4.5: Diagonalize a Symmetric Matrix

Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{3}{2} \end{bmatrix}$. Find an orthogonal matrix U such that $U^T AU$ is a diagonal matrix.

Solution

In this case, the eigenvalues are 2 (with multiplicity one) and 1 (with multiplicity two). First we will find an eigenvector for the eigenvalue 2. The appropriate augmented matrix and resulting reduced row-echelon form are given by

$$\left[\begin{array}{ccc|c} 2-1 & 0 & 0 & 0 \\ 0 & 2-\frac{3}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 2-\frac{3}{2} & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and so an eigenvector is

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

However, it is desired that the eigenvectors be unit vectors and so dividing this vector by its length gives

$$\begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

Next find the eigenvectors corresponding to the eigenvalue equal to 1. The appropriate augmented matrix and resulting reduced row-echelon form are given by:

$$\left[\begin{array}{ccc|c} 1-1 & 0 & 0 & 0 \\ 0 & 1-\frac{3}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1-\frac{3}{2} & 0 \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{ccc|c} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore, the eigenvectors are of the form

$$\begin{bmatrix} s \\ -t \\ t \end{bmatrix}$$

Two of these which are orthonormal are $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, choosing $s = 1$ and $t = 0$, and $\begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$, letting $s = 0$, $t = 1$ and normalizing the resulting vector.

To obtain the desired orthogonal matrix, we let the orthonormal eigenvectors computed above be the columns.

$$\begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

To verify, compute $U^T AU$ as follows:

$$U^T AU = \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{3}{2} \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = D$$

the desired diagonal matrix. Notice that the eigenvectors, which construct the columns of U, are in the same order as the eigenvalues in D.

We conclude this section with a Theorem that generalizes earlier results.

Theorem 7.4.5: Triangulation of a Matrix

Let A be an $n \times n$ matrix. If A has n real eigenvalues, then an orthogonal matrix U can be found to result in the upper triangular matrix $U^T AU$.

triangulation

This Theorem provides a useful Corollary.

Corollary 7.4.2: Determinant and Trace

Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \cdots, \lambda_n$. Then it follows that $\det(A)$ is equal to the product of the λ_i, while $trace(A)$ is equal to the sum of the λ_i.

Proof

By Theorem 7.4.5 there exists an orthogonal matrix U such that $U^TAU = P$, where P is an upper triangular matrix. Since P is similar to A, the eigenvalues of P are $\lambda_1, \lambda_2, \ldots, \lambda_n$. Furthermore, since P is (upper) triangular, the entries on the main diagonal of P are its eigenvalues, so $\det(P) = \lambda_1\lambda_2\cdots\lambda_n$ and $trace(P) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$. Since P and A are similar, $\det(A) = \det(P)$ and $trace(A) = trace(P)$, and therefore the results follow.

The Singular Value Decomposition

We begin this section with an important definition.

Definition 7.4.4: Singular Values

Let A be an $m \times n$ matrix. The singular values of A are the square roots of the positive eigenvalues of A^TA.

Below is a video on singular values of a matrix.



Singular Value Decomposition (SVD) can be thought of as a generalization of orthogonal diagonalization of a symmetric matrix to an arbitrary $m \times n$ matrix. This decomposition is the focus of this section.

The following is a useful result that will help when computing the SVD of matrices.

Proposition 7.4.1: Same Nonzero Eigenvalues

Let A be an $m \times n$ matrix. Then A^TA and AA^T have the same **nonzero eigenvalues.**

Proof

Suppose A is an $m \times n$ matrix, and suppose that λ is a nonzero eigenvalue of A^TA. Then there exists a nonzero vector $X \in \mathbb{R}^n$ such that

$$(A^TA)X = \lambda X. \tag{7.4.1}$$

Multiplying both sides of this equation by A yields:

$$A(A^TA)X = A\lambda X$$
$$(AA^T)(AX) = \lambda(AX).$$

Since $\lambda \neq 0$ and $X \neq 0_n$, $\lambda X \neq 0_n$, and thus by equation (7.4.1), $(A^TA)X \neq 0_m$; thus $A^T(AX) \neq 0_m$, implying that $AX \neq 0_m$.

Therefore AX is an eigenvector of AA^T corresponding to eigenvalue λ. An analogous argument can be used to show that every nonzero eigenvalue of AA^T is an eigenvalue of A^TA, thus completing the proof.

Given an $m \times n$ matrix A, we will see how to express A as a product

$$A = U\Sigma V^T$$

where

- U is an $m \times m$ orthogonal matrix whose columns are eigenvectors of AA^T.
- V is an $n \times n$ orthogonal matrix whose columns are eigenvectors of A^TA.
- Σ is an $m \times n$ matrix whose only nonzero values lie on its main diagonal, and are the singular values of A.

How can we find such a decomposition? We are aiming to decompose A in the following form:

$$A = U \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix} V^T$$

where σ is of the form

$$\sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_k \end{bmatrix}$$

Thus $A^T = V \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix} U^T$ and it follows that

$$A^TA = V \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix} U^TU \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix} V^T = V \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix} V^T$$

and so $A^TAV = V \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix}$. Similarly, $AA^TU = U \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix}$. Therefore, you would find an orthonormal basis of eigenvectors for AA^T make them the columns of a matrix such that the corresponding eigenvalues are decreasing. This gives U. You could then do the same for A^TA to get V.

We formalize this discussion in the following theorem.

Theorem 7.4.6: Singular Value Decomposition

Let A be an $m \times n$ matrix. Then there exist orthogonal matrices U and V of the appropriate size such that $A = U\Sigma V^T$ where Σ is of the form

$$\Sigma = \begin{bmatrix} \sigma & 0 \\ 0 & 0 \end{bmatrix}$$

and σ is of the form

$$\sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_k \end{bmatrix}$$

for the σ_i the singular values of A.

Proof

There exists an orthonormal basis, $\{\vec{v}_i\}_{i=1}^n$ such that $A^TA\vec{v}_i = \sigma_i^2\vec{v}_i$ where $\sigma_i^2 > 0$ for $i = 1, \cdots, k, (\sigma_i > 0)$ and equals zero if $i > k$. Thus for $i > k$, $A\vec{v}_i = \vec{0}$ because

$$A\vec{v}_i \cdot A\vec{v}_i = A^TA\vec{v}_i \cdot \vec{v}_i = \vec{0} \cdot \vec{v}_i = 0.$$

For $i = 1, \cdots, k$, define $\vec{u}_i \in \mathbb{R}^m$ by

$$\vec{u}_i = \sigma_i^{-1} A\vec{v}_i.$$

Thus $A\vec{v}_i = \sigma_i\vec{u}_i$. Now

$$\begin{aligned}\vec{u}_i \cdot \vec{u}_j &= \sigma_i^{-1} A\vec{v}_i \cdot \sigma_j^{-1} A\vec{v}_j = \sigma_i^{-1}\vec{v}_i \cdot \sigma_j^{-1} A^T A\vec{v}_j \\ &= \sigma_i^{-1}\vec{v}_i \cdot \sigma_j^{-1}\sigma_j^2\vec{v}_j = \frac{\sigma_j}{\sigma_i}(\vec{v}_i \cdot \vec{v}_j) = \delta_{ij}.\end{aligned}$$

Thus $\{\vec{u}_i\}_{i=1}^k$ is an orthonormal set of vectors in $\mathbb{R}^m$. Also,

$$AA^T\vec{u}_i = AA^T\sigma_i^{-1}A\vec{v}_i = \sigma_i^{-1}AA^TA\vec{v}_i = \sigma_i^{-1}A\sigma_i^2\vec{v}_i = \sigma_i^2\vec{u}_i.$$

Now extend $\{\vec{u}_i\}_{i=1}^k$ to an orthonormal basis for all of $\mathbb{R}^m$, $\{\vec{u}_i\}_{i=1}^m$ and let

$$U = \begin{bmatrix}\vec{u}_1 & \cdots & \vec{u}_m\end{bmatrix}$$

while $V = (\vec{v}_1 \cdots \vec{v}_n)$. Thus U is the matrix which has the $\vec{u}_i$ as columns and V is defined as the matrix which has the $\vec{v}_i$ as columns. Then

$$U^TAV = \begin{bmatrix}\vec{u}_1^T \\ \vdots \\ \vec{u}_k^T \\ \vdots \\ \vec{u}_m^T\end{bmatrix} A\begin{bmatrix}\vec{v}_1 \cdots \vec{v}_n\end{bmatrix}$$

$$= \begin{bmatrix}\vec{u}_1^T \\ \vdots \\ \vec{u}_k^T \\ \vdots \\ \vec{u}_m^T\end{bmatrix}\begin{bmatrix}\sigma_1\vec{u}_1 & \cdots & \sigma_k\vec{u}_k & \vec{0} & \cdots & \vec{0}\end{bmatrix} = \begin{bmatrix}\sigma & 0 \\ 0 & 0\end{bmatrix}$$

where σ is given in the statement of the theorem.

The singular value decomposition has as an immediate corollary which is given in the following interesting result.

Corollary 7.4.3: Rank and Singular Values

Let A be an $m \times n$ matrix. Then the rank of A and A^T equals the number of singular values.

Let's compute the Singular Value Decomposition of a simple matrix.

Example 7.4.6: Singular Value Decomposition

Let $A = \begin{bmatrix}1 & -1 & 3 \\ 3 & 1 & 1\end{bmatrix}$. Find the Singular Value Decomposition (SVD) of A.

Solution

To begin, we compute AA^T and A^TA.

$$AA^T = \begin{bmatrix}1 & -1 & 3 \\ 3 & 1 & 1\end{bmatrix}\begin{bmatrix}1 & 3 \\ -1 & 1 \\ 3 & 1\end{bmatrix} = \begin{bmatrix}11 & 5 \\ 5 & 11\end{bmatrix}.$$

$$A^TA = \begin{bmatrix}1 & 3 \\ -1 & 1 \\ 3 & 1\end{bmatrix}\begin{bmatrix}1 & -1 & 3 \\ 3 & 1 & 1\end{bmatrix} = \begin{bmatrix}10 & 2 & 6 \\ 2 & 2 & -2 \\ 6 & -2 & 10\end{bmatrix}.$$

Since AA^T is 2×2 while A^TA is 3×3, and AA^T and A^TA have the same *nonzero* eigenvalues (by Proposition 7.4.1), we compute the characteristic polynomial $c_{AA^T}(x)$ (because it's easier to compute than $c_{A^TA}(x)$).

$$\begin{aligned} c_{AA^T}(x) &= \det(xI - AA^T) = \begin{vmatrix} x-11 & -5 \\ -5 & x-11 \end{vmatrix} \\ &= (x-11)^2 - 25 \\ &= x^2 - 22x + 121 - 25 \\ &= x^2 - 22x + 96 \\ &= (x-16)(x-6) \end{aligned}$$

Therefore, the eigenvalues of AA^T are $\lambda_1 = 16$ and $\lambda_2 = 6$.

The eigenvalues of A^TA are $\lambda_1 = 16$, $\lambda_2 = 6$, and $\lambda_3 = 0$, and the singular values of A are $\sigma_1 = \sqrt{16} = 4$ and $\sigma_2 = \sqrt{6}$. By convention, we list the eigenvalues (and corresponding singular values) in non increasing order (i.e., from largest to smallest).

To find the matrix V:

To construct the matrix V we need to find eigenvectors for A^TA. Since the eigenvalues of AA^T are distinct, the corresponding eigenvectors are orthogonal, and we need only normalize them.

$\lambda_1 = 16$: solve $(16I - A^TA)Y = 0$.

$$\left[\begin{array}{ccc|c} 6 & -2 & -6 & 0 \\ -2 & 14 & 2 & 0 \\ -6 & 2 & 6 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \text{ so } Y = \begin{bmatrix} t \\ 0 \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, t \in \mathbb{R}.$$

$\lambda_2 = 6$: solve $(6I - A^TA)Y = 0$.

$$\left[\begin{array}{ccc|c} -4 & -2 & -6 & 0 \\ -2 & 4 & 2 & 0 \\ -6 & 2 & -4 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \text{ so } Y = \begin{bmatrix} -s \\ -s \\ s \end{bmatrix} = s \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, s \in \mathbb{R}.$$

$\lambda_3 = 0$: solve $(-A^TA)Y = 0$.

$$\left[\begin{array}{ccc|c} -10 & -2 & -6 & 0 \\ -2 & -2 & 2 & 0 \\ -6 & 2 & -10 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right], \text{ so } Y = \begin{bmatrix} -r \\ 2r \\ r \end{bmatrix} = r \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, r \in \mathbb{R}.$$

Let

$$V_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, V_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, V_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Then

$$V = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & -\sqrt{2} & -1 \\ 0 & -\sqrt{2} & 2 \\ \sqrt{3} & \sqrt{2} & 1 \end{bmatrix}.$$

Also,

$$\Sigma = \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix},$$

and we use A, V^T, and Σ to find U.

Since V is orthogonal and $A = U\Sigma V^T$, it follows that $AV = U\Sigma$. Let $V = [\, V_1 \quad V_2 \quad V_3 \,]$, and let $U = [\, U_1 \quad U_2 \,]$, where U_1 and U_2 are the two columns of U.

Then we have

$$A\begin{bmatrix} V_1 & V_2 & V_3 \end{bmatrix} = \begin{bmatrix} U_1 & U_2 \end{bmatrix}\Sigma$$
$$\begin{bmatrix} AV_1 & AV_2 & AV_3 \end{bmatrix} = \begin{bmatrix} \sigma_1 U_1 + 0U_2 & 0U_1 + \sigma_2 U_2 & 0U_1 + 0U_2 \end{bmatrix}$$
$$= \begin{bmatrix} \sigma_1 U_1 & \sigma_2 U_2 & 0 \end{bmatrix}$$

which implies that $AV_1 = \sigma_1 U_1 = 4U_1$ and $AV_2 = \sigma_2 U_2 = \sqrt{6}U_2$.

Thus,

$$U_1 = \frac{1}{4}AV_1 = \frac{1}{4}\begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{4\sqrt{2}}\begin{bmatrix} 4 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and

$$U_2 = \frac{1}{\sqrt{6}}AV_2 = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix}\frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{3\sqrt{2}}\begin{bmatrix} 3 \\ -3 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Therefore,

$$U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

and

$$A = \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix}$$
$$= \left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right)\begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix}\left(\frac{1}{\sqrt{6}}\begin{bmatrix} \sqrt{3} & 0 & \sqrt{3} \\ -\sqrt{2} & -\sqrt{2} & \sqrt{2} \\ -1 & 2 & 1 \end{bmatrix}\right).$$

Below is a video on the singular value decomposition of a matrix.



Here is another example.

Example 7.4.7: Finding the SVD

Find an SVD for $A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$.

Solution

Since A is 3×1, $A^T A$ is a 1×1 matrix whose eigenvalues are easier to find than the eigenvalues of the 3×3 matrix AA^T.

$$A^TA = \begin{bmatrix} -1 & 2 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \end{bmatrix}.$$

Thus A^TA has eigenvalue $\lambda_1 = 9$, and the eigenvalues of AA^T are $\lambda_1 = 9$, $\lambda_2 = 0$, and $\lambda_3 = 0$. Furthermore, A has only one singular value, $\sigma_1 = 3$.

To find the matrix V: To do so we find an eigenvector for A^TA and normalize it. In this case, finding a unit eigenvector is trivial: $V_1 = \begin{bmatrix} 1 \end{bmatrix}$, and

$$V = \begin{bmatrix} 1 \end{bmatrix}.$$

Also, $\Sigma = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$, and we use A, V^T, and Σ to find U.

Now $AV = U\Sigma$, with $V = \begin{bmatrix} V_1 \end{bmatrix}$, and $U = \begin{bmatrix} U_1 & U_2 & U_3 \end{bmatrix}$, where U_1, U_2, and U_3 are the columns of U. Thus

$$\begin{aligned} A\begin{bmatrix} V_1 \end{bmatrix} &= \begin{bmatrix} U_1 & U_2 & U_3 \end{bmatrix} \Sigma \\ \begin{bmatrix} AV_1 \end{bmatrix} &= \begin{bmatrix} \sigma_1 U_1 + 0U_2 + 0U_3 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 U_1 \end{bmatrix} \end{aligned}$$

This gives us $AV_1 = \sigma_1 U_1 = 3U_1$, so

$$U_1 = \frac{1}{3} AV_1 = \frac{1}{3} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

The vectors U_2 and U_3 are eigenvectors of AA^T corresponding to the eigenvalue $\lambda_2 = \lambda_3 = 0$. Instead of solving the system $(0I - AA^T)X = 0$ and then using the Gram-Schmidt process on the resulting set of two basic eigenvectors, the following approach may be used.

Find vectors U_2 and U_3 by first extending $\{U_1\}$ to a basis of $\mathbb{R}^3$, then using the Gram-Schmidt algorithm to orthogonalize the basis, and finally normalizing the vectors.

Starting with $\{3U_1\}$ instead of $\{U_1\}$ makes the arithmetic a bit easier. It is easy to verify that

$$\left\{ \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

is a basis of $\mathbb{R}^3$. Set

$$E_1 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, X_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, X_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

and apply the Gram-Schmidt algorithm to $\{E_1, X_2, X_3\}$.

This gives us

$$E_2 = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} \text{ and } E_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Therefore,

$$U_2 = \frac{1}{\sqrt{18}} \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}, U_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

and

$$U = \begin{bmatrix} -\frac{1}{3} & \frac{4}{\sqrt{18}} & 0 \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Finally,

$$A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{4}{\sqrt{18}} & 0 \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} [1].$$

Consider another example.

Example 7.4.8: Find the SVD

Find a singular value decomposition for the matrix

$$A = \begin{bmatrix} \frac{2}{5}\sqrt{2}\sqrt{5} & \frac{4}{5}\sqrt{2}\sqrt{5} & 0 \\ \frac{2}{5}\sqrt{2}\sqrt{5} & \frac{4}{5}\sqrt{2}\sqrt{5} & 0 \end{bmatrix}$$

Solution

First consider $A^T A$

$$\begin{bmatrix} \frac{16}{5} & \frac{32}{5} & 0 \\ \frac{32}{5} & \frac{64}{5} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

What are some eigenvalues and eigenvectors? Some computing shows these are

$$\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{2}{5}\sqrt{5} \\ \frac{1}{5}\sqrt{5} \\ 0 \end{bmatrix} \right\} \leftrightarrow 0, \left\{ \begin{bmatrix} \frac{1}{5}\sqrt{5} \\ \frac{2}{5}\sqrt{5} \\ 0 \end{bmatrix} \right\} \leftrightarrow 16$$

Thus the matrix V is given by

$$V = \begin{bmatrix} \frac{1}{5}\sqrt{5} & -\frac{2}{5}\sqrt{5} & 0 \\ \frac{2}{5}\sqrt{5} & \frac{1}{5}\sqrt{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Next consider AA^T

$$\begin{bmatrix} 8 & 8 \\ 8 & 8 \end{bmatrix}$$

Eigenvectors and eigenvalues are

$$\left\{ \begin{bmatrix} -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix} \right\} \leftrightarrow 0, \left\{ \begin{bmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix} \right\} \leftrightarrow 16$$

Thus you can let U be given by

$$U = \begin{bmatrix} \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix}$$

Lets check this. $U^T AV =$

$$\begin{bmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \\ -\frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{2}{5}\sqrt{2}\sqrt{5} & \frac{4}{5}\sqrt{2}\sqrt{5} & 0 \\ \frac{2}{5}\sqrt{2}\sqrt{5} & \frac{4}{5}\sqrt{2}\sqrt{5} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{5}\sqrt{5} & -\frac{2}{5}\sqrt{5} & 0 \\ \frac{2}{5}\sqrt{5} & \frac{1}{5}\sqrt{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Below is a video on determining the fundamental subspaces of a matrix given its singular value decomposition.



This illustrates that if you have a good way to find the eigenvectors and eigenvalues for a Hermitian matrix which has nonnegative eigenvalues, then you also have a good way to find the singular value decomposition of an arbitrary matrix.

Positive Definite Matrices

Positive definite matrices are often encountered in applications such mechanics and statistics.

We begin with a definition.

Definition 7.4.5: Positive Definite Matrix

Let A be an $n \times n$ symmetric matrix. Then A is positive definite if all of its eigenvalues are positive.

The relationship between a negative definite matrix and positive definite matrix is as follows.

Lemma 7.4.2: Negative Definite Matrix

An $n \times n$ matrix A is negative definite if and only if $-A$ is positive definite

Consider the following lemma.

Lemma 7.4.3: Positive Definite Matrix and Invertibility

If A is positive definite, then it is invertible.

Proof

If $A\vec{v} = \vec{0}$, then 0 is an eigenvalue if $\vec{v}$ is nonzero, which does not happen for a positive definite matrix. Hence $\vec{v} = \vec{0}$ and so A is one to one. This is sufficient to conclude that it is invertible.

Notice that this lemma implies that if a matrix A is positive definite, then $\det(A) > 0$.

The following theorem provides another characterization of positive definite matrices. It gives a useful test for verifying if a matrix is positive definite.

Theorem 7.4.7: Positive Definite Matrix

Let A be a symmetric matrix. Then A is positive definite if and only if $\vec{x}^T A\vec{x}$ is positive for all nonzero $\vec{x} \in \mathbb{R}^n$.

Proof

Since A is symmetric, there exists an orthogonal matrix U so that

$$U^T AU = diag(\lambda_1, \lambda_2, \ldots, \lambda_n) = D,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the (not necessarily distinct) eigenvalues of A. Let $\vec{x} \in \mathbb{R}^n$, $\vec{x} \neq \vec{0}$, and define $\vec{y} = U^T\vec{x}$. Then

$$\vec{x}^T A\vec{x} = \vec{x}^T(UDU^T)\vec{x} = (\vec{x}^T U)D(U^T\vec{x}) = \vec{y}^T D\vec{y}.$$

Writing $\vec{y}^T = [\, y_1 \quad y_2 \quad \cdots \quad y_n \,]$,

$$\vec{x}^T A\vec{x} = [\, y_1 \quad y_2 \quad \cdots \quad y_n \,]\, diag(\lambda_1, \lambda_2, \ldots, \lambda_n) \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots \lambda_n y_n^2.$$

($\Rightarrow$) First we will assume that A is positive definite and prove that $\vec{x}^T A\vec{x}$ is positive.

Suppose A is positive definite, and $\vec{x} \in \mathbb{R}^n$, $\vec{x} \neq \vec{0}$. Since U^T is invertible, $\vec{y} = U^T\vec{x} \neq \vec{0}$, and thus $y_j \neq 0$ for some j, implying $y_j^2 > 0$ for some j. Furthermore, since all eigenvalues of A are positive, $\lambda_i y_i^2 \geq 0$ for all i and $\lambda_j y_j^2 > 0$. Therefore, $\vec{x}^T A\vec{x} > 0$.

($\Leftarrow$) Now we will assume $\vec{x}^T A\vec{x}$ is positive and show that A is positive definite.

If $\vec{x}^T A\vec{x} > 0$ whenever $\vec{x} \neq \vec{0}$, choose $\vec{x} = U\vec{e}_j$, where $\vec{e}_j$ is the jth column of I_n. Since U is invertible, $\vec{x} \neq \vec{0}$, and thus

$$\vec{y} = U^T\vec{x} = U^T(U\vec{e}_j) = \vec{e}_j.$$

Thus $y_j = 1$ and $y_i = 0$ when $i \neq j$, so

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots \lambda_n y_n^2 = \lambda_j,$$

i.e., $\lambda_j = \vec{x}^T A\vec{x} > 0$. Therefore, A is positive definite.

There are some other very interesting consequences which result from a matrix being positive definite. First one can note that the property of being positive definite is transferred to each of the principal submatrices which we will now define.

Definition 7.4.6: The Submatrix A_k

Let A be an $n \times n$ matrix. Denote by A_k the $k \times k$ matrix obtained by deleting the $k+1, \cdots, n$ columns and the $k+1, \cdots, n$ rows from A. Thus $A_n = A$ and A_k is the $k \times k$ submatrix of A which occupies the upper left corner of A.

Lemma 7.4.4: Positive Definite and Submatrices

Let A be an $n \times n$ positive definite matrix. Then each submatrix A_k is also positive definite.

Proof

This follows right away from the above definition. Let $\vec{x} \in \mathbb{R}^k$ be nonzero. Then

$$\vec{x}^T A_k\vec{x} = [\, \vec{x}^T \quad 0 \,]\, A \begin{bmatrix} \vec{x} \\ 0 \end{bmatrix} > 0$$

by the assumption that A is positive definite.

There is yet another way to recognize whether a matrix is positive definite which is described in terms of these submatrices. We state the result, the proof of which can be found in more advanced texts.

Theorem 7.4.8: Positive Matrix and Determinant of A_k

Let A be a symmetric matrix. Then A is positive definite if and only if $\det(A_k)$ is greater than 0 for every submatrix A_k, $k = 1, \cdots, n$.

Proof

We prove this theorem by induction on n. It is clearly true if $n = 1$. Suppose then that it is true for $n-1$ where $n \geq 2$. Since $\det(A) = \det(A_n) > 0$, it follows that all the eigenvalues are nonzero. We need to show that they are all positive. Suppose not. Then there is some even number of them which are negative, even because the product of all the eigenvalues is known to be positive, equaling $\det(A)$. Pick two, λ_1 and λ_2 and let $A\vec{u}_i = \lambda_i \vec{u}_i$ where $\vec{u}_i \neq \vec{0}$ for $i = 1, 2$ and $\vec{u}_1 \cdot \vec{u}_2 = 0$. Now if $\vec{y} \equiv \alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2$ is an element of $span\{\vec{u}_1, \vec{u}_2\}$, then since these are eigenvalues and $\vec{u}_1 \cdot \vec{u}_2 = 0$, a short computation shows

$$\begin{aligned} &(\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2)^T A (\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2) \\ &= |\alpha_1|^2 \lambda_1 \vec{u}_1^2 + |\alpha_2|^2 \lambda_2 \vec{u}_2^2 < 0. \end{aligned}$$

Now letting $\vec{x} \in \mathbb{R}^{n-1}$, we can use the induction hypothesis to write

$$\begin{bmatrix} x^T & 0 \end{bmatrix} A \begin{bmatrix} \vec{x} \\ 0 \end{bmatrix} = \vec{x}^T A_{n-1} \vec{x} > 0.$$

Now the dimension of $\{\vec{z} \in \mathbb{R}^n : z_n = 0\}$ is $n-1$ and the dimension of $span\{\vec{u}_1, \vec{u}_2\} = 2$ and so there must be some nonzero $\vec{x} \in \mathbb{R}^n$ which is in both of these subspaces of $\mathbb{R}^n$. However, the first computation would require that $\vec{x}^T A \vec{x} < 0$ while the second would require that $\vec{x}^T A \vec{x} > 0$. This contradiction shows that all the eigenvalues must be positive. This proves the if part of the theorem. The converse can also be shown to be correct, but it is the direction which was just shown which is of most interest.

Corollary 7.4.4: Symmetric and Negative Definite Matrix

Let A be symmetric. Then A is negative definite if and only if

$$(-1)^k \det(A_k) > 0$$

for every $k = 1, \cdots, n$.

Proof

This is immediate from the above theorem when we notice, that A is negative definite if and only if $-A$ is positive definite. Therefore, if $\det(-A_k) > 0$ for all $k = 1, \cdots, n$, it follows that A is negative definite. However, $\det(-A_k) = (-1)^k \det(A_k)$.

The Cholesky Factorization

Another important theorem is the existence of a specific factorization of positive definite matrices. It is called the Cholesky Factorization and factors the matrix into the product of an upper triangular matrix and its transpose.

Theorem 7.4.9: Cholesky Factorization

Let A be a positive definite matrix. Then there exists an upper triangular matrix U whose main diagonal entries are positive, such that A can be written

$$A = U^T U$$

This factorization is unique.

The process for finding such a matrix U relies on simple row operations.

Procedure 7.4.1: Finding the Cholesky Factorization

Let A be a positive definite matrix. The matrix U that creates the Cholesky Factorization can be found through two steps.

1. Using only type 3 elementary row operations (multiples of rows added to other rows) put A in upper triangular form. Call this matrix $\hat{U}$. Then $\hat{U}$ has positive entries on the main diagonal.
2. Divide each row of $\hat{U}$ by the square root of the diagonal entry in that row. The result is the matrix U.

Of course you can always verify that your factorization is correct by multiplying U and U^T to ensure the result is the original matrix A.

Consider the following example.

Example 7.4.9: Cholesky Factorization

Show that $A = \begin{bmatrix} 9 & -6 & 3 \\ -6 & 5 & -3 \\ 3 & -3 & 6 \end{bmatrix}$ is positive definite, and find the Cholesky factorization of A.

Solution

First we show that A is positive definite. By Theorem 7.4.8 it suffices to show that the determinant of each submatrix is positive.

$$A_1 = [\,9\,] \text{ and } A_2 = \begin{bmatrix} 9 & -6 \\ -6 & 5 \end{bmatrix},$$

so $\det(A_1) = 9$ and $\det(A_2) = 9$. Since $\det(A) = 36$, it follows that A is positive definite.

Now we use Procedure 7.4.1 to find the Cholesky Factorization. Row reduce (using only type 3 row operations) until an upper triangular matrix is obtained.

$$\begin{bmatrix} 9 & -6 & 3 \\ -6 & 5 & -3 \\ 3 & -3 & 6 \end{bmatrix} \to \begin{bmatrix} 9 & -6 & 3 \\ 0 & 1 & -1 \\ 0 & -1 & 5 \end{bmatrix} \to \begin{bmatrix} 9 & -6 & 3 \\ 0 & 1 & -1 \\ 0 & 0 & 4 \end{bmatrix}$$

Now divide the entries in each row by the square root of the diagonal entry in that row, to give

$$U = \begin{bmatrix} 3 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{bmatrix}$$

You can verify that $U^T U = A$.

Example 7.4.10: Cholesky Factorization

Let A be a positive definite matrix given by

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 4 & 2 \\ 1 & 2 & 5 \end{bmatrix}$$

Determine its Cholesky factorization.

Solution

You can verify that A is in fact positive definite.

To find the Cholesky factorization we first row reduce to an upper triangular matrix.

$$\begin{bmatrix} 3 & 1 & 1 \\ 1 & 4 & 2 \\ 1 & 2 & 5 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 1 & 1 \\ 0 & \frac{11}{3} & \frac{5}{3} \\ 0 & \frac{5}{3} & \frac{14}{5} \end{bmatrix} \rightarrow \begin{bmatrix} 3 & 1 & 1 \\ 0 & \frac{11}{3} & \frac{5}{3} \\ 0 & 0 & \frac{43}{11} \end{bmatrix}$$

Now divide the entries in each row by the square root of the diagonal entry in that row and simplify.

$$U = \begin{bmatrix} \sqrt{3} & \frac{1}{3}\sqrt{3} & \frac{1}{3}\sqrt{3} \\ 0 & \frac{1}{3}\sqrt{3}\sqrt{11} & \frac{5}{33}\sqrt{3}\sqrt{11} \\ 0 & 0 & \frac{1}{11}\sqrt{11}\sqrt{43} \end{bmatrix}$$

QR Factorization

In this section, a reliable factorization of matrices is studied. Called the QR factorization of a matrix, it *always* exists. While much can be said about the QR factorization, this section will be limited to real matrices. Therefore we assume the dot product used below is the usual dot product. We begin with a definition.

Definition 7.4.7: QR Factorization

Let A be a real $m \times n$ matrix. Then a QR factorization of A consists of two matrices, Q orthogonal and R upper triangular, such that $A = QR$.

qrfactorization

The following theorem claims that such a factorization exists.

Theorem 7.4.10: Existence of QR Factorization

Let A be any real $m \times n$ matrix with linearly independent columns. Then there exists an orthogonal matrix Q and an upper triangular matrix R having non-negative entries on the main diagonal such that

$$A = QR$$

The procedure for obtaining the QR factorization for any matrix A is as follows.

Procedure 7.4.2: QR Factorization

Let A be an $m \times n$ matrix given by $A = \begin{bmatrix} A_1 & A_2 & \cdots & A_n \end{bmatrix}$ where the A_i are the linearly independent columns of A.

1. Apply the Gram-Schmidt Process 4.11.1 to the columns of A, writing B_i for the resulting columns.
2. Normalize the B_i, to find $C_i = \frac{1}{B_i} B_i$.
3. Construct the orthogonal matrix Q as $Q = \begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix}$.
4. Construct the upper triangular matrix R as

$$R = \begin{bmatrix} B_1 & A_2 \cdot C_1 & A_3 \cdot C_1 & \cdots & A_n \cdot C_1 \\ 0 & B_2 & A_3 \cdot C_2 & \cdots & A_n \cdot C_2 \\ 0 & 0 & B_3 & \cdots & A_n \cdot C_3 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & B_n \end{bmatrix}$$

5. Finally, write $A = QR$ where Q is the orthogonal matrix and R is the upper triangular matrix obtained above.

Notice that Q is an orthogonal matrix as the C_i form an orthonormal set. Since $B_i > 0$ for all i (since the length of a vector is always positive), it follows that R is an upper triangular matrix with positive entries on the main diagonal.

Consider the following example.

Example 7.4.11: Finding a QR Factorization

Let

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Find an orthogonal matrix Q and upper triangular matrix R such that $A = QR$.

Solution

First, observe that A_1, A_2, the columns of A, are linearly independent. Therefore we can use the Gram-Schmidt Process to create a corresponding orthogonal set $\{B_1, B_2\}$ as follows:

$$\begin{aligned} B_1 &= A_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\ B_2 &= A_2 - \frac{A_2 \cdot B_1}{B_1^2} B_1 \\ &= \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \end{aligned}$$

Normalize each vector to create the set $\{C_1, C_2\}$ as follows:

$$\begin{aligned} C_1 &= \frac{1}{B_1} B_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\ C_2 &= \frac{1}{B_2} B_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \end{aligned}$$

Now construct the orthogonal matrix Q as

$$\begin{aligned} Q &= \begin{bmatrix} C_1 & C_2 & \cdots & C_n \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{bmatrix} \end{aligned}$$

Finally, construct the upper triangular matrix R as

$$\begin{aligned} R &= \begin{bmatrix} B_1 & A_2 \cdot C_1 \\ 0 & B_2 \end{bmatrix} \\ &= \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & \sqrt{3} \end{bmatrix} \end{aligned}$$

It is left to the reader to verify that $A = QR$.

Below is a video on the QR decomposition of a matrix.



The QR Factorization and Eigenvalues

The QR factorization of a matrix has a very useful application. It turns out that it can be used repeatedly to estimate the eigenvalues of a matrix. Consider the following procedure.

Procedure 7.4.3: Using the QR Factorization to Estimate Eigenvalues

Let A be an invertible matrix. Define the matrices $A_1, A_2, \cdots$ as follows:

1. $A_1 = A$ factored as $A_1 = Q_1 R_1$
2. $A_2 = R_1 Q_1$ factored as $A_2 = Q_2 R_2$
3. $A_3 = R_2 Q_2$ factored as $A_3 = Q_3 R_3$

Continue in this manner, where in general $A_k = Q_k R_k$ and $A_{k+1} = R_k Q_k$.

Then it follows that this sequence of A_i converges to an upper triangular matrix which is similar to A. Therefore the eigenvalues of A can be approximated by the entries on the main diagonal of this upper triangular matrix.

Below is another video on the QR decomposition of a matrix.



Power Methods

While the QR algorithm can be used to compute eigenvalues, there is a useful and fairly elementary technique for finding the eigenvector and associated eigenvalue nearest to a given complex number which is called the *shifted inverse power method*. It tends to work extremely well provided you start with something which is fairly close to an eigenvalue.

Power methods are based the consideration of powers of a given matrix. Let $\{\vec{x}_1, \cdots, \vec{x}_n\}$ be a basis of eigenvectors for $\mathbb{C}^n$ such that $A\vec{x}_n = \lambda_n \vec{x}_n$. Now let $\vec{u}_1$ be some nonzero vector. Since $\{\vec{x}_1, \cdots, \vec{x}_n\}$ is a basis, there exists unique scalars, c_i such that

$$\vec{u}_1 = \sum_{k=1}^{n} c_k \vec{x}_k$$

Assume you have not been so unlucky as to pick $\vec{u}_1$ in such a way that $c_n = 0$. Then let $A\vec{u}_k = \vec{u}_{k+1}$ so that

$$\vec{u}_m = A^m \vec{u}_1 = \sum_{k=1}^{n-1} c_k \lambda_k^m \vec{x}_k + \lambda_n^m c_n \vec{x}_n. \tag{7.4.2}$$

For large m the last term, $\lambda_n^m c_n \vec{x}_n$, determines quite well the direction of the vector on the right. This is because $|\lambda_n|$ is larger than $|\lambda_k|$ for $k < n$ and so for a large m, the sum, $\sum_{k=1}^{n-1} c_k \lambda_k^m \vec{x}_k$, on the right is fairly insignificant. Therefore, for large m, $\vec{u}_m$ is essentially a multiple of the eigenvector $\vec{x}_n$, the one which goes with λ_n. The only problem is that there is no control of the size of the vectors $\vec{u}_m$. You can fix this by scaling. Let S_2 denote the entry of $A\vec{u}_1$ which is largest in absolute value. We call this a **scaling factor**. Then $\vec{u}_2$ will not be just $A\vec{u}_1$ but $A\vec{u}_1/S_2$. Next let S_3 denote the entry of $A\vec{u}_2$ which has largest absolute value and define $\vec{u}_3 \equiv A\vec{u}_2/S_3$. Continue this way. The scaling just described does not destroy the relative insignificance of the term involving a sum in (7.4.2). Indeed it amounts to nothing more than changing the units of length. Also note that from this scaling procedure, the absolute value of the largest element of $\vec{u}_k$ is always equal to 1. Therefore, for large m,

$$\vec{u}_m = \frac{\lambda_n^m c_n \vec{x}_n}{S_2 S_3 \cdots S_m} + (\text{relatively insignificant term}).$$

Therefore, the entry of $A\vec{u}_m$ which has the largest absolute value is essentially equal to the entry having largest absolute value of

$$A\left(\frac{\lambda_n^m c_n \vec{x}_n}{S_2 S_3 \cdots S_m}\right) = \frac{\lambda_n^{m+1} c_n \vec{x}_n}{S_2 S_3 \cdots S_m} \approx \lambda_n \vec{u}_m$$

and so for large m, it must be the case that $\lambda_n \approx S_{m+1}$. This suggests the following procedure.

Procedure 7.4.4: Finding the Largest Eigenvalue with its Eigenvector

1. Start with a vector $\vec{u}_1$ which you hope has a component in the direction of $\vec{x}_n$. The vector $(1, \cdots, 1)^T$ is usually a pretty good choice.
2. If $\vec{u}_k$ is known,
$$\vec{u}_{k+1} = \frac{A\vec{u}_k}{S_{k+1}}$$
where S_{k+1} is the entry of $A\vec{u}_k$ which has largest absolute value.
3. When the scaling factors, S_k are not changing much, S_{k+1} will be close to the eigenvalue and $\vec{u}_{k+1}$ will be close to an eigenvector.
4. Check your answer to see if it worked well.

The shifted inverse power method involves finding the eigenvalue closest to a given complex number along with the associated eigenvalue. If μ is a complex number and you want to find λ which is closest to μ, you could consider the eigenvalues and eigenvectors of $(A - \mu I)^{-1}$. Then $A\vec{x} = \lambda\vec{x}$ if and only if

$$(A - \mu I)\vec{x} = (\lambda - \mu)\vec{x}$$

If and only if

$$\frac{1}{\lambda - \mu}\vec{x} = (A - \mu I)^{-1}\vec{x}$$

Thus, if λ is the closest eigenvalue of A to μ then out of all eigenvalues of $(A - \mu I)^{-1}$, you would have $\frac{1}{\lambda - \mu}$ would be the largest. Thus all you have to do is apply the power method to $(A - \mu I)^{-1}$ and the eigenvector you get will be the eigenvector which corresponds to λ where λ is the closest to μ of all eigenvalues of A. You could use the eigenvector to determine this directly.

Example 7.4.12: Finding Eigenvalue and Eigenvector

Find the eigenvalue and eigenvector for

$$\begin{bmatrix} 3 & 2 & 1 \\ -2 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix}$$

which is closest to $.9+.9i$.

Solution

$$\left(\begin{bmatrix} 3 & 2 & 1 \\ -2 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix} - (.9+.9i) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)^{-1}$$

$$= \begin{bmatrix} -0.619\,19 - 10.545i & -5.524\,9 - 4.972\,4i & -0.370\,57 - 5.821\,3i \\ 5.524\,9 + 4.972\,4i & 5.276\,2 + 0.248\,62i & 2.762\,4 + 2.486\,2i \\ 0.741\,14 + 11.643i & 5.524\,9 + 4.972\,4i & 0.492\,52 + 6.918\,9i \end{bmatrix}$$

Then pick an initial guess an multiply by this matrix raised to a large power.

$$= \begin{bmatrix} -0.619\,19 - 10.545i & -5.524\,9 - 4.972\,4i & -0.370\,57 - 5.821\,3i \\ 5.524\,9 + 4.972\,4i & 5.276\,2 + 0.248\,62i & 2.762\,4 + 2.486\,2i \\ 0.741\,14 + 11.643i & 5.524\,9 + 4.972\,4i & 0.492\,52 + 6.918\,9i \end{bmatrix}^{15} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

This equals

$$\begin{bmatrix} 1.562\,9 \times 10^{13} - 3.899\,3 \times 10^{12} i \\ -5.864\,5 \times 10^{12} + 9.764\,2 \times 10^{12} i \\ -1.562\,9 \times 10^{13} + 3.899\,9 \times 10^{12} i \end{bmatrix}$$

Now divide by an entry to make the vector have reasonable size. This yields

$$\begin{bmatrix} -0.999\,99 - 3.614\,0 \times 10^{-5} i \\ 0.499\,99 - 0.499\,99 i \\ 1.0 \end{bmatrix}$$

which is close to

$$\begin{bmatrix} -1 \\ 0.5 - 0.5i \\ 1.0 \end{bmatrix}$$

Then

$$\begin{bmatrix} 3 & 2 & 1 \\ -2 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0.5 - 0.5i \\ 1.0 \end{bmatrix} = \begin{bmatrix} -1.0 - 1.0i \\ 1.0 \\ 1.0 + 1.0i \end{bmatrix}$$

Now to determine the eigenvalue, you could just take the ratio of corresponding entries. Pick the two corresponding entries which have the largest absolute values. In this case, you would get the eigenvalue is $1+i$ which happens to be the exact eigenvalue. Thus an eigenvector and eigenvalue are

$$\begin{bmatrix} -1 \\ 0.5 - 0.5i \\ 1.0 \end{bmatrix}, 1+i$$

Usually it won't work out so well but you can still find what is desired. Thus, once you have obtained approximate eigenvalues using the QR algorithm, you can find the eigenvalue more exactly along with an eigenvector associated with it by using the shifted inverse power method.

Quadratic Forms

One of the applications of orthogonal diagonalization is that of quadratic forms and graphs of level curves of a quadratic form. This section has to do with rotation of axes so that with respect to the new axes, the graph of the level curve of a quadratic form is oriented parallel to the coordinate axes. This makes it much easier to understand. For example, we all know that $x_1^2+x_2^2=1$ represents the equation in two variables whose graph in $\mathbb{R}^2$ is a circle of radius 1. But how do we know what the graph of the equation $5x_1^2+4x_1x_2+3x_2^2=1$ represents?

We first formally define what is meant by a quadratic form. In this section we will work with only *real* quadratic forms, which means that the coefficients will all be real numbers.

Definition 7.4.8: Quadratic Form

A **quadratic form** is a polynomial of degree two in n variables $x_1, x_2, \cdots, x_n$, written as a linear combination of x_i^2 terms and x_ix_j terms.

Consider the quadratic form $q=a_{11}x_1^2+a_{22}x_2^2+\cdots+a_{nn}x_n^2+a_{12}x_1x_2+\cdots$. We can write $\vec{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ as the vector whose entries are the variables contained in the quadratic form.

Similarly, let $A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$ be the matrix whose entries are the coefficients of x_i^2 and x_ix_j from q. Note that the matrix A is not unique, and we will consider this further in the example below. Using this matrix A, the quadratic form can be written as $q=\vec{x}^TA\vec{x}$.

$$\begin{aligned} q &= \vec{x}^TA\vec{x} \\ &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11}x_1+a_{21}x_2+\cdots+a_{n1}x_n \\ a_{12}x_1+a_{22}x_2+\cdots+a_{n2}x_n \\ \vdots \\ a_{1n}x_1+a_{2n}x_2+\cdots+a_{nn}x_n \end{bmatrix} \\ &= a_{11}x_1^2+a_{22}x_2^2+\cdots+a_{nn}x_n^2+a_{12}x_1x_2+\cdots \end{aligned}$$

Let's explore how to find this matrix A. Consider the following example.

Example 7.4.13: Matrix of a Quadratic Form

Let a quadratic form q be given by

$$q=6x_1^2+4x_1x_2+3x_2^2$$

Write q in the form $\vec{x}^TA\vec{x}$.

Solution

First, let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.

Then, writing $q = \vec{x}^T A \vec{x}$ gives

$$\begin{aligned} q &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= a_{11}x_1^2 + a_{21}x_1x_2 + a_{12}x_1x_2 + a_{22}x_2^2 \end{aligned}$$

Notice that we have an x_1x_2 term as well as an x_2x_1 term. Since multiplication is commutative, these terms can be combined. This means that q can be written

$$q = a_{11}x_1^2 + (a_{21} + a_{12})\, x_1x_2 + a_{22}x_2^2$$

Equating this to q as given in the example, we have

$$a_{11}x_1^2 + (a_{21} + a_{12})\, x_1x_2 + a_{22}x_2^2 = 6x_1^2 + 4x_1x_2 + 3x_2^2$$

Therefore,

$$\begin{aligned} a_{11} &= 6 \\ a_{22} &= 3 \\ a_{21} + a_{12} &= 4 \end{aligned}$$

This demonstrates that the matrix A is not unique, as there are several correct solutions to $a_{21} + a_{12} = 4$. However, we will *always* choose the coefficients such that $a_{21} = a_{12} = \frac{1}{2}(a_{21} + a_{12})$. This results in $a_{21} = a_{12} = 2$. This choice is key, as it will ensure that A turns out to be a symmetric matrix.

Hence,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}$$

You can verify that $q = \vec{x}^T A \vec{x}$ holds for this choice of A.

The above procedure for choosing A to be symmetric applies for any quadratic form q. We will *always* choose coefficients such that $a_{ij} = a_{ji}$.

We now turn our attention to the focus of this section. Our goal is to start with a quadratic form q as given above and find a way to rewrite it to eliminate the x_ix_j terms. This is done through a change of variables. In other words, we wish to find y_i such that

$$q = d_{11}y_1^2 + d_{22}y_2^2 + \cdots + d_{nn}y_n^2$$

Letting $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ and $D = [d_{ij}]$, we can write $q = \vec{y}^T D \vec{y}$ where D is the matrix of coefficients from q. There is something special about this matrix D that is crucial. Since no y_iy_j terms exist in q, it follows that $d_{ij} = 0$ for all $i \neq j$. Therefore, D is a diagonal matrix. Through this change of variables, we find the **principal axes** $y_1, y_2, \cdots, y_n$ of the quadratic form.

This discussion sets the stage for the following essential theorem.

Theorem 7.4.11: Diagonalizing a Quadratic Form

Let q be a quadratic form in the variables $x_1, \cdots, x_n$. It follows that q can be written in the form $q = \vec{x}^T A \vec{x}$ where

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

and $A = [a_{ij}]$ is the symmetric matrix of coefficients of q.

New variables $y_1, y_2, \cdots, y_n$ can be found such that $q = \vec{y}^T D\vec{y}$ where

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

and $D = [d_{ij}]$ is a diagonal matrix. The matrix D contains the eigenvalues of A and is found by orthogonally diagonalizing A.

While not a formal proof, the following discussion should convince you that the above theorem holds. Let q be a quadratic form in the variables $x_1, \cdots, x_n$. Then, q can be written in the form $q = \vec{x}^T A\vec{x}$ for a symmetric matrix A. By Theorem 7.4.3 we can orthogonally diagonalize the matrix A such that $U^T AU = D$ for an orthogonal matrix U and diagonal matrix D.

Then, the vector $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ is found by $\vec{y} = U^T\vec{x}$. To see that this works, rewrite $\vec{y} = U^T\vec{x}$ as $\vec{x} = U\vec{y}$. Letting $q = \vec{x}^T A\vec{x}$, proceed as follows:

$$\begin{aligned} q &= \vec{x}^T A\vec{x} \\ &= (U\vec{y})^T A(U\vec{y}) \\ &= \vec{y}^T (U^T AU)\vec{y} \\ &= \vec{y}^T D\vec{y} \end{aligned}$$

The following procedure details the steps for the change of variables given in the above theorem.

Procedure 7.4.5: Diagonalizing a Quadratic Form

Let q be a quadratic form in the variables $x_1, \cdots, x_n$ given by

$$q = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + a_{12}x_1x_2 + \cdots$$

Then, q can be written as $q = d_{11}y_1^2 + \cdots + d_{nn}y_n^2$ as follows:

1. Write $q = \vec{x}^T A\vec{x}$ for a symmetric matrix A.
2. Orthogonally diagonalize A to be written as $U^T AU = D$ for an orthogonal matrix U and diagonal matrix D.
3. Write $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$. Then, $\vec{x} = U\vec{y}$.
4. The quadratic form q will now be given by

$$q = d_{11}y_1^2 + \cdots + d_{nn}y_n^2 = \vec{y}^T D\vec{y}$$

where $D = [d_{ij}]$ is the diagonal matrix found by orthogonally diagonalizing A.

Consider the following example.

Example 7.4.14: Choosing New Axes to Simplify a Quadratic Form

Consider the following level curve

$$6x_1^2 + 4x_1x_2 + 3x_2^2 = 7$$

shown in the following graph.

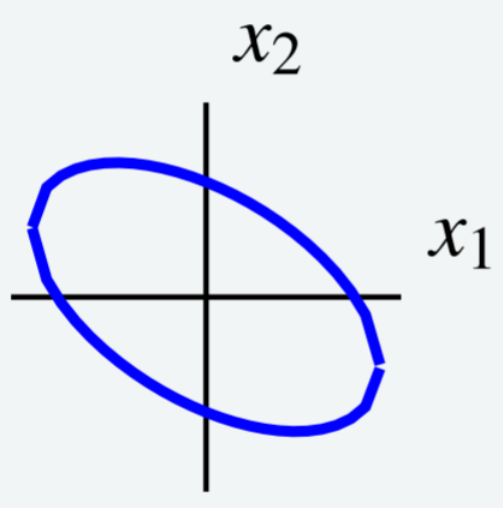


Figure 7.4.1

Use a change of variables to choose new axes such that the ellipse is oriented parallel to the new coordinate axes. In other words, use a change of variables to rewrite q to eliminate the x_1x_2 term.

Solution

Notice that the level curve is given by $q = 7$ for $q = 6x_1^2 + 4x_1x_2 + 3x_2^2$. This is the same quadratic form that we examined earlier in Example 7.4.13. Therefore we know that we can write $q = \vec{x}^T A\vec{x}$ for the matrix

$$A = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}$$

Now we want to orthogonally diagonalize A to write $U^T AU = D$ for an orthogonal matrix U and diagonal matrix D. The details are left to the reader, and you can verify that the resulting matrices are

$$U = \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$
$$D = \begin{bmatrix} 7 & 0 \\ 0 & 2 \end{bmatrix}$$

Next we write $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. It follows that $\vec{x} = U\vec{y}$.

We can now express the quadratic form q in terms of y, using the entries from D as coefficients as follows:

$$\begin{aligned} q &= d_{11}y_1^2 + d_{22}y_2^2 \\ &= 7y_1^2 + 2y_2^2 \end{aligned}$$

Hence the level curve can be written $7y_1^2 + 2y_2^2 = 7$. The graph of this equation is given by:

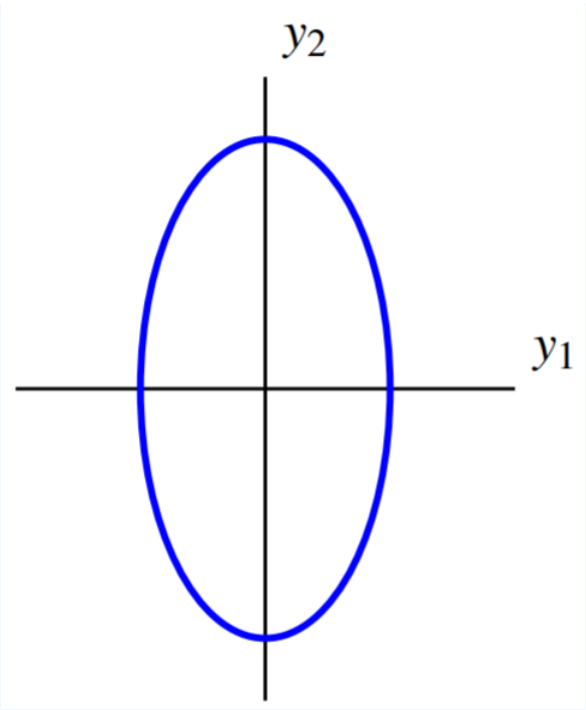


Figure 7.4.2

The change of variables results in new axes such that with respect to the new axes, the ellipse is oriented parallel to the coordinate axes. These are called the **principal axes** of the quadratic form.

The following is another example of diagonalizing a quadratic form.

Example 7.4.15: Choosing New Axes to Simplify a Quadratic Form

Consider the level curve

$$5x_1^2 - 6x_1x_2 + 5x_2^2 = 8$$

shown in the following graph.

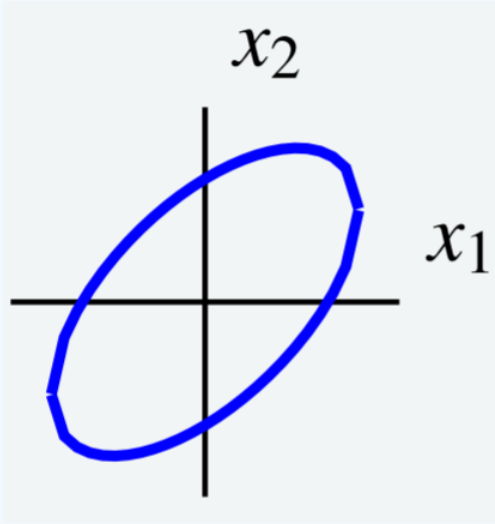


Figure 7.4.3

Use a change of variables to choose new axes such that the ellipse is oriented parallel to the new coordinate axes. In other words, use a change of variables to rewrite q to eliminate the x_1x_2 term.

Solution

First, express the level curve as $\vec{x}^T A\vec{x}$ where $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and A is symmetric. Let $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. Then $q = \vec{x}^T A\vec{x}$ is given by

$$\begin{aligned} q &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + a_{22}x_2^2 \end{aligned}$$

Equating this to the given description for q, we have

$$5x_1^2 - 6x_1x_2 + 5x_2^2 = a_{11}x_1^2 + (a_{12} + a_{21})x_1x_2 + a_{22}x_2^2$$

This implies that $a_{11} = 5, a_{22} = 5$ and in order for A to be symmetric, $a_{12} = a_{22} = \frac{1}{2}(a_{12} + a_{21}) = -3$. The result is $A = \begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}$. We can write $q = \vec{x}^T A\vec{x}$ as

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 8$$

Next, orthogonally diagonalize the matrix A to write $U^T AU = D$. The details are left to the reader and the necessary matrices are given by

$$\begin{aligned} U &= \begin{bmatrix} \frac{1}{2}\sqrt{2} & \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} & -\frac{1}{2}\sqrt{2} \end{bmatrix} \\ D &= \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \end{aligned}$$

Write $\vec{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, such that $\vec{x} = U\vec{y}$. Then it follows that q is given by

$$\begin{aligned} q &= d_{11}y_1^2 + d_{22}y_2^2 \\ &= 2y_1^2 + 8y_2^2 \end{aligned}$$

Therefore the level curve can be written as $2y_1^2 + 8y_2^2 = 8$.

This is an ellipse which is parallel to the coordinate axes. Its graph is of the form

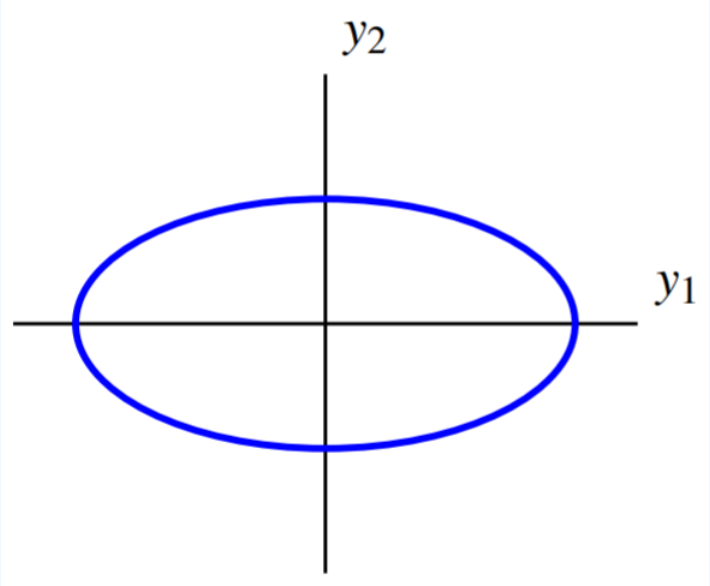


Figure 7.4.4

Thus this change of variables chooses new axes such that with respect to these new axes, the ellipse is oriented parallel to the coordinate axes.

This page titled 7.4: Orthogonality is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **7.4: Orthogonality** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

7.E: Exercises

Exercise 7.E.1

If A is an invertible $n \times n$ matrix, compare the eigenvalues of A and A^{-1}. More generally, for m an arbitrary integer, compare the eigenvalues of A and A^m.

Answer

$A^m X = \lambda^m X$ for any integer. In the case of -1, $A^{-1}\lambda X = AA^{-1}X = X$ so $A^{-1}X = \lambda^{-1}X$. Thus the eigenvalues of A^{-1} are just λ^{-1} where λ is an eigenvalue of A.

Exercise 7.E.2

If A is an $n \times n$ matrix and c is a nonzero constant, compare the eigenvalues of A and cA.

Answer

Say $AX = \lambda X$. Then $cAX = c\lambda X$ and so the eigenvalues of cA are just $c\lambda$ where λ is an eigenvalue of A.

Exercise 7.E.3

Let A, B be invertible $n \times n$ matrices which commute. That is, $AB = BA$. Suppose X is an eigenvector of B. Show that then AX must also be an eigenvector for B.

Answer

$BAX = ABX = A\lambda X = \lambda AX$. Here it is assumed that $BX = \lambda X$.

Exercise 7.E.4

Suppose A is an $n \times n$ matrix and it satisfies $A^m = A$ for some m a positive integer larger than 1. Show that if λ is an eigenvalue of A then $|\lambda|$ equals either 0 or 1.

Answer

Let X be the eigenvector. Then $A^m X = \lambda^m X$, $A^m X = AX = \lambda X$ and so

$$\lambda^m = \lambda$$

Hence if $\lambda \neq 0$, then

$$\lambda^{m-1} = 1$$

and so $|\lambda| = 1$.

Exercise 7.E.5

Show that if $AX = \lambda X$ and $AY = \lambda Y$, then whenever k, p are scalars,

$$A(kX + pY) = \lambda(kX + pY)$$

Does this imply that $kX + pY$ is an eigenvector? Explain.

Answer

The formula follows from properties of matrix multiplications. However, this vector might not be an eigenvector because it might equal 0 and eigenvectors cannot equal 0.

Exercise 7.E.6

Suppose A is a 3×3 matrix and the following information is available.

$$A\begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix} = 0\begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix}$$
$$A\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = -2\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$
$$A\begin{bmatrix} -2 \\ -3 \\ -2 \end{bmatrix} = -2\begin{bmatrix} -2 \\ -3 \\ -2 \end{bmatrix}$$

Find $A\begin{bmatrix} 1 \\ -4 \\ 3 \end{bmatrix}$.

Exercise 7.E.7

Suppose A is a 3×3 matrix and the following information is available.

$$A\begin{bmatrix} -1 \\ -2 \\ -2 \end{bmatrix} = 1\begin{bmatrix} -1 \\ -2 \\ -2 \end{bmatrix}$$
$$A\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$
$$A\begin{bmatrix} -1 \\ -4 \\ -3 \end{bmatrix} = 2\begin{bmatrix} -1 \\ -4 \\ -3 \end{bmatrix}$$

Find $A\begin{bmatrix} 3 \\ -4 \\ 3 \end{bmatrix}$.

Exercise 7.E.8

Suppose A is a 3×3 matrix and the following information is available.

$$A\begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix} = 2\begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix}$$
$$A\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$
$$A\begin{bmatrix} -3 \\ -5 \\ -4 \end{bmatrix} = -3\begin{bmatrix} -3 \\ -5 \\ -4 \end{bmatrix}$$

Find $A\begin{bmatrix} 2 \\ -3 \\ 3 \end{bmatrix}$.

Exercise 7.E.9

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} -6 & -92 & 12 \\ 0 & 0 & 0 \\ -2 & -31 & 4 \end{bmatrix}$$

One eigenvalue is -2.

Exercise 7.E.10

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} -2 & -17 & -6 \\ 0 & 0 & 0 \\ 1 & 9 & 3 \end{bmatrix}$$

One eigenvalue is 1.

Exercise 7.E.11

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 9 & 2 & 8 \\ 2 & -6 & -2 \\ -8 & 2 & -5 \end{bmatrix}$$

One eigenvalue is -3.

Exercise 7.E.12

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 6 & 76 & 16 \\ -2 & -21 & -4 \\ 2 & 64 & 17 \end{bmatrix}$$

One eigenvalue is -2.

Exercise 7.E.13

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 3 & 5 & 2 \\ -8 & -11 & -4 \\ 10 & 11 & 3 \end{bmatrix}$$

One eigenvalue is -3.

Exercise 7.E.14

Is it possible for a nonzero matrix to have only 0 as an eigenvalue?

Answer

Yes. $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ works.

Exercise 7.E.15

If A is the matrix of a linear transformation which rotates all vectors in $\mathbb{R}^2$ through 60°, explain why A cannot have any real eigenvalues. Is there an angle such that rotation through this angle would have a real eigenvalue? What eigenvalues would be obtainable in this way?

Exercise 7.E.16

Let A be the 2×2 matrix of the linear transformation which rotates all vectors in $\mathbb{R}^2$ through an angle of θ. For which values of θ does A have a real eigenvalue?

Answer

When you think of this geometrically, it is clear that the only two values of θ are 0 and π or these added to integer multiples of 2π.

Exercise 7.E.17

Let T be the linear transformation which reflects vectors about the x axis. Find a matrix for T and then find its eigenvalues and eigenvectors.

Answer

The matrix of T is $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. The eigenvectors and eigenvalues are:

$$\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} \leftrightarrow -1, \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \leftrightarrow 1$$

Exercise 7.E.18

Let T be the linear transformation which rotates all vectors in $\mathbb{R}^2$ counterclockwise through an angle of $\pi/2$. Find a matrix of T and then find eigenvalues and eigenvectors.

Answer

The matrix of T is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. The eigenvectors and eigenvalues are:

$$\left\{ \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\} \leftrightarrow -i, \left\{ \begin{bmatrix} i \\ 1 \end{bmatrix} \right\} \leftrightarrow i$$

Exercise 7.E.19

Let T be the linear transformation which reflects all vectors in $\mathbb{R}^3$ through the xy plane. Find a matrix for T and then obtain its eigenvalues and eigenvectors.

Answer

The matrix of T is $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$. The eigenvectors and eigenvalues are:

$$\left\{\begin{bmatrix}0\\0\\1\end{bmatrix}\right\} \leftrightarrow -1, \left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\} \leftrightarrow 1$$

Exercise 7.E.20

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix}5 & -18 & -32\\0 & 5 & 4\\2 & -5 & -11\end{bmatrix}$$

One eigenvalue is 1. Diagonalize if possible.

Answer

The eigenvalues are $-1, -1, 1$. The eigenvectors corresponding to the eigenvalues are:

$$\{\begin{bmatrix}10 & -2 & 3\end{bmatrix}\} \leftrightarrow -1, \left\{\begin{bmatrix}7\\-2\\2\end{bmatrix}\right\} \leftrightarrow 1$$

Therefore this matrix is not diagonalizable.

Exercise 7.E.21

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix}-12 & -28 & 28\\4 & 9 & -8\\-4 & -8 & 9\end{bmatrix}$$

One eigenvalue is 3. Diagonalize if possible.

Answer

The eigenvectors and eigenvalues are:

$$\left\{\begin{bmatrix}2\\0\\1\end{bmatrix}\right\} \leftrightarrow 1, \left\{\begin{bmatrix}-2\\1\\0\end{bmatrix}\right\} \leftrightarrow 1, \left\{\begin{bmatrix}7\\-2\\2\end{bmatrix}\right\} \leftrightarrow 3$$

The matrix P needed to diagonalize the above matrix is

$$\begin{bmatrix}2 & -2 & 7\\0 & 1 & -2\\1 & 0 & 2\end{bmatrix}$$

and the diagonal matrix D is

$$\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 3\end{bmatrix}$$

Exercise 7.E.22

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 89 & 38 & 268 \\ 14 & 2 & 40 \\ -30 & -12 & -90 \end{bmatrix}$$

One eigenvalue is -3. Diagonalize if possible.

Answer

The eigenvectors and eigenvalues are:

$$\left\{\begin{bmatrix} -6 \\ -1 \\ -2 \end{bmatrix}\right\} \leftrightarrow 6, \left\{\begin{bmatrix} -5 \\ -2 \\ 2 \end{bmatrix}\right\} \leftrightarrow -3, \left\{\begin{bmatrix} -8 \\ -2 \\ 3 \end{bmatrix}\right\} \leftrightarrow 2$$

The matrix P needed to diagonalize the above matrix is

$$\begin{bmatrix} -6 & -5 & -8 \\ -1 & -2 & -2 \\ 2 & 2 & 3 \end{bmatrix}$$

and the diagonal matrix D is

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Exercise 7.E.23

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 1 & 90 & 0 \\ 0 & -2 & 0 \\ 3 & 89 & -2 \end{bmatrix}$$

One eigenvalue is 1. Diagonalize if possible.

Exercise 7.E.24

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 11 & 45 & 30 \\ 10 & 26 & 20 \\ -20 & -60 & -44 \end{bmatrix}$$

One eigenvalue is 1. Diagonalize if possible.

Exercise 7.E.25

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 95 & 25 & 24 \\ -196 & -53 & -48 \\ -164 & -42 & -43 \end{bmatrix}$$

One eigenvalue is 5. Diagonalize if possible.

Exercise 7.E.26

Suppose A is an $n \times n$ matrix and let V be an eigenvector such that $AV = \lambda V$. Also suppose the characteristic polynomial of A is

$$\det(xI - A) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

Explain why

$$(A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I)V = 0$$

If A is diagonalizable, give a proof of the Cayley Hamilton theorem based on this. This theorem says A satisfies its characteristic equation

$$A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I = 0$$

Exercise 7.E.27

Suppose the characteristic polynomial of an $n \times n$ matrix A is $1 - X^n$. Find A^{mn} where m is an integer.

Answer

The eigenvalues are distinct because they are the nth roots of 1. Hence if X is a given vector with

$$X = \sum_{j=1}^{n} a_j V_j$$

then

$$A^{nm}X = A^{nm}\sum_{j=1}^{n} a_j V_j = \sum_{j=1}^{n} a_j A^{nm} V_j = \sum_{j=1}^{n} a_j V_j = X$$

so $A^{nm} = I$.

Exercise 7.E.28

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 15 & -24 & 7 \\ -6 & 5 & -1 \\ -58 & 76 & -20 \end{bmatrix}$$

One eigenvalue is -2. Diagonalize if possible. Hint: This one has some complex eigenvalues.

Exercise 7.E.29

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 15 & -25 & 6 \\ -13 & 23 & -4 \\ -91 & 155 & -30 \end{bmatrix}$$

One eigenvalue is 2. Diagonalize if possible. Hint: This one has some complex eigenvalues.

Exercise 7.E.30

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} -11 & -12 & 4 \\ 8 & 17 & -4 \\ -4 & 28 & -3 \end{bmatrix}$$

One eigenvalue is 1. Diagonalize if possible. Hint: This one has some complex eigenvalues.

Exercise 7.E.31

Find the eigenvalues and eigenvectors of the matrix

$$\begin{bmatrix} 14 & -12 & 5 \\ -6 & 2 & -1 \\ -69 & 51 & -21 \end{bmatrix}$$

One eigenvalue is -3. Diagonalize if possible. Hint: This one has some complex eigenvalues.

Exercise 7.E.32

Suppose A is an $n \times n$ matrix consisting entirely of real entries but $a+ib$ is a complex eigenvalue having the eigenvector, $X+iY$ Here X and Y are real vectors. Show that then $a-ib$ is also an eigenvalue with the eigenvector, $X-iY$. Hint: You should remember that the conjugate of a product of complex numbers equals the product of the conjugates. Here $a+ib$ is a complex number whose conjugate equals $a-ib$.

Answer

$AX = (a+ib)X$. Now take conjugates of both sides. Since A is real,

$$A\overline{X} = (a-ib)\overline{X}$$

Exercise 7.E.33

Let $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. Diagonalize A to find A^{10}.

Answer

First we write $A = PDP^{-1}$.

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Therefore $A^{10} = PD^{10}P^{-1}$.

$$\begin{aligned} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}^{10} &= \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}^{10} \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (-1)^{10} & 0 \\ 0 & 3^{10} \end{bmatrix} \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \\ &= \begin{bmatrix} 29525 & 29524 \\ 29524 & 29525 \end{bmatrix} \end{aligned}$$

Exercise 7.E.34

Let $A = \begin{bmatrix} 1 & 4 & 1 \\ 0 & 2 & 5 \\ 0 & 0 & 5 \end{bmatrix}$. Diagonalize A to find A^{50}.

Exercise 7.E.35

Let $A = \begin{bmatrix} 1 & -2 & -1 \\ 2 & -1 & 1 \\ -2 & 3 & 1 \end{bmatrix}$. Diagonalize A to find A^{100}.

Exercise 7.E.36

The following is a Markov (migration) matrix for three locations

$$\begin{bmatrix} \frac{7}{10} & \frac{1}{9} & \frac{1}{5} \\ \frac{1}{10} & \frac{7}{9} & \frac{2}{5} \\ \frac{1}{5} & \frac{1}{9} & \frac{2}{5} \end{bmatrix}$$

a. Initially, there are 90 people in location 1, 81 in location 2, and 85 in location 3. How many are in each location after one time period?
b. The total number of individuals in the migration process is 256. After a long time, how many are in each location?

Answer

a. Multiply the given matrix by the initial state vector given by $\begin{bmatrix} 90 \\ 81 \\ 85 \end{bmatrix}$. After one time period there are 89 people in location 1, 106 in location 2, and 61 in location 3.

b. Solve the system given by $(I - A)X_s = 0$ where A is the migration matrix and $X_s = \begin{bmatrix} x_{1s} \\ x_{2s} \\ x_{3s} \end{bmatrix}$ is the steady state vector. The solution to this system is given by

$$x_{1s} = \frac{8}{5}x_{3s}$$
$$x_{2s} = \frac{63}{25}x_{3s}$$

Letting $x_{3s} = t$ and using the fact that there are a total of 256 individuals, we must solve

$$\frac{8}{5}t + \frac{63}{25}t + t = 256$$

We find that $t = 50$. Therefore after a long time, there are 80 people in location 1, 126 in location 2, and 50 in location 3.

Exercise 7.E.37

The following is a Markov (migration) matrix for three locations

$$\begin{bmatrix} \frac{1}{5} & \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \\ \frac{2}{5} & \frac{2}{5} & \frac{2}{5} \end{bmatrix}$$

a. Initially, there are 130 individuals in location 1, 300 in location 2, and 70 in location 3. How many are in each location after two time periods?
b. The total number of individuals in the migration process is 500. After a long time, how many are in each location?

Exercise $7.E.38$

The following is a Markov (migration) matrix for three locations

$$\begin{bmatrix} \frac{3}{10} & \frac{3}{8} & \frac{1}{3} \\ \frac{1}{10} & \frac{3}{8} & \frac{1}{3} \\ \frac{3}{5} & \frac{1}{4} & \frac{1}{3} \end{bmatrix}$$

The total number of individuals in the migration process is 480. After a long time, how many are in each location?

Answer

We solve $(I-A)X_s = 0$ to find the steady state vector $X_s = \begin{bmatrix} x_{1s} \\ x_{2s} \\ x_{3s} \end{bmatrix}$. The solution to the system is given by

$$x_{1s} = \frac{5}{6}x_{3s}$$
$$x_{2s} = \frac{2}{3}x_{3s}$$

Letting $x_{3s} = t$ and using the fact that there are a total of 480 individuals, we must solve

$$\frac{5}{6}t + \frac{2}{3}t + t = 480$$

We find that $t = 192$. Therefore after a long time, there are 160 people in location 1, 128 in location 2, and 192 in location 3.

Exercise $7.E.39$

The following is a Markov (migration) matrix for three locations

$$\begin{bmatrix} \frac{3}{10} & \frac{1}{3} & \frac{1}{5} \\ \frac{3}{10} & \frac{1}{3} & \frac{7}{10} \\ \frac{2}{5} & \frac{1}{3} & \frac{1}{10} \end{bmatrix}$$

The total number of individuals in the migration process is 1155. After a long time, how many are in each location?

Exercise $7.E.40$

The following is a Markov (migration) matrix for three locations

$$\begin{bmatrix} \frac{2}{5} & \frac{1}{10} & \frac{1}{8} \\ \frac{3}{10} & \frac{2}{5} & \frac{5}{8} \\ \frac{3}{10} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

The total number of individuals in the migration process is 704. After a long time, how many are in each location?

Exercise 7.E.41

A person sets off on a random walk with three possible locations. The Markov matrix of probabilities $A = [a_{ij}]$ is given by

$$\begin{bmatrix} 0.1 & 0.3 & 0.7 \\ 0.1 & 0.3 & 0.2 \\ 0.8 & 0.4 & 0.1 \end{bmatrix}$$

If the walker starts in location 2, what is the probability of ending back in location 2 at time $n = 3$?

Answer

$$X_3 = \begin{bmatrix} 0.38 \\ 0.18 \\ 0.44 \end{bmatrix}$$

Therefore the probability of ending up back in location 2 is 0.18.

Exercise 7.E.42

A person sets off on a random walk with three possible locations. The Markov matrix of probabilities $A = [a_{ij}]$ is given by

$$\begin{bmatrix} 0.5 & 0.1 & 0.6 \\ 0.2 & 0.9 & 0.2 \\ 0.3 & 0 & 0.2 \end{bmatrix}$$

It is unknown where the walker starts, but the probability of starting in each location is given by

$$X_0 = \begin{bmatrix} 0.2 \\ 0.25 \\ 0.55 \end{bmatrix}$$

What is the probability of the walker being in location 1 at time $n = 2$?

Answer

$$X_2 = \begin{bmatrix} 0.367 \\ 0.4625 \\ 0.1705 \end{bmatrix}$$

Therefore the probability of ending up in location 1 is 0.367.

Exercise 7.E.43

You own a trailer rental company in a large city and you have four locations, one in the South East, one in the North East, one in the North West, and one in the South West. Denote these locations by SE, NE, NW, and SW respectively. Suppose that the following table is observed to take place.

Table 7.E.1

	SE	NE	NW	SW
SE	$\frac{1}{3}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{5}$
NE	$\frac{1}{3}$	$\frac{7}{10}$	$\frac{1}{5}$	$\frac{1}{10}$
NW	$\frac{2}{9}$	$\frac{1}{10}$	$\frac{3}{5}$	$\frac{1}{5}$
SW	$\frac{1}{9}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{2}$

In this table, the probability that a trailer starting at NE ends in NW is $1/10$, the probability that a trailer starting at SW ends in NW is $1/5$, and so forth. Approximately how many will you have in each location after a long time if the total number of trailers is 413?

Answer

The migration matrix is

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{10} & \frac{1}{10} & \frac{1}{5} \\ \frac{1}{3} & \frac{7}{10} & \frac{1}{5} & \frac{1}{10} \\ \frac{2}{9} & \frac{1}{10} & \frac{3}{5} & \frac{1}{5} \\ \frac{1}{9} & \frac{1}{10} & \frac{1}{10} & \frac{1}{2} \end{bmatrix}$$

To find the number of trailers in each location after a long time we solve system $(I - A)X_s = 0$ for the steady state vector $X_s = \begin{bmatrix} x_{1s} \\ x_{2s} \\ x_{3s} \\ x_{4s} \end{bmatrix}$. The solution to the system is

$$x_{1s} = \frac{9}{10}x_{4s}$$
$$x_{2s} = \frac{12}{5}x_{4s}$$
$$x_{3s} = \frac{8}{5}x_{4s}$$

Letting $x_{4s} = t$ and using the fact that there are a total of 413 trailers we must solve

$$\frac{9}{10}t + \frac{12}{5}t + \frac{8}{5}t + t = 413$$

We find that $t = 70$. Therefore after a long time, there are 63 trailers in the SE, 168 in the NE, 112 in the NW and 70 in the SW.

Exercise 7.E.44

You own a trailer rental company in a large city and you have four locations, one in the South East, one in the North East, one in the North West, and one in the South West. Denote these locations by SE, NE, NW, and SW respectively. Suppose that the following table is observed to take place.

Table 7.E.2

	SE	NE	NW	SW
SE	$\frac{1}{7}$	$\frac{1}{4}$	$\frac{1}{10}$	$\frac{1}{5}$
NE	$\frac{2}{7}$	$\frac{1}{4}$	$\frac{1}{5}$	$\frac{1}{10}$
NW	$\frac{1}{7}$	$\frac{1}{4}$	$\frac{3}{5}$	$\frac{1}{5}$
SW	$\frac{3}{7}$	$\frac{1}{4}$	$\frac{1}{10}$	$\frac{1}{2}$

In this table, the probability that a trailer starting at NE ends in NW is 1/10, the probability that a trailer starting at SW ends in NW is 1/5, and so forth. Approximately how many will you have in each location after a long time if the total number of trailers is 1469.

Exercise 7.E.45

The following table describes the transition probabilities between the states rainy, partly cloudy and sunny. The symbol p.c. indicates partly cloudy. Thus if it starts off p.c. it ends up sunny the next day with probability $\frac{1}{5}$. If it starts off sunny, it ends up sunny the next day with probability $\frac{2}{5}$ and so forth.

Table 7.E.3

	rains	sunny	p.c.
rains	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{3}$
sunny	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{1}{3}$
p.c.	$\frac{3}{5}$	$\frac{2}{5}$	$\frac{1}{3}$

Given this information, what are the probabilities that a given day is rainy, sunny, or partly cloudy?

Exercise 7.E.46

The following table describes the transition probabilities between the states rainy, partly cloudy and sunny. The symbol p.c. indicates partly cloudy. Thus if it starts off p.c. it ends up sunny the next day with probability $\frac{1}{10}$. If it starts off sunny, it ends up sunny the next day with probability $\frac{2}{5}$ and so forth.

Table 7.E.4

	rains	sunny	p.c.
rains	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{3}$
sunny	$\frac{1}{10}$	$\frac{2}{5}$	$\frac{4}{9}$
p.c.	$\frac{7}{10}$	$\frac{2}{5}$	$\frac{2}{9}$

Given this information, what are the probabilities that a given day is rainy, sunny, or partly cloudy?

Exercise 7.E.47

You own a trailer rental company in a large city and you have four locations, one in the South East, one in the North East, one in the North West, and one in the South West. Denote these locations by SE, NE, NW, and SW respectively. Suppose that the following table is observed to take place.

Table 7.E.5

	SE	NE	NW	SW
SE	$\frac{5}{11}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{5}$
NE	$\frac{1}{11}$	$\frac{7}{10}$	$\frac{1}{5}$	$\frac{1}{10}$
NW	$\frac{2}{11}$	$\frac{1}{10}$	$\frac{3}{5}$	$\frac{1}{5}$
SW	$\frac{3}{11}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{2}$

In this table, the probability that a trailer starting at NE ends in NW is 1/10, the probability that a trailer starting at SW ends in NW is 1/5, and so forth. Approximately how many will you have in each location after a long time if the total number of trailers is 407?

Exercise 7.E.48

The University of Poohbah offers three degree programs, scouting education (SE), dance appreciation (DA), and engineering (E). It has been determined that the probabilities of transferring from one program to another are as in the following table.

Table 7.E.6

	SE	DA	E
SE	.8	.1	.3
DA	.1	.7	.5
E	.1	.2	.2

where the number indicates the probability of transferring from the top program to the program on the left. Thus the probability of going from DA to E is .2. Find the probability that a student is enrolled in the various programs.

Exercise 7.E.49

In the city of Nabal, there are three political persuasions, republicans (R), democrats (D), and neither one (N). The following table shows the transition probabilities between the political parties, the top row being the initial political party and the side row being the political affiliation the following year.

Table 7.E.7

	R	D	N
R	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{2}{7}$
D	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{4}{7}$
N	$\frac{3}{5}$	$\frac{1}{2}$	$\frac{1}{7}$

Find the probabilities that a person will be identified with the various political persuasions. Which party will end up being most important?

Exercise 7.E.50

The following table describes the transition probabilities between the states rainy, partly cloudy and sunny. The symbol p.c. indicates partly cloudy. Thus if it starts off p.c. it ends up sunny the next day with probability $\frac{1}{5}$. If it starts off sunny, it ends up sunny the next day with probability $\frac{2}{7}$ and so forth.

Table 7.E.8

	rains	sunny	p.c.
rains	$\frac{1}{5}$	$\frac{2}{7}$	$\frac{5}{9}$
sunny	$\frac{1}{5}$	$\frac{2}{7}$	$\frac{1}{3}$
p.c.	$\frac{3}{5}$	$\frac{3}{7}$	$\frac{1}{9}$

Given this information, what are the probabilities that a given day is rainy, sunny, or partly cloudy?

Exercise 7.E.51

Find the solution to the initial value problem

$$\begin{bmatrix} x \\ y \end{bmatrix}' = \begin{bmatrix} 0 & -1 \\ 6 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$\begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Hint: form the matrix exponential e^{At} and then the solution is $e^{At}C$ where C is the initial vector,

Answer

The solution is

$$e^{At}C = \begin{bmatrix} 8e^{2t} - 6e^{3t} \\ 18e^{3t} - 16e^{2t} \end{bmatrix}$$

Exercise 7.E.52

Find the solution to the initial value problem

$$\begin{bmatrix} x \\ y \end{bmatrix}' = \begin{bmatrix} -4 & -3 \\ 6 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$\begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Hint: form the matrix exponential e^{At} and then the solution is $e^{At}C$ where C is the initial vector.

Exercise 7.E.53

Find the solution to the initial value problem

$$\begin{bmatrix} x \\ y \end{bmatrix}' = \begin{bmatrix} -1 & 2 \\ -4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$\begin{bmatrix} x(0) \\ y(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Hint: form the matrix exponential e^{At} and then the solution is $e^{At}C$ where C is the initial vector.

Exercise 7.E.54

Find the eigenvalues and an orthonormal basis of eigenvectors for A.

$$A = \begin{bmatrix} 11 & -1 & -4 \\ -1 & 11 & -4 \\ -4 & -4 & 14 \end{bmatrix}$$

Hint: Two eigenvalues are 12 and 18.

Answer

The eigenvectors and eigenvalues are:

$$\left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \leftrightarrow 6, \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\} \leftrightarrow 12, \left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \right\} \leftrightarrow 18$$

Exercise $7.E.55$

Find the eigenvalues and an orthonormal basis of eigenvectors for A.

$$A = \begin{bmatrix} 4 & 1 & -2 \\ 1 & 4 & -2 \\ -2 & -2 & 7 \end{bmatrix}$$

Hint: One eigenvalue is 3.

Answer

The eigenvectors and eigenvalues are:

$$\left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \leftrightarrow 3, \left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix} \right\} \leftrightarrow 9$$

Exercise $7.E.56$

Find the eigenvalues and an orthonormal basis of eigenvectors for A. Diagonalize A by finding an orthogonal matrix U and a diagonal matrix D such that $U^T AU = D$.

$$A = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

Hint: One eigenvalue is -2.

Answer

The eigenvectors and eigenvalues are:

$$\left\{ \begin{bmatrix} \frac{1}{3}\sqrt{3} \\ \frac{1}{3}\sqrt{3} \\ \frac{1}{3}\sqrt{3} \end{bmatrix} \right\} \leftrightarrow 1, \left\{ \begin{bmatrix} -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{6}\sqrt{6} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{2}\sqrt{3} \end{bmatrix} \right\} \leftrightarrow -2$$

$$\begin{bmatrix} \sqrt{3}/3 & -\sqrt{2}/2 & -\sqrt{6}/6 \\ \sqrt{3}/3 & \sqrt{2}/2 & -\sqrt{6}/6 \\ \sqrt{3}/3 & 0 & \frac{1}{3}\sqrt{2}\sqrt{3} \end{bmatrix}^T \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} \sqrt{3}/3 & -\sqrt{2}/2 & -\sqrt{6}/6 \\ \sqrt{3}/3 & \sqrt{2}/2 & -\sqrt{6}/6 \\ \sqrt{3}/3 & 0 & \frac{1}{3}\sqrt{2}\sqrt{3} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Exercise $7.E.57$

Find the eigenvalues and an orthonormal basis of eigenvectors for A. Diagonalize A by finding an orthogonal matrix U and a diagonal matrix D such that $U^T AU = D$.

$$A = \begin{bmatrix} 17 & -7 & -4 \\ -7 & 17 & -4 \\ -4 & -4 & 14 \end{bmatrix}$$

Hint: Two eigenvalues are 18 and 24.

Answer

The eigenvectors and eigenvalues are:

$$\left\{ \begin{bmatrix} \frac{1}{3}\sqrt{3} \\ \frac{1}{3}\sqrt{3} \\ \frac{1}{3}\sqrt{3} \end{bmatrix} \right\} \leftrightarrow 6, \left\{ \begin{bmatrix} -\frac{1}{6}\sqrt{6} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{2}\sqrt{3} \end{bmatrix} \right\} \leftrightarrow 18, \left\{ \begin{bmatrix} -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{bmatrix} \right\} \leftrightarrow 24$$

The matrix U has these as its columns.

Exercise 7.E.58

Find the eigenvalues and an orthonormal basis of eigenvectors for A. Diagonalize A by finding an orthogonal matrix U and a diagonal matrix D such that $U^T AU = D$.

$$A = \begin{bmatrix} 13 & 1 & 4 \\ 1 & 13 & 4 \\ 4 & 4 & 10 \end{bmatrix}$$

Hint: Two eigenvalues are 12 and 18.

Answer

The eigenvectors and eigenvalues are:

$$\left\{ \begin{bmatrix} -\frac{1}{6}\sqrt{6} \\ -\frac{1}{6}\sqrt{6} \\ \frac{1}{3}\sqrt{2}\sqrt{3} \end{bmatrix} \right\} \leftrightarrow 6, \left\{ \begin{bmatrix} -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \\ 0 \end{bmatrix} \right\} \leftrightarrow 12, \left\{ \begin{bmatrix} \frac{1}{3}\sqrt{3} \\ \frac{1}{3}\sqrt{3} \\ \frac{1}{3}\sqrt{3} \end{bmatrix} \right\} \leftrightarrow 18.$$

The matrix U has these as its columns.

Exercise 7.E.59

Find the eigenvalues and an orthonormal basis of eigenvectors for A. Diagonalize A by finding an orthogonal matrix U and a diagonal matrix D such that $U^T AU = D$.

$$A = \begin{bmatrix} -\frac{5}{3} & \frac{1}{15}\sqrt{6}\sqrt{5} & \frac{8}{15}\sqrt{5} \\ \frac{1}{15}\sqrt{6}\sqrt{5} & -\frac{14}{5} & -\frac{1}{15}\sqrt{6} \\ \frac{8}{15}\sqrt{5} & -\frac{1}{15}\sqrt{6} & \frac{7}{15} \end{bmatrix}$$

Hint: The eigenvalues are $-3, -2, 1$.

Answer

The eigenvectors and eigenvalues are:

$$\left\{ \begin{bmatrix} \frac{1}{6}\sqrt{6} \\ 0 \\ \frac{1}{6}\sqrt{5}\sqrt{6} \end{bmatrix} \right\} \leftrightarrow 1, \left\{ \begin{bmatrix} -\frac{1}{3}\sqrt{2}\sqrt{3} \\ -\frac{1}{5}\sqrt{5} \\ \frac{1}{15}\sqrt{2}\sqrt{15} \end{bmatrix} \right\} - 2, \left\{ \begin{bmatrix} -\frac{1}{6}\sqrt{6} \\ \frac{2}{5}\sqrt{5} \\ \frac{1}{30}\sqrt{30} \end{bmatrix} \right\} \leftrightarrow -3$$

These vectors are the columns of U.

Exercise 7.E.60

Find the eigenvalues and an orthonormal basis of eigenvectors for A. Diagonalize A by finding an orthogonal matrix U and a diagonal matrix D such that $U^T AU = D$.

$$A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & \frac{3}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{3}{2} \end{bmatrix}$$

Answer

The eigenvectors and eigenvalues are:

$$\left\{ \begin{bmatrix} 0 \\ -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix} \right\} \leftrightarrow 1, \left\{ \begin{bmatrix} 0 \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix} \right\} \leftrightarrow 2, \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\} \leftrightarrow 3.$$

These vectors are the columns of the matrix U.

Exercise 7.E.61

Find the eigenvalues and an orthonormal basis of eigenvectors for A. Diagonalize A by finding an orthogonal matrix U and a diagonal matrix D such that $U^T AU = D$.

$$A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 1 \\ 0 & 1 & 5 \end{bmatrix}$$

Answer

The eigenvectors and eigenvalues are:

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\} \leftrightarrow 2, \left\{ \begin{bmatrix} 0 \\ -\frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix} \right\} \leftrightarrow 4, \left\{ \begin{bmatrix} 0 \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix} \right\} \leftrightarrow 6.$$

These vectors are the columns of U.

Exercise 7.E.62

Find the eigenvalues and an orthonormal basis of eigenvectors for A. Diagonalize A by finding an orthogonal matrix U and a diagonal matrix D such that $U^T AU = D$.

$$A = \begin{bmatrix} \frac{4}{3} & \frac{1}{3}\sqrt{3}\sqrt{2} & \frac{1}{3}\sqrt{2} \\ \frac{1}{3}\sqrt{3}\sqrt{2} & 1 & -\frac{1}{3}\sqrt{3} \\ \frac{1}{3}\sqrt{2} & -\frac{1}{3}\sqrt{3} & \frac{5}{3} \end{bmatrix}$$

Hint: The eigenvalues are $0, 2, 2$ where 2 is listed twice because it is a root of multiplicity 2.

Answer

The eigenvectors and eigenvalues are:

$$\left\{\begin{bmatrix} -\frac{1}{5}\sqrt{2}\sqrt{5} \\ \frac{1}{5}\sqrt{3}\sqrt{5} \\ \frac{1}{5}\sqrt{5} \end{bmatrix}\right\} \leftrightarrow 0, \left\{\begin{bmatrix} \frac{1}{3}\sqrt{3} \\ 0 \\ \frac{1}{3}\sqrt{2}\sqrt{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{5}\sqrt{2}\sqrt{5} \\ \frac{1}{5}\sqrt{3}\sqrt{5} \\ -\frac{1}{5}\sqrt{5} \end{bmatrix}\right\} \leftrightarrow 2.$$

The columns are these vectors.

Exercise 7.E.63

Find the eigenvalues and an orthonormal basis of eigenvectors for A. Diagonalize A by finding an orthogonal matrix U and a diagonal matrix D such that $U^T AU = D$.

$$A = \begin{bmatrix} 1 & \frac{1}{6}\sqrt{3}\sqrt{2} & \frac{1}{6}\sqrt{3}\sqrt{6} \\ \frac{1}{6}\sqrt{3}\sqrt{2} & \frac{3}{2} & \frac{1}{12}\sqrt{2}\sqrt{6} \\ \frac{1}{6}\sqrt{3}\sqrt{6} & \frac{1}{12}\sqrt{2}\sqrt{6} & \frac{1}{2} \end{bmatrix}$$

Hint: The eigenvalues are $2, 1, 0$.

Answer

The eigenvectors and eigenvalues are:

$$\left\{\begin{bmatrix} -\frac{1}{3}\sqrt{3} \\ 0 \\ \frac{1}{3}\sqrt{2}\sqrt{3} \end{bmatrix}\right\} \leftrightarrow 0, \left\{\begin{bmatrix} \frac{1}{3}\sqrt{3} \\ -\frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{bmatrix}\right\} \leftrightarrow 1, \left\{\begin{bmatrix} \frac{1}{3}\sqrt{3} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{bmatrix}\right\} \leftrightarrow 2.$$

The columns are these vectors.

Exercise 7.E.64

Find the eigenvalues and an orthonormal basis of eigenvectors for the matrix

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{6}\sqrt{3}\sqrt{2} & -\frac{7}{18}\sqrt{3}\sqrt{6} \\ \frac{1}{6}\sqrt{3}\sqrt{2} & \frac{3}{2} & -\frac{1}{12}\sqrt{2}\sqrt{6} \\ -\frac{7}{18}\sqrt{3}\sqrt{6} & -\frac{1}{12}\sqrt{2}\sqrt{6} & -\frac{5}{6} \end{bmatrix}$$

Hint: The eigenvalues are $1, 2, -2$.

Answer

The eigenvectors and eigenvalues are:

$$\left\{\begin{bmatrix} -\frac{1}{3}\sqrt{3} \\ \frac{1}{2}\sqrt{2} \\ \frac{1}{6}\sqrt{6} \end{bmatrix}\right\} \leftrightarrow 1, \left\{\begin{bmatrix} \frac{1}{3}\sqrt{3} \\ 0 \\ \frac{1}{3}\sqrt{2}\sqrt{3} \end{bmatrix}\right\} \leftrightarrow -2, \left\{\begin{bmatrix} \frac{1}{3}\sqrt{3} \\ \frac{1}{2}\sqrt{2} \\ -\frac{1}{6}\sqrt{6} \end{bmatrix}\right\} \leftrightarrow 2.$$

Then the columns of U are these vectors.

Exercise 7.E.65

Find the eigenvalues and an orthonormal basis of eigenvectors for the matrix

$$A = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{5}\sqrt{6}\sqrt{5} & \frac{1}{10}\sqrt{5} & \\ -\frac{1}{5}\sqrt{6}\sqrt{5} & \frac{7}{5} & -\frac{1}{5}\sqrt{6} & \\ \frac{1}{10} & \sqrt{5} & -\frac{1}{5}\sqrt{6} & -\frac{9}{10} \end{bmatrix}$$

Hint: The eigenvalues are $-1, 2, -1$ where -1 is listed twice because it has multiplicity 2 as a zero of the characteristic equation.

Answer

The eigenvectors and eigenvalues are:

$$\left\{ \begin{bmatrix} -\frac{1}{6}\sqrt{6} \\ 0 \\ \frac{1}{6}\sqrt{5}\sqrt{6} \end{bmatrix}, \begin{bmatrix} \frac{1}{3}\sqrt{2}\sqrt{3} \\ \frac{1}{5}\sqrt{5} \\ \frac{1}{15}\sqrt{2}\sqrt{15} \end{bmatrix} \right\} \leftrightarrow -1, \left\{ \begin{bmatrix} \frac{1}{6}\sqrt{6} \\ -\frac{2}{5}\sqrt{5} \\ \frac{1}{30}\sqrt{30} \end{bmatrix} \right\} \leftrightarrow 2.$$

The columns of U are these vectors.

$$\begin{bmatrix} -\frac{1}{6}\sqrt{6} & \frac{1}{3}\sqrt{2}\sqrt{3} & \frac{1}{6}\sqrt{6} \\ 0 & \frac{1}{5}\sqrt{5} & -\frac{2}{5}\sqrt{5} \\ \frac{1}{6}\sqrt{5}\sqrt{6} & \frac{1}{15}\sqrt{2}\sqrt{15} & \frac{1}{30}\sqrt{30} \end{bmatrix}^T \begin{bmatrix} -\frac{1}{2} & -\frac{1}{5}\sqrt{6}\sqrt{5} & \frac{1}{10}\sqrt{5} \\ -\frac{1}{5}\sqrt{6}\sqrt{5} & \frac{7}{5} & -\frac{1}{5}\sqrt{6} \\ \frac{1}{10}\sqrt{5} & -\frac{1}{5}\sqrt{6} & -\frac{9}{10} \end{bmatrix}.$$

$$\begin{bmatrix} -\frac{1}{6}\sqrt{6} & \frac{1}{3}\sqrt{2}\sqrt{3} & \frac{1}{6}\sqrt{6} \\ 0 & \frac{1}{5}\sqrt{5} & -\frac{2}{5}\sqrt{5} \\ \frac{1}{6}\sqrt{5}\sqrt{6} & \frac{1}{15}\sqrt{2}\sqrt{15} & \frac{1}{30}\sqrt{30} \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Exercise $7.E.66$

Explain why a matrix A is symmetric if and only if there exists an orthogonal matrix U such that $A = U^T DU$ for D a diagonal matrix.

Answer

If A is given by the formula, then

$$A^T = U^T D^T U = U^T DU = A$$

Next suppose $A = A^T$. Then by the theorems on symmetric matrices, there exists an orthogonal matrix U such that

$$UAU^T = D$$

for D diagonal. Hence

$$A = U^T DU$$

Exercise $7.E.67$

Show that if A is a real symmetric matrix and λ and μ are two different eigenvalues, then if X is an eigenvector for λ and Y is an eigenvector for μ, then $X \bullet Y = 0$. Also all eigenvalues are real. Supply reasons for each step in the following argument. First

$$\lambda X^T \overline{X} = (AX)^T \overline{X} = X^T A\overline{X} = X^T \overline{AX} = X^T \overline{\lambda X} = \overline{\lambda} X^T \overline{X}$$

and so $\lambda = \overline{\lambda}$. This shows that all eigenvalues are real. It follows all the eigenvectors are real. Why? Now let X, Y, μ and λ be given as above.

$$\lambda(X \bullet Y) = \lambda X \bullet Y = AX \bullet Y = X \bullet AY = X \bullet \mu Y = \mu(X \bullet Y) = \mu(X \bullet Y)$$

and so

$$(\lambda - \mu) X \bullet Y = 0$$

Why does it follow that $X \bullet Y = 0$?

Answer

Since $\lambda \neq \mu$, it follows $X \bullet Y = 0$.

Exercise 7.E.68

Find the Cholesky factorization for the matrix

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 6 & 4 \\ 0 & 4 & 10 \end{bmatrix}$$

Exercise 7.E.69

Find the Cholesky factorization for the matrix

$$\begin{bmatrix} 4 & 8 & 0 \\ 8 & 17 & 2 \\ 0 & 2 & 13 \end{bmatrix}$$

Exercise 7.E.70

Find the Cholesky factorization for the matrix

$$\begin{bmatrix} 4 & 8 & 0 \\ 8 & 20 & 8 \\ 0 & 8 & 20 \end{bmatrix}$$

Exercise 7.E.71

Find the Cholesky factorization for the matrix

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 8 & 10 \\ 1 & 10 & 18 \end{bmatrix}$$

Exercise 7.E.72

Find the Cholesky factorization for the matrix

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 8 & 10 \\ 1 & 10 & 26 \end{bmatrix}$$

Exercise 7.E.73

Suppose you have a lower triangular matrix L and it is invertible. Show that LL^T must be positive definite.

Exercise 7.E.74

Using the Gram Schmidt process or the QR factorization, find an orthonormal basis for the following span:

$$span\left\{\begin{bmatrix}1\\2\\1\end{bmatrix},\begin{bmatrix}2\\-1\\3\end{bmatrix},\begin{bmatrix}1\\0\\0\end{bmatrix}\right\}$$

Answer

Using the QR factorization, we have:

$$\begin{bmatrix}1&2&1\\2&-1&0\\1&3&0\end{bmatrix}=\begin{bmatrix}\frac{1}{6}\sqrt{6}&\frac{3}{10}\sqrt{2}&\frac{7}{15}\sqrt{3}\\\frac{1}{3}\sqrt{6}&-\frac{2}{5}\sqrt{2}&-\frac{1}{15}\sqrt{3}\\\frac{1}{6}\sqrt{6}&\frac{1}{2}\sqrt{2}&-\frac{1}{3}\sqrt{3}\end{bmatrix}\begin{bmatrix}\sqrt{6}&\frac{1}{2}\sqrt{6}&\frac{1}{6}\sqrt{6}\\0&\frac{5}{2}\sqrt{2}&\frac{3}{10}\sqrt{2}\\0&0&\frac{7}{15}\sqrt{3}\end{bmatrix}$$

A solution is then

$$\begin{bmatrix}\frac{1}{6}\sqrt{6}\\\frac{1}{3}\sqrt{6}\\\frac{1}{6}\sqrt{6}\end{bmatrix},\begin{bmatrix}\frac{3}{10}\sqrt{2}\\-\frac{2}{5}\sqrt{2}\\\frac{1}{2}\sqrt{2}\end{bmatrix},\begin{bmatrix}\frac{7}{15}\sqrt{3}\\-\frac{1}{15}\sqrt{3}\\-\frac{1}{3}\sqrt{3}\end{bmatrix}$$

Exercise 7.E.75

Using the Gram Schmidt process or the QR factorization, find an orthonormal basis for the following span:

$$span\left\{\begin{bmatrix}1\\2\\1\\0\end{bmatrix},\begin{bmatrix}2\\-1\\3\\1\end{bmatrix},\begin{bmatrix}1\\0\\0\\1\end{bmatrix}\right\}$$

Answer

$$\begin{bmatrix}1&2&1\\2&-1&0\\1&3&0\\0&1&1\end{bmatrix}=\begin{bmatrix}\frac{1}{6}\sqrt{6}&\frac{1}{6}\sqrt{2}\sqrt{3}&\frac{5}{111}\sqrt{3}\sqrt{37}&\frac{7}{111}\sqrt{111}\\\frac{1}{3}\sqrt{6}&-\frac{2}{9}\sqrt{2}\sqrt{3}&\frac{1}{333}\sqrt{3}\sqrt{37}&-\frac{2}{111}\sqrt{111}\\\frac{1}{6}\sqrt{6}&\frac{5}{18}\sqrt{2}\sqrt{3}&-\frac{17}{333}\sqrt{3}\sqrt{37}&-\frac{1}{37}\sqrt{111}\\0&\frac{1}{9}\sqrt{2}\sqrt{3}&\frac{22}{333}\sqrt{3}\sqrt{37}&-\frac{7}{111}\sqrt{111}\end{bmatrix}$$
$$\begin{bmatrix}\sqrt{6}&\frac{1}{2}\sqrt{6}&\frac{1}{6}\sqrt{6}\\0&\frac{3}{2}\sqrt{2}\sqrt{3}&\frac{5}{18}\sqrt{2}\sqrt{3}\\0&0&\frac{1}{9}\sqrt{3}\sqrt{37}\\0&0&0\end{bmatrix}$$

Then a solution is

$$\begin{bmatrix}\frac{1}{6}\sqrt{6}&\\\frac{1}{3}\sqrt{6}&\\\frac{1}{6}\sqrt{6}&0\end{bmatrix},\begin{bmatrix}\frac{1}{6}\sqrt{2}\sqrt{3}\\-\frac{2}{9}\sqrt{2}\sqrt{3}\\\frac{5}{18}\sqrt{2}\sqrt{3}\\\frac{1}{9}\sqrt{2}\sqrt{3}\end{bmatrix},\begin{bmatrix}\frac{5}{111}\sqrt{3}\sqrt{37}\\\frac{1}{333}\sqrt{3}\sqrt{37}\\-\frac{17}{333}\sqrt{3}\sqrt{37}\\\frac{22}{333}\sqrt{3}\sqrt{37}\end{bmatrix}$$

Exercise 7.E.76

a. $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix}$

b. $\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}$

d. $\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}$

e. $\begin{bmatrix} \sqrt{11} & 1 & 3\sqrt{6} \\ \sqrt{11} & 7 & -\sqrt{6} \\ 2\sqrt{11} & -4 & -\sqrt{6} \end{bmatrix}$ Hint: Notice that the columns are orthogonal.

Exercise 7.E.77

Using a computer algebra system, find a QR factorization for the following matrices.

a. $\begin{bmatrix} 1 & 1 & 2 \\ 3 & -2 & 3 \\ 2 & 1 & 1 \end{bmatrix}$

b. $\begin{bmatrix} 1 & 2 & 1 & 3 \\ 4 & 5 & -4 & 3 \\ 2 & 1 & 2 & 1 \end{bmatrix}$

c. $\begin{bmatrix} 1 & 2 \\ 3 & 2 \\ 1 & -4 \end{bmatrix}$ Find the thin QR factorization of this one.

Exercise 7.E.78

A quadratic form in three variables is an expression of the form $a_1x^2+a_2y^2+a_3z^2+a_4xy+a_5xz+a_6yz$. Show that every such quadratic form may be written as

$$\begin{bmatrix} x & y & z \end{bmatrix} A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

where A is a symmetric matrix.

Answer

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a_1 & a_4/2 & a_5/2 \\ a_4/2 & a_2 & a_6/2 \\ a_5/2 & a_6/2 & a_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Exercise 7.E.79

Given a quadratic form in three variables, x, y, and z, show there exists an orthogonal matrix U and variables x', y', z' such that

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = U \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

with the property that in terms of the new varaibles, the quadratic form is

$$\lambda_1(x')^2+\lambda_2(y')^2+\lambda_3(z')^2$$

where the numbers, λ_1, λ_2, and λ_3 are the eigenvalues of the matrix A in Exercise $7.E.78$.

Answer

The quadratic form may be written as

$$\vec{x}^T A\vec{x}$$

where $A=A^T$. By the theorem about diagonalizing a symmetric matrix, there exists an orthogonal matrix U such that

$$U^TAU=D,\ A=UDU^T$$

Then the quadratic form is

$$\vec{x}^TUDU^T\vec{x}=(U^T\vec{x})^TD(U^T\vec{x})$$

where D is a diagonal matrix having the real eigenvalues of A down the main diagonal. Now simply let

$$\vec{x}'=U^T\vec{x}$$

? Exercise $7.E.80$

Consider the quadratic form q given by $q=3x_1^2-12x_1x_2-2x_2^2$.

a. Write q in the form $\vec{x}^TA\vec{x}$ for an appropriate symmetric matrix A.
b. Use a change of variables to rewrite q to eliminate the x_1x_2 term.

? Exercise $7.E.81$

Consider the quadratic form q given by $q=-2x_1^2+2x_1x_2-2x_2^2$.

a. Write q in the form $\vec{x}^TA\vec{x}$ for an appropriate symmetric matrix A.
b. Use a change of variables to rewrite q to eliminate the x_1x_2 term.

? Exercise $7.E.82$

Consider the quadratic form q given by $q=7x_1^2+6x_1x_2-x_2^2$.

a. Write q in the form $\vec{x}^TA\vec{x}$ for an appropriate symmetric matrix A.
b. Use a change of variables to rewrite q to eliminate the x_1x_2 term.

7.E: Exercises is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

8: Some Curvilinear Coordinate Systems

8.1: Polar Coordinates and Polar Graphs

8.2: Spherical and Cylindrical Coordinates

8.E: Exercises

This page titled 8: Some Curvilinear Coordinate Systems is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

8.1: Polar Coordinates and Polar Graphs

Outcomes

A. Understand polar coordinates.
B. Convert points between Cartesian and polar coordinates.

You have likely encountered the Cartesian coordinate system in many aspects of mathematics. There is an alternative way to represent points in space, called **polar coordinates**. The idea is suggested in the following picture.

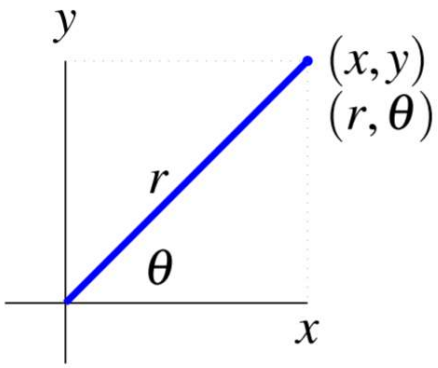


Figure 8.1.1

Consider the point above, which would be specified as (x, y) in Cartesian coordinates. We can also specify this point using polar coordinates, which we write as (r, θ). The number r is the distance from the origin$(0, 0)$ to the point, while θ is the angle shown between the positive x axis and the line from the origin to the point. In this way, the point can be specified in polar coordinates as (r, θ).

Now suppose we are given an ordered pair (r, θ) where r and θ are real numbers. We want to determine the point specified by this ordered pair. We can use θ to identify a ray from the origin as follows. Let the ray pass from $(0, 0)$ through the point $(\cos\theta, \sin\theta)$ as shown.

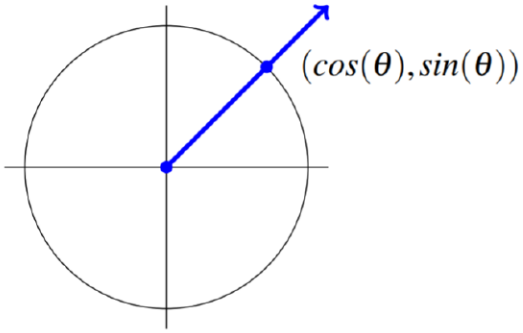


Figure 8.1.2

The ray is identified on the graph as the line from the origin, through the point $(\cos(\theta), \sin(\theta))$. Now if $r > 0$, go a distance equal to r in the direction of the displayed arrow starting at $(0, 0)$. If $r < 0$, move in the opposite direction a distance of $|r|$. This is the point determined by (r, θ).

It is common to assume that θ is in the interval $[0, 2\pi)$ and $r > 0$. In this case, there is a very simple relationship between the Cartesian and polar coordinates, given by

$$x = r\cos(\theta), \quad y = r\sin(\theta) \tag{8.1.1}$$

These equations demonstrate how to find the Cartesian coordinates when we are given the polar coordinates of a point. They can also be used to find the polar coordinates when we know (x, y). A simpler way to do this is the following equations:

$$r = \sqrt{x^2 + y^2}$$

$$\tan(\theta) = \frac{y}{x} \tag{8.1.2}$$

In the next example, we look at how to find the Cartesian coordinates of a point specified by polar coordinates.

Example 8.1.1: Finding Cartesian Coordinates

The polar coordinates of a point in the plane are $(5, \pi/6)$. Find the Cartesian coordinates of this point.

Solution

The point is specified by the polar coordinates $(5, \pi/6)$. Therefore $r = 5$ and $\theta = \pi/6$. From (8.1.1)

$$x = r\cos(\theta) = 5\cos\left(\frac{\pi}{6}\right) = \frac{5}{2}\sqrt{3}$$

$$y = r\sin(\theta) = 5\sin\left(\frac{\pi}{6}\right) = \frac{5}{2}$$

Thus the Cartesian coordinates are $\left(\frac{5}{2}\sqrt{3}, \frac{5}{2}\right)$. The point is shown in the below graph.

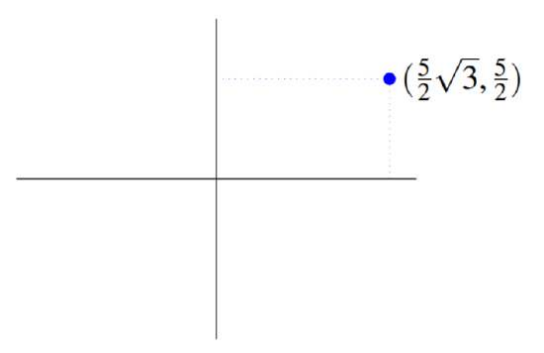


Figure 8.1.3

Consider the following example of the case where $r < 0$.

Example 8.1.2: Finding Cartesian Coordinates

The polar coordinates of a point in the plane are $(-5, \pi/6)$. Find the Cartesian coordinates.

Solution

For the point specified by the polar coordinates $(-5, \pi/6)$, $r = -5$, and $x\theta = \pi/6$. From (8.1.1)

$$x = r\cos(\theta) = -5\cos\left(\frac{\pi}{6}\right) = -\frac{5}{2}\sqrt{3}$$

$$y = r\sin(\theta) = -5\sin\left(\frac{\pi}{6}\right) = -\frac{5}{2}$$

Thus the Cartesian coordinates are $\left(-\frac{5}{2}\sqrt{3}, -\frac{5}{2}\right)$. The point is shown in the following graph.

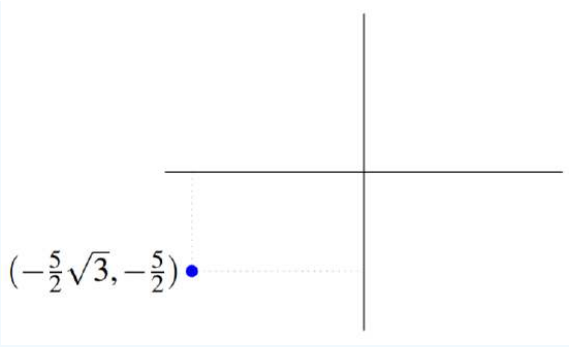


Figure 8.1.4

Recall from the previous example that for the point specified by $(5, \pi/6)$, the Cartesian coordinates are $\left(\frac{5}{2}\sqrt{3}, \frac{5}{2}\right)$. Notice that in this example, by multiplying r by -1, the resulting Cartesian coordinates are also multiplied by -1.

The following picture exhibits both points in the above two examples to emphasize how they are just on opposite sides of $(0, 0)$ but at the same distance from $(0, 0)$.

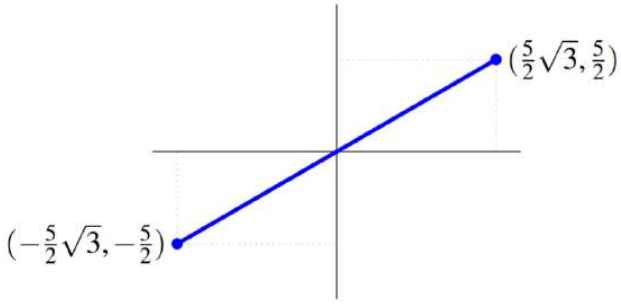


Figure 8.1.5

In the next two examples, we look at how to convert Cartesian coordinates to polar coordinates.

Example 8.1.3: Finding Polar Coordinates

Suppose the Cartesian coordinates of a point are $(3, 4)$. Find a pair of polar coordinates which correspond to this point.

Solution

Using equation (8.1.2), we can find r and θ. Hence $r = \sqrt{3^2 + 4^2} = 5$. It remains to identify the angle θ between the positive x axis and the line from the origin to the point. Since both the x and y values are positive, the point is in the first quadrant. Therefore, θ is between 0 and $\pi/2$. Using this and (8.1.2), we have to solve:

$$\tan(\theta) = \frac{4}{3}$$

Conversely, we can use equation (8.1.1) as follows:

$$3 = 5\cos(\theta)$$

$$4 = 5\sin(\theta)$$

Solving these equations, we find that, approximately, $\theta = 0.927\,295$ radians.

Consider the following example.

Example 8.1.4: Finding Polar Coordinates

Suppose the Cartesian coordinates of a point are $\left(-\sqrt{3},1\right)$. Find the polar coordinates which correspond to this point.

Solution

Given the point $\left(-\sqrt{3},1\right)$,

$$\begin{aligned} r &= \sqrt{1^2+\left(-\sqrt{3}\right)^2} \\ &= \sqrt{1+3} \\ &= 2 \end{aligned}$$

In this case, the point is in the second quadrant since the x value is negative and the y value is positive. Therefore, θ will be between $\pi/2$ and π. Solving the equations

$$-\sqrt{3} = 2\cos(\theta)$$

$$1 = 2\sin(\theta)$$

we find that $\theta = 5\pi/6$. Hence the polar coordinates for this point are $(2, 5\pi/6)$.

Consider this example. Suppose we used $r=-2$ and $\theta = 2\pi - (\pi/6) = 11\pi/6$. These coordinates specify the same point as above. Observe that there are infinitely many ways to identify this particular point with polar coordinates. In fact, every point can be represented with polar coordinates in infinitely many ways. Because of this, it will usually be the case that θ is confined to lie in some interval of length 2π and $r>0$, for real numbers r and θ.

Just as with Cartesian coordinates, it is possible to use relations between the polar coordinates to specify points in the plane. The process of sketching the graphs of these relations is very similar to that used to sketch graphs of functions in Cartesian coordinates. Consider a relation between polar coordinates of the form, $r = f(\theta)$. To graph such a relation, first make a table of the form

Table 8.1.1

θ	r
θ_1	$f(\theta_1)$
θ_2	$f(\theta_2)$
$\vdots$	$\vdots$

Graph the resulting points and connect them with a curve. The following picture illustrates how to begin this process.

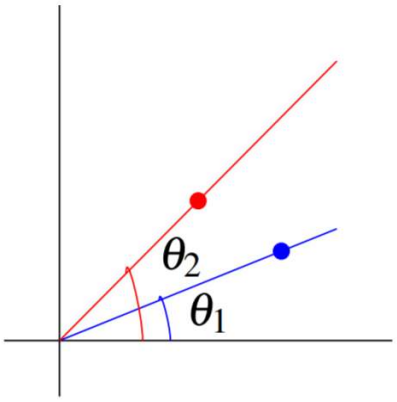


Figure 8.1.6

To find the point in the plane corresponding to the ordered pair $(f(\theta),\theta)$, we follow the same process as when finding the point corresponding to (r,θ).

Consider the following example of this procedure, incorporating computer software.

Example 8.1.5

Graphing a Polar Equation Graph the polar equation $r = 1 + \cos\theta$.

Solution

We will use the computer software *Maple* to complete this example. The command which produces the polar graph of the above equation is: > plot(1+cos(t),t= 0..2*Pi,coords=polar). Here we use t to represent the variable θ for convenience. The command tells Maple that r is given by $1 + \cos(t)$ and that $t \in [0, 2\pi]$.

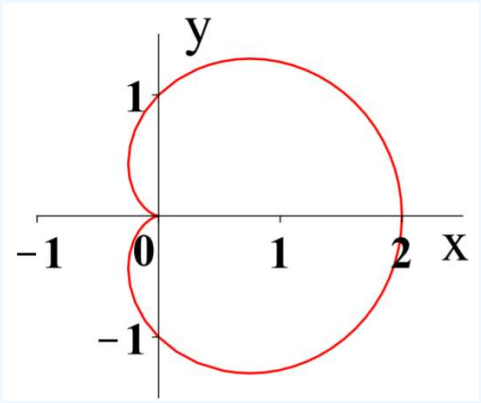


Figure 8.1.7

The above graph makes sense when considered in terms of trigonometric functions. Suppose $\theta = 0, r = 2$ and let θ increase to $\pi/2$. As θ increases, $\cos\theta$ decreases to 0. Thus the line from the origin to the point on the curve should get shorter as θ goes from 0 to $\pi/2$. As θ goes from $\pi/2$ to π, $\cos\theta$ decreases, eventually equaling -1 at $\theta = \pi$. Thus $r = 0$ at this point. This scenario is depicted in the above graph, which shows a function called a **cardioid**.

The following picture illustrates the above procedure for obtaining the polar graph of $r = 1 + cos(\theta)$. In this picture, the concentric circles correspond to values of r while the rays from the origin correspond to the angles which are shown on the picture. The dot on the ray corresponding to the angle $\pi/6$ is located at a distance of $r = 1 + cos(\pi/6)$ from the origin. The dot on the ray corresponding to the angle $\pi/3$ is located at a distance of $r = 1 + cos(\pi/3)$ from the origin and so forth. The polar graph is obtained by connecting such points with a smooth curve, with the result being the figure shown above.

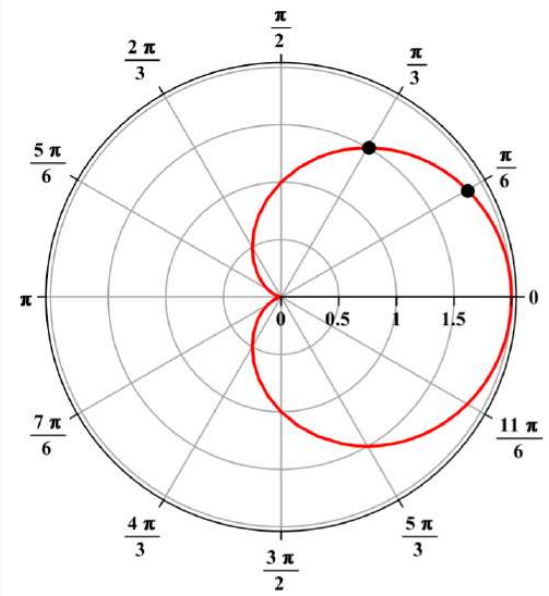


Figure 8.1.8

Consider another example of constructing a polar graph.

Example 8.1.6: A Polar Graph

Graph $r = 1 + 2\cos\theta$ for $\theta \in [0, 2\pi]$.

Solution

The graph of the polar equation $r = 1 + 2\cos\theta$ for $\theta \in [0, 2\pi]$ is given as follows.

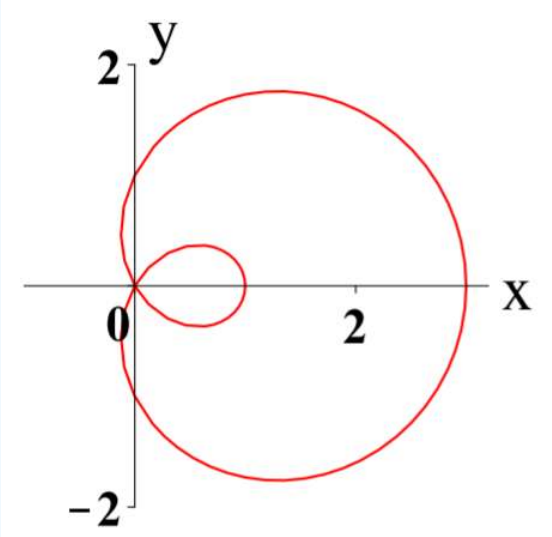


Figure 8.1.9

To see the way this is graphed, consider the following picture. First the indicated points were graphed and then the curve was drawn to connect the points. When done by a computer, many more points are used to create a more accurate picture.

Consider first the following table of points.

Table 8.1.2

θ	$\pi/6$	$\pi/3$	$\pi/2$	$5\pi/6$	π	$4\pi/3$	$7\pi/6$	$5\pi/3$
r	$\sqrt{3}+1$	2	1	$1-\sqrt{3}$	-1	0	$1-\sqrt{3}$	2

Note how some entries in the table have $r < 0$. To graph these points, simply move in the opposite direction. These types of points are responsible for the small loop on the inside of the larger loop in the graph.

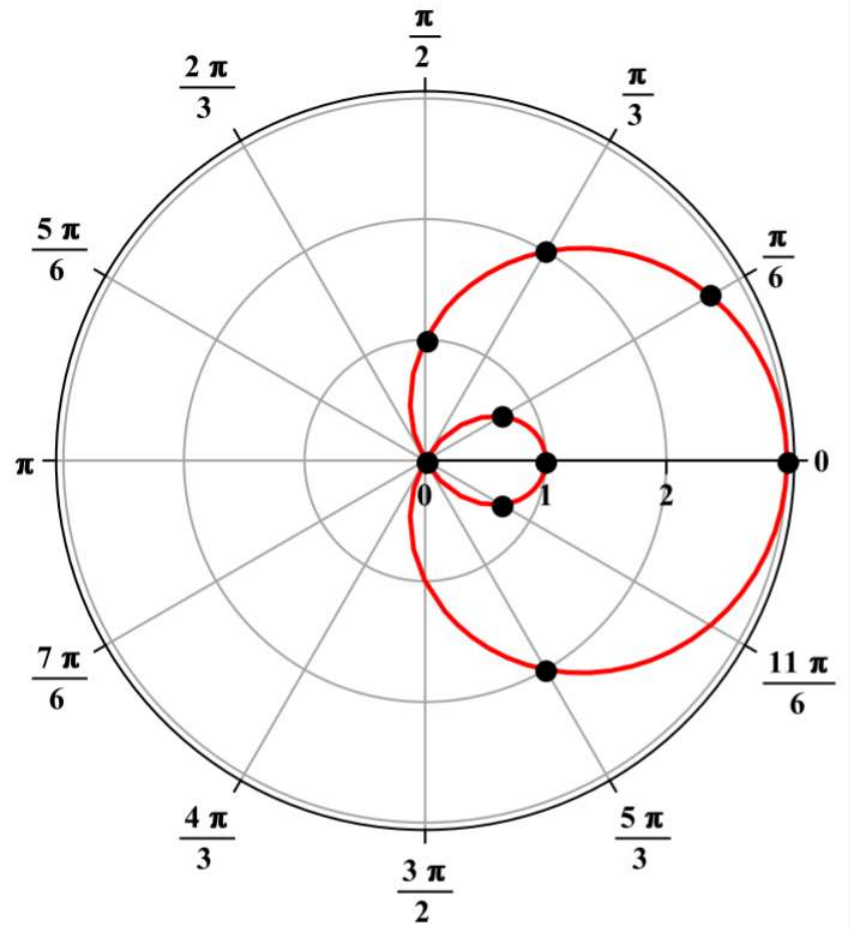


Figure 8.1.10

The process of constructing these graphs can be greatly facilitated by computer software. However, the use of such software should not replace understanding the steps involved.

The next example shows the graph for the equation $r = 3 + \sin\left(\frac{7\theta}{6}\right)$. For complicated polar graphs, computer software is used to facilitate the process.

Example 8.1.7: A Polar Graph

Graph $r = 3 + \sin\left(\frac{7\theta}{6}\right)$ for $\theta \in [0, 14\pi]$.

Solution

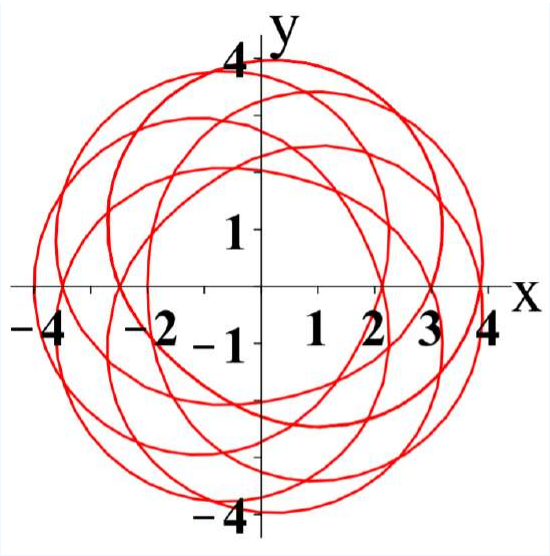


Figure 8.1.11

The next example shows another situation in which r can be negative.

Example 8.1.8: A Polar Graph: Negative r

Graph $r = 3sin(4\theta)$ for $\theta \in [0, 2\pi]$.

Solution

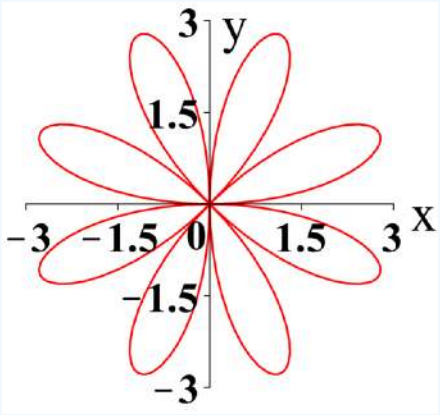


Figure 8.1.12

We conclude this section with an interesting graph of a simple polar equation.

Example 8.1.9: The Graph of a Spiral

Graph $r = \theta$ for $\theta \in [0, 2\pi]$.

Solution

The graph of this polar equation is a spiral. This is the case because as θ increases, so does r.

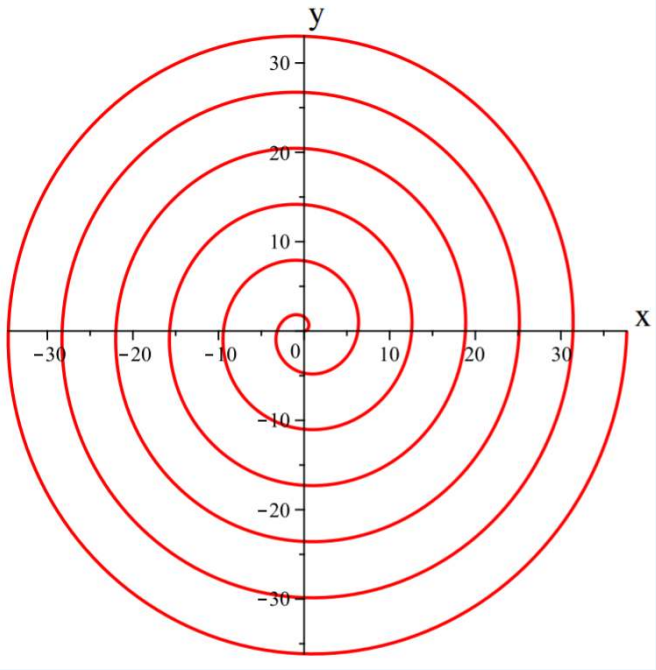


Figure 8.1.13

In the next section, we will look at two ways of generalizing polar coordinates to three dimensions.

This page titled 8.1: Polar Coordinates and Polar Graphs is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **8.1: Polar Coordinates and Polar Graphs** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

8.2: Spherical and Cylindrical Coordinates

Outcomes

A. Understand cylindrical and spherical coordinates.
B. Convert points between Cartesian, cylindrical, and spherical coordinates.

Spherical and cylindrical coordinates are two generalizations of polar coordinates to three dimensions. We will first look at **cylindrical coordinates** .

When moving from polar coordinates in two dimensions to cylindrical coordinates in three dimensions, we use the polar coordinates in the xy plane and add a z coordinate. For this reason, we use the notation (r, θ, z) to express cylindrical coordinates. The relationship between Cartesian coordinates (x, y, z) and cylindrical coordinates (r, θ, z) is given by

$$\begin{aligned} x &= r\cos(\theta) \\ y &= r\sin(\theta) \\ z &= z \end{aligned}$$

where $r \geq 0$, $\theta \in [0, 2\pi)$, and z is simply the Cartesian coordinate. Notice that x and y are defined as the usual polar coordinates in the xy-plane. Recall that r is defined as the length of the ray from the origin to the point $(x, y, 0)$, while θ is the angle between the positive x-axis and this same ray.

To illustrate this coordinate system, consider the following two pictures. In the first of these, both r and z are known. The cylinder corresponds to a given value for r. A useful way to think of r is as the distance between a point in three dimensions and the z-axis. Every point on the cylinder shown is at the same distance from the z-axis. Giving a value for z results in a horizontal circle, or cross section of the cylinder at the given height on the z axis (shown below as a black line on the cylinder). In the second picture, the point is specified completely by also knowing θ as shown.

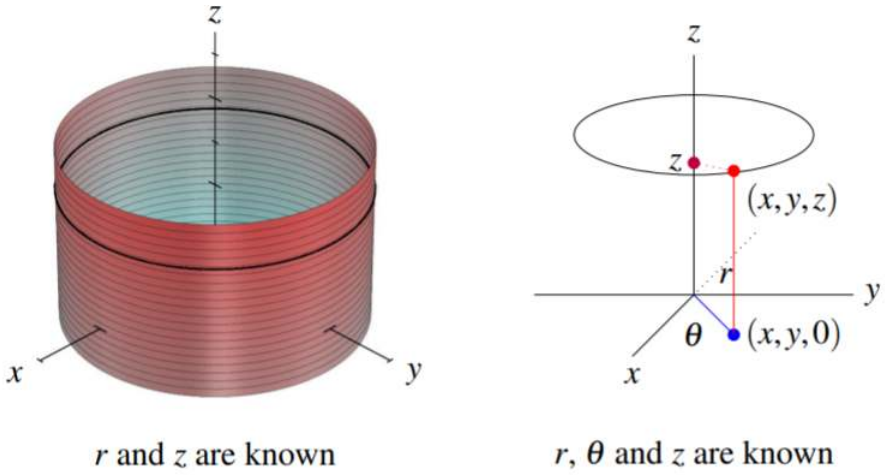


Figure 8.2.1

Every point of three dimensional space other than the z axis has unique cylindrical coordinates. Of course there are infinitely many cylindrical coordinates for the origin and for the z-axis. Any θ will work if $r = 0$ and z is given.

Consider now **spherical coordinates**, the second generalization of polar form in three dimensions. For a point (x, y, z) in three dimensional space, the spherical coordinates are defined as follows.

ρ : the length of the ray from the origin to the point
θ : the angle between the positive x-axis and the ray from the origin to the point $(x, y, 0)$
ϕ : the angle between the positive z-axis and the ray from the origin to the point of interest

The spherical coordinates are determined by (ρ, ϕ, θ). The relation between these and the Cartesian coordinates (x, y, z) for a point are as follows.

$$x = \rho \sin(\phi)\cos(\theta),\ \phi \in [0, \pi]$$
$$y = \rho \sin(\phi)\sin(\theta),\ \theta \in [0, 2\pi)$$
$$z = \rho \cos\phi,\ \rho \geq 0.$$

Consider the pictures below. The first illustrates the surface when ρ is known, which is a sphere of radius ρ. The second picture corresponds to knowing both ρ and ϕ, which results in a circle about the z-axis. Suppose the first picture demonstrates a graph of the Earth. Then the circle in the second picture would correspond to a particular latitude.

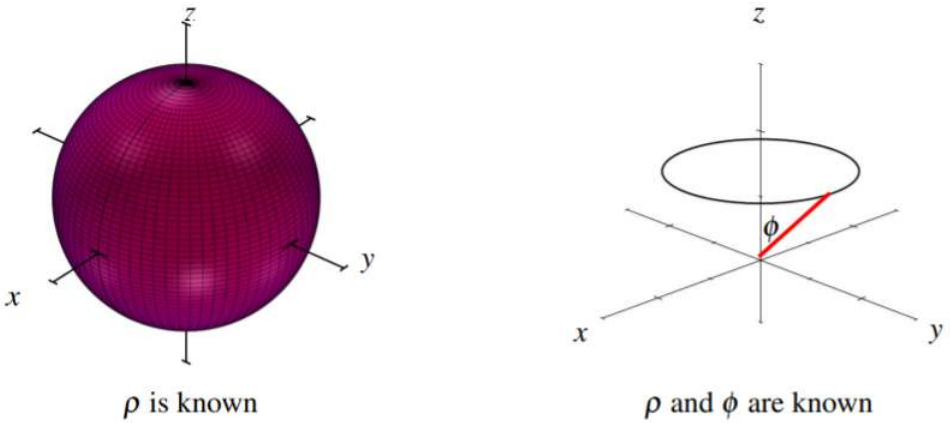


Figure 8.2.2

Giving the third coordinate, θ completely specifies the point of interest. This is demonstrated in the following picture. If the latitude corresponds to ϕ, then we can think of θ as the longitude.

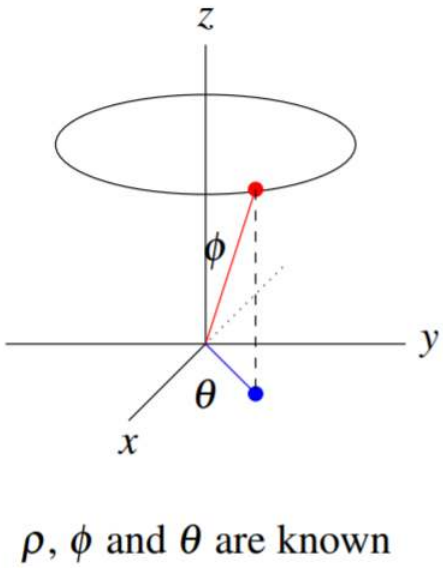


Figure 8.2.3

The following picture summarizes the geometric meaning of the three coordinate systems.

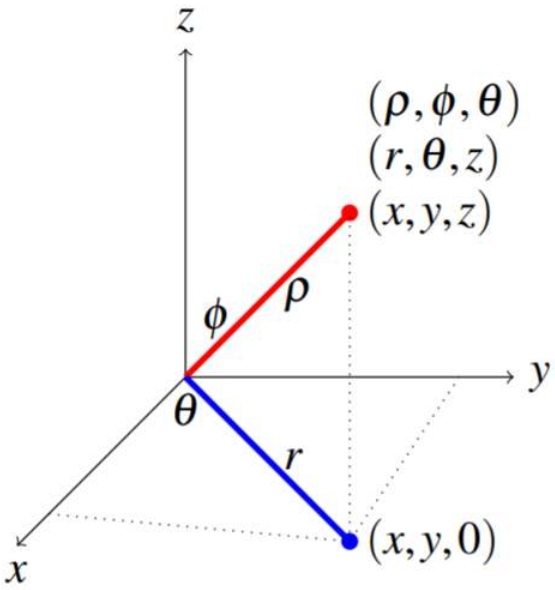


Figure 8.2.4

Therefore, we can represent the same point in three ways, using Cartesian coordinates, (x, y, z), cylindrical coordinates, (r, θ, z), and spherical coordinates (ρ, ϕ, θ).

Using this picture to review, call the point of interest P for convenience. The Cartesian coordinates for P are (x, y, z). Then ρ is the distance between the origin and the point P. The angle between the positive z axis and the line between the origin and P is denoted by ϕ. Then θ is the angle between the positive x axis and the line joining the origin to the point $(x, y, 0)$ as shown. This gives the spherical coordinates, (ρ, ϕ, θ). Given the line from the origin to $(x, y, 0)$, $r = \rho \sin(\phi)$ is the length of this line. Thus r and θ determine a point in the xy-plane. In other words, r and θ are the usual polar coordinates and $r \geq 0$ and $\theta \in [0, 2\pi)$. Letting z denote the usual z coordinate of a point in three dimensions, (r, θ, z) are the cylindrical coordinates of P.

The relation between spherical and cylindrical coordinates is that $r = \rho \sin(\phi)$ and the θ is the same as the θ of cylindrical and polar coordinates.

We will now consider some examples.

Example 8.2.1: Describing a Surface in Spherical Coordinates

Express the surface $z = \frac{1}{\sqrt{3}}\sqrt{x^2 + y^2}$ in spherical coordinates.

Solution

We will use the equations from above:

$$\begin{aligned} x &= \rho \sin(\phi) \cos(\theta), \phi \in [0, \pi] \\ y &= \rho \sin(\phi) \sin(\theta), \ \theta \in [0, 2\pi) \\ z &= \rho \cos \phi, \ \rho \geq 0 \end{aligned}$$

To express the surface in spherical coordinates, we substitute these expressions into the equation. This is done as follows:

$$\rho \cos(\phi) = \frac{1}{\sqrt{3}}\sqrt{\left(\rho \sin(\phi) \cos(\theta)\right)^2 + \left(\rho \sin(\phi) \sin(\theta)\right)^2} = \frac{1}{3}\sqrt{3}\rho \sin(\phi).$$

This reduces to

$$\tan(\phi) = \sqrt{3}$$

and so $\phi = \pi/3$.

Example 8.2.2: Describing a Surface in Spherical Coordinates

Express the surface $y = x$ in terms of spherical coordinates.

Solution

Using the same procedure as the previous example, this says $\rho \sin(\phi) \sin(\theta) = \rho \sin(\phi) \cos(\theta)$. Simplifying, $\sin(\theta) = \cos(\theta)$, which you could also write $\tan(\theta) = 1$.

We conclude this section with an example of how to describe a surface using cylindrical coordinates.

Example 8.2.3: Describing a Surface in Cylindrical Coordinates

Express the surface $x^2 + y^2 = 4$ in cylindrical coordinates.

Solution

Recall that to convert from Cartesian to cylindrical coordinates, we can use the following equations:

$$x = r\cos(\theta), y = r\sin(\theta), z = z$$

Substituting these equations in for x, y, z in the equation for the surface, we have

$$r^2 \cos^2(\theta) + r^2 \sin^2(\theta) = 4$$

This can be written as $r^2(\cos^2(\theta) + \sin^2(\theta)) = 4$. Recall that $\cos^2(\theta) + \sin^2(\theta) = 1$. Thus $r^2 = 4$ or $r = 2$.

This page titled 8.2: Spherical and Cylindrical Coordinates is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **8.2: Spherical and Cylindrical Coordinates** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

8.E: Exercises

Exercise 8.E.1

In the following, polar coordinates (r, θ) for a point in the plane are given. Find the corresponding Cartesian coordinates.

a. $(2, \pi/4)$
b. $(-2, \pi/4)$
c. $(3, \pi/3)$
d. $(-3, \pi/3)$
e. $(2, 5\pi/6)$
f. $(-2, 11\pi/6)$
g. $(2, \pi/2)$
h. $(1, 3\pi/2)$
i. $(-3, 3\pi/4)$
j. $(3, 5\pi/4)$
k. $(-2, \pi/6)$

Exercise 8.E.2

Consider the following Cartesian coordinates (x, y). Find polar coordinates corresponding to these points.

a. $(-1, 1)$
b. $(\sqrt{3}, -1)$
c. $(0, 2)$
d. $(-5, 0)$
e. $(-2\sqrt{3}, 2)$
f. $(2, -2)$
g. $(-1, \sqrt{3})$
h. $(-1, -\sqrt{3})$

Exercise 8.E.3

The following relations are written in terms of Cartesian coordinates (x, y). Rewrite them in terms of polar coordinates, (r, θ).

a. $y = x^2$
b. $y = 2x + 6$
c. $x^2 + y^2 = 4$
d. $x^2 - y^2 = 1$

Exercise 8.E.4

Use a calculator or computer algebra system to graph the following polar relations.

a. $r = 1 - \sin(2\theta),\ \theta \in [0, 2\pi]$
b. $r = \sin(4\theta),\ \theta \in [0, 2\pi]$
c. $r = \cos(3\theta) + \sin(2\theta),\ \theta \in [0, 2\pi]$
d. $r = \theta,\ \theta \in [0, 15]$

Exercise 8.E.5

Graph the polar equation $r = 1 + \sin\theta$ for $\theta \in [0, 2\pi]$.

Exercise 8.E.6

Graph the polar equation $r = 2 + \sin\theta$ for $\theta \in [0, 2\pi]$.

Exercise 8.E.7

Graph the polar equation $r = 1 + 2\sin\theta$ for $\theta \in [0, 2\pi]$.

Exercise 8.E.8

Graph the polar equation $r = 2 + \sin(2\theta)$ for $\theta \in [0, 2\pi]$.

Exercise 8.E.9

Graph the polar equation $r = 1 + \sin(2\theta)$ for $\theta \in [0, 2\pi]$.

Exercise 8.E.10

Graph the polar equation $r = 1 + \sin(3\theta)$ for $\theta \in [0, 2\pi]$.

Exercise 8.E.11

Describe how to solve for r and θ in terms of x and y in polar coordinates.

Exercise 8.E.12

This problem deals with parabolas, ellipses, and hyperbolas and their equations. Let l, $e > 0$ and consider

$$r = \frac{l}{1 \pm e\cos\theta}$$

Show that if $e = 0$, the graph of this equation gives a circle. Show that if $0 < e < 1$, the graph is an ellipse, if $e = 1$ it is a parabola and if $e > 1$, it is a hyperbola.

Exercise 8.E.13

The following are the cylindrical coordinates of points, (r, θ, z). Find the Cartesian and spherical coordinates of each point.

a. $(5, \frac{5\pi}{6}, -3)$
b. $(3, \frac{\pi}{3}, 4)$
c. $(4, \frac{2\pi}{3}, 1)$
d. $(2, \frac{3\pi}{4}, -2)$
e. $(3, \frac{3\pi}{2}, -1)$
f. $(8, \frac{11\pi}{6}, -11)$

Exercise 8.E.14

The following are the Cartesian coordinates of points, (x, y, z). Find the cylindrical and spherical coordinates of these points.

a. $(\frac{5}{2}\sqrt{2}, \frac{5}{2}\sqrt{2}, -3)$
b. $(\frac{3}{2}, \frac{3}{2}\sqrt{3}, 2)$
c. $(-\frac{5}{2}\sqrt{2}, \frac{5}{2}\sqrt{2}, 11)$
d. $(-\frac{5}{2}, \frac{5}{2}\sqrt{3}, 23)$
e. $(-\sqrt{3}, -1, -5)$

f. $(\frac{3}{2}, -\frac{3}{2}\sqrt{3}, -7)$
g. $(\sqrt{2}, \sqrt{6}, 2\sqrt{2})$
h. $(-\frac{1}{2}\sqrt{3}, \frac{3}{2}, 1)$
i. $(-\frac{3}{4}\sqrt{2}, \frac{3}{4}\sqrt{2}, -\frac{3}{2}\sqrt{3})$
j. $(-\sqrt{31}, 2\sqrt{3})$
k. $(-\frac{1}{4}\sqrt{2}, \frac{1}{4}\sqrt{6}, -\frac{1}{2}\sqrt{2})$

Exercise 8.E.15

The following are spherical coordinates of points in the form (ρ, φ, θ). Find the Cartesian and cylindrical coordinates of each point.

a. $(4, \frac{\pi}{4}, \frac{5\pi}{6})$
b. $(2, \frac{\pi}{3}, \frac{2\pi}{3})$
c. $(3, \frac{5\pi}{6}, \frac{3\pi}{2})$
d. $(4, \frac{\pi}{2}, \frac{7\pi}{4})$
e. $(4, \frac{2\pi}{3}, \frac{\pi}{6})$
f. $(4, \frac{3\pi}{4}, \frac{5\pi}{3})$

Exercise 8.E.16

Describe the surface $\varphi = \pi/4$ in Cartesian coordinates, where φ is the polar angle in spherical coordinates.

Exercise 8.E.17

Describe the surface $\theta = \pi/4$ in spherical coordinates, where θ is the angle measured from the positive x axis.

Exercise 8.E.18

Describe the surface $r = 5$ in Cartesian coordinates, where r is one of the cylindrical coordinates.

Exercise 8.E.19

Describe the surface $\rho = 4$ in Cartesian coordinates, where ρ is the distance to the origin.

Exercise 8.E.20

Give the cone described by $z = \sqrt{x^2 + y^2}$ in cylindrical coordinates and in spherical coordinates.

Exercise 8.E.21

The following are described in Cartesian coordinates. Rewrite them in terms of spherical coordinates.

a. $z = x^2 + y^2$
b. $x^2 - y^2 = 1$
c. $z^2 + x^2 + y^2 = 6$
d. $z = \sqrt{x^2 + y^2}$
e. $y = x$
f. $z = x$

Exercise $8.E.22$

The following are described in Cartesian coordinates. Rewrite them in terms of cylindrical coordinates.

a. $z = x^2 + y^2$
b. $x^2 - y^2 = 1$
c. $z^2 + x^2 + y^2 = 6$
d. $z = \sqrt{x^2 + y^2}$
e. $y = x$
f. $z = x$

This page titled 8.E: Exercises is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

CHAPTER OVERVIEW

9: Vector Spaces

9.1: Algebraic Considerations
9.2: Spanning Sets
9.3: Linear Independence
9.4: Subspaces and Basis
9.5: Sums and Intersections
9.6: Linear Transformations
9.7: Isomorphisms
9.8: The Kernel and Image of a Linear Map
9.9: The Matrix of a Linear Transformation
9.E: Exercises

This page titled 9: Vector Spaces is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

9.1: Algebraic Considerations

Outcomes

A. Develop the abstract concept of a vector space through axioms.
B. Deduce basic properties of vector spaces.
C. Use the vector space axioms to determine if a set and its operations constitute a vector space.

In this section we consider the idea of an abstract vector space. A vector space is something which has two operations satisfying the following vector space axioms.

Definition 9.1.1: Vector Space

A vector space V is a set of vectors with two operations defined, addition and scalar multiplication, which satisfy the axioms of addition and scalar multiplication.

In the following definition we define two operations; vector addition, denoted by $+$ and scalar multiplication denoted by placing the scalar next to the vector. A vector space need not have usual operations, and for this reason the operations will always be given in the definition of the vector space. The below axioms for addition (written +) and scalar multiplication must hold for however addition and scalar multiplication are defined for the vector space.

It is important to note that we have seen much of this content before, in terms of $\mathbb{R}^n$. We will prove in this section that $\mathbb{R}^n$ is an example of a vector space and therefore all discussions in this chapter will pertain to $\mathbb{R}^n$. While it may be useful to consider all concepts of this chapter in terms of $\mathbb{R}^n$, it is also important to understand that these concepts apply to *all* vector spaces.

In the following definition, we will choose scalars a, b to be real numbers and are thus dealing with *real* vector spaces. However, we could also choose scalars which are complex numbers. In this case, we would call the vector space V complex.

Definition 9.1.2: Axioms of Addition

Let $\vec{v}, \vec{w}, \vec{z}$ be vectors in a vector space V. Then they satisfy the following axioms of addition:

- Closed under Addition

$$\text{If } \vec{v}, \vec{w} \text{ are in } V, \text{ then } \vec{v}+\vec{w} \text{ is also in } V.$$

- The Commutative Law of Addition

$$\vec{v}+\vec{w}=\vec{w}+\vec{v}$$

- The Associative Law of Addition

$$(\vec{v}+\vec{w})+\vec{z}=\vec{v}+(\vec{w}+\vec{z})$$

- The Existence of an Additive Identity

$$\vec{v}+\vec{0}=\vec{v}$$

- The Existence of an Additive Inverse

$$\vec{v}+(-\vec{v})=\vec{0}$$

vectorspaceaxiomsaddition

Definition 9.1.3: Axioms of Scalar Multiplication

Let $a, b \in \mathbb{R}$ and let $\vec{v}, \vec{w}, \vec{z}$ be vectors in a vector space V. Then they satisfy the following axioms of scalar multiplication:

- Closed under Scalar Multiplication

If a is a real number, and $\vec{v}$ is in V, then $a\vec{v}$ is in V.

- $a\left(\vec{v}+\vec{w}\right)=a\vec{v}+a\vec{w}$
- $\left(a+b\right)\vec{v}=a\vec{v}+b\vec{v}$
- $a\left(b\vec{v}\right)=\left(ab\right)\vec{v}$
- $1\vec{v}=\vec{v}$

Below is a video on what a vector space is.



Consider the following example, in which we prove that $\mathbb{R}^n$ is in fact a vector space.

Example 9.1.1: $\mathbb{R}^n$

$\mathbb{R}^n$, under the usual operations of vector addition and scalar multiplication, is a vector space.

Solution

To show that $\mathbb{R}^n$ is a vector space, we need to show that the above axioms hold. Let $\vec{x},\vec{y},\vec{z}$ be vectors in $\mathbb{R}^n$. We first prove the axioms for vector addition.

- To show that $\mathbb{R}^n$ is closed under addition, we must show that for two vectors in $\mathbb{R}^n$ their sum is also in $\mathbb{R}^n$. The sum $\vec{x}+\vec{y}$ is given by:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix}$$

  The sum is a vector with n entries, showing that it is in $\mathbb{R}^n$. Hence $\mathbb{R}^n$ is closed under vector addition.
- To show that addition is commutative, consider the following:

$$\begin{aligned}\vec{x}+\vec{y} &= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \\ &= \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} \\ &= \begin{bmatrix} y_1+x_1 \\ y_2+x_2 \\ \vdots \\ y_n+x_n \end{bmatrix} \\ &= \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \vec{y}+\vec{x}\end{aligned}$$

Hence addition of vectors in $\mathbb{R}^n$ is commutative.

- We will show that addition of vectors in $\mathbb{R}^n$ is associative in a similar way.

$$\begin{aligned}(\vec{x}+\vec{y})+\vec{z} &= \left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right) + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \\ &= \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \\ &= \begin{bmatrix} (x_1+y_1)+z_1 \\ (x_2+y_2)+z_2 \\ \vdots \\ (x_n+y_n)+z_n \end{bmatrix} \\ &= \begin{bmatrix} x_1+(y_1+z_1) \\ x_2+(y_2+z_2) \\ \vdots \\ x_n+(y_n+z_n) \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1+z_1 \\ y_2+z_2 \\ \vdots \\ y_n+z_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \left(\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \right) \\ &= \vec{x}+(\vec{y}+\vec{z})\end{aligned}$$

Hence addition of vectors is associative.

- Next, we show the existence of an additive identity. Let $\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

$$\begin{aligned} \vec{x} + \vec{0} &= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} x_1 + 0 \\ x_2 + 0 \\ \vdots \\ x_n + 0 \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \vec{x} \end{aligned}$$

Hence the zero vector $\vec{0}$ is an additive identity.

- Next, we prove the existence of an additive inverse. Let $-\vec{x} = \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix}$.

$$\begin{aligned} \vec{x} + (-\vec{x}) &= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{bmatrix} \\ &= \begin{bmatrix} x_1 - x_1 \\ x_2 - x_2 \\ \vdots \\ x_n - x_n \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \\ &= \vec{0} \end{aligned}$$

Hence $-\vec{x}$ is an additive inverse.

We now need to prove the axioms related to scalar multiplication. Let a, b be real numbers and let $\vec{x}, \vec{y}$ be vectors in $\mathbb{R}^n$.

- We first show that $\mathbb{R}^n$ is closed under scalar multiplication. To do so, we show that $a\vec{x}$ is also a vector with n entries.

$$a\vec{x} = a \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix}$$

The vector $a\vec{x}$ is again a vector with n entries, showing that $\mathbb{R}^n$ is closed under scalar multiplication.

- We wish to show that $a(\vec{x}+\vec{y}) = a\vec{x}+a\vec{y}$.

$$\begin{aligned} a(\vec{x}+\vec{y}) &= a\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} x_1 \\ x_2 \\ \vdots x_n \end{bmatrix}\right) \\ &= a\begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} \\ &= \begin{bmatrix} a(x_1+y_1) \\ a(x_2+y_2) \\ \vdots \\ a(x_n+y_n) \end{bmatrix} \\ &= \begin{bmatrix} ax_1+ay_1 \\ ax_2+ay_2 \\ \vdots \\ ax_n+ay_n \end{bmatrix} \\ &= \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix} + \begin{bmatrix} ay_1 \\ ay_2 \\ \vdots \\ ay_n \end{bmatrix} \\ &= a\vec{x}+a\vec{y} \end{aligned}$$

- Next, we wish to show that $(a+b)\vec{x} = a\vec{x}+b\vec{x}$.

$$\begin{aligned} (a+b)\vec{x} &= (a+b)\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} (a+b)x_1 \\ (a+b)x_2 \\ \vdots \\ (a+b)x_n \end{bmatrix} \\ &= \begin{bmatrix} ax_1+bx_1 \\ ax_2+bx_2 \\ \vdots \\ ax_n+bx_n \end{bmatrix} \\ &= \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{bmatrix} + \begin{bmatrix} bx_1 \\ bx_2 \\ \vdots \\ bx_n \end{bmatrix} \\ &= a\vec{x}+b\vec{x} \end{aligned}$$

- We wish to show that $a(b\vec{x}) = (ab)\vec{x}$.

$$
\begin{aligned}
a(b\vec{x}) &= a\left(b\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) \\
&= a\left(\begin{bmatrix} bx_1 \\ bx_2 \\ \vdots \\ bx_n \end{bmatrix}\right) \\
&= \begin{bmatrix} a(bx_1) \\ a(bx_2) \\ \vdots \\ a(bx_n) \end{bmatrix} \\
&= \begin{bmatrix} (ab)x_1 \\ (ab)x_2 \\ \vdots \\ (ab)x_n \end{bmatrix} \\
&= (ab)\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= (ab)\vec{x}
\end{aligned}
$$

- Finally, we need to show that $1\vec{x} = \vec{x}$.

$$
\begin{aligned}
1\vec{x} &= 1\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} 1x_1 \\ 1x_2 \\ \vdots \\ 1x_n \end{bmatrix} \\
&= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \vec{x}
\end{aligned}
$$

By the above proofs, it is clear that $\mathbb{R}^n$ satisfies the vector space axioms. Hence, $\mathbb{R}^n$ is a vector space under the usual operations of vector addition and scalar multiplication.

Below is a video on proving R^n is a vector space.



We now consider some examples of vector spaces.

Example 9.1.2: Vector Space of Polynomials

Let $\mathbb{P}_2$ be the set of all polynomials of at most degree 2 as well as the zero polynomial. Define addition to be the standard addition of polynomials, and scalar multiplication the usual multiplication of a polynomial by a number. Then $\mathbb{P}_2$ is a vector space.

Solution

We can write $\mathbb{P}_2$ explicitly as

$$\mathbb{P}_2 = \left\{ a_2x^2 + a_1x + a_0 | a_i \in \mathbb{R} \text{ for all } i \right\}$$

To show that $\mathbb{P}_2$ is a vector space, we verify the axioms. Let $p(x), q(x), r(x)$ be polynomials in $\mathbb{P}_2$ and let a, b, c be real numbers. Write $p(x) = p_2x^2 + p_1x + p_0$, $q(x) = q_2x^2 + q_1x + q_0$, and $r(x) = r_2x^2 + r_1x + r_0$.

- We first prove that addition of polynomials in $\mathbb{P}_2$ is closed. For two polynomials in $\mathbb{P}_2$ we need to show that their sum is also a polynomial in $\mathbb{P}_2$. From the definition of $\mathbb{P}_2$, a polynomial is contained in $\mathbb{P}_2$ if it is of degree at most 2 or the zero polynomial.

$$\begin{aligned} p(x) + q(x) &= p_2x^2 + p_1x + p_0 + q_2x^2 + q_1x + q_0 \\ &= (p_2 + q_2)x^2 + (p_1 + q_1)x + (p_0 + q_0) \end{aligned}$$

  The sum is a polynomial of degree 2 and therefore is in $\mathbb{P}_2$. It follows that $\mathbb{P}_2$ is closed under addition.
- We need to show that addition is commutative, that is $p(x) + q(x) = q(x) + p(x)$.

$$\begin{aligned} p(x) + q(x) &= p_2x^2 + p_1x + p_0 + q_2x^2 + q_1x + q_0 \\ &= (p_2 + q_2)x^2 + (p_1 + q_1)x + (p_0 + q_0) \\ &= (q_2 + p_2)x^2 + (q_1 + p_1)x + (q_0 + p_0) \\ &= q_2x^2 + q_1x + q_0 + p_2x^2 + p_1x + p_0 \\ &= q(x) + p(x) \end{aligned}$$

- Next, we need to show that addition is associative. That is, that $(p(x) + q(x)) + r(x) = p(x) + (q(x) + r(x))$.

$$\begin{aligned} (p(x) + q(x)) + r(x) &= \left(p_2x^2 + p_1x + p_0 + q_2x^2 + q_1x + q_0\right) + r_2x^2 + r_1x + r_0 \\ &= (p_2 + q_2)x^2 + (p_1 + q_1)x + (p_0 + q_0) + r_2x^2 + r_1x + r_0 \\ &= (p_2 + q_2 + r_2)x^2 + (p_1 + q_1 + r_1)x + (p_0 + q_0 + r_0) \\ &= p_2x^2 + p_1x + p_0 + (q_2 + r_2)x^2 + (q_1 + r_1)x + (q_0 + r_0) \\ &= p_2x^2 + p_1x + p_0 + \left(q_2x^2 + q_1x + q_0 + r_2x^2 + r_1x + r_0\right) \\ &= p(x) + (q(x) + r(x)) \end{aligned}$$

- Next, we must prove that there exists an additive identity. Let $0(x) = 0x^2 + 0x + 0$.

$$\begin{aligned} p(x)+0(x) &= p_2x^2+p_1x+p_0+0x^2+0x+0 \\ &= (p_2+0)x^2+(p_1+0)x+(p_0+0) \\ &= p_2x^2+p_1x+p_0 \\ &= p(x) \end{aligned}$$

Hence an additive identity exists, specifically the zero polynomial.

- Next we must prove that there exists an additive inverse. Let $-p(x)=-p_2x^2-p_1x-p_0$ and consider the following:

$$\begin{aligned} p(x)+(-p(x)) &= p_2x^2+p_1x+p_0+\left(-p_2x^2-p_1x-p_0\right) \\ &= (p_2-p_2)x^2+(p_1-p_1)x+(p_0-p_0) \\ &= 0x^2+0x+0 \\ &= 0(x) \end{aligned}$$

Hence an additive inverse $-p(x)$ exists such that $p(x)+(-p(x))=0(x)$.

We now need to verify the axioms related to scalar multiplication.

- First we prove that $\mathbb{P}_2$ is closed under scalar multiplication. That is, we show that $ap(x)$ is also a polynomial of degree at most 2.

$$ap(x)=a\left(p_2x^2+p_1x+p_0\right)=ap_2x^2+ap_1x+ap_0$$

Therefore $\mathbb{P}_2$ is closed under scalar multiplication.

- We need to show that $a(p(x)+q(x))=ap(x)+aq(x)$.

$$\begin{aligned} a(p(x)+q(x)) &= a\left(p_2x^2+p_1x+p_0+q_2x^2+q_1x+q_0\right) \\ &= a\left((p_2+q_2)x^2+(p_1+q_1)x+(p_0+q_0)\right) \\ &= a(p_2+q_2)x^2+a(p_1+q_1)x+a(p_0+q_0) \\ &= (ap_2+aq_2)x^2+(ap_1+aq_1)x+(ap_0+aq_0) \\ &= ap_2x^2+ap_1x+ap_0+aq_2x^2+aq_1x+aq_0 \\ &= ap(x)+aq(x) \end{aligned}$$

- Next we show that $(a+b)p(x)=ap(x)+bp(x)$.

$$\begin{aligned} (a+b)p(x) &= (a+b)(p_2x^2+p_1x+p_0) \\ &= (a+b)p_2x^2+(a+b)p_1x+(a+b)p_0 \\ &= ap_2x^2+ap_1x+ap_0+bp_2x^2+bp_1x+bp_0 \\ &= ap(x)+bp(x) \end{aligned}$$

- The next axiom which needs to be verified is $a(bp(x))=(ab)p(x)$.

$$\begin{aligned} a(bp(x)) &= a\left(b\left(p_2x^2+p_1x+p_0\right)\right) \\ &= a\left(bp_2x^2+bp_1x+bp_0\right) \\ &= abp_2x^2+abp_1x+abp_0 \\ &= (ab)(p_2x^2+p_1x+p_0) \\ &= (ab)p(x) \end{aligned}$$

- Finally, we show that $1p(x)=p(x)$.

$$\begin{aligned} 1p(x) &= 1\left(p_2x^2+p_1x+p_0\right) \\ &= 1p_2x^2+1p_1x+1p_0 \\ &= p_2x^2+p_1x+p_0 \\ &= p(x) \end{aligned}$$

Since the above axioms hold, we know that $\mathbb{P}_2$ as described above is a vector space.

Below is a video on which sets are vector spaces.



Another important example of a vector space is the set of all matrices of the same size.

Example 9.1.3: Vector Space of Matrices

Let $\mathbb{M}_{2,3}$ be the set of all 2×3 matrices. Using the usual operations of matrix addition and scalar multiplication, show that $\mathbb{M}_{2,3}$ is a vector space.

Solution

Let A, B be 2×3 matrices in $\mathbb{M}_{2,3}$. We first prove the axioms for addition.

- In order to prove that $\mathbb{M}_{2,3}$ is closed under matrix addition, we show that the sum $A+B$ is in $\mathbb{M}_{2,3}$. This means showing that $A+B$ is a 2×3 matrix.

$$\begin{aligned} A+B &= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{bmatrix} \end{aligned}$$

  You can see that the sum is a 2×3 matrix, so it is in $\mathbb{M}_{2,3}$. It follows that $\mathbb{M}_{2,3}$ is closed under matrix addition.
- The remaining axioms regarding matrix addition follow from properties of matrix addition. Therefore $\mathbb{M}_{2,3}$ satisfies the axioms of matrix addition.

We now turn our attention to the axioms regarding scalar multiplication. Let A, B be matrices in $\mathbb{M}_{2,3}$ and let c be a real number.

- We first show that $\mathbb{M}_{2,3}$ is closed under scalar multiplication. That is, we show that cA a 2×3 matrix.

$$\begin{aligned} cA &= c \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \\ &= \begin{bmatrix} ca_{11} & ca_{12} & ca_{13} \\ ca_{21} & ca_{22} & ca_{23} \end{bmatrix} \end{aligned}$$

  This is a 2×3 matrix in $\mathbb{M}_{2,3}$ which proves that the set is closed under scalar multiplication.
- The remaining axioms of scalar multiplication follow from properties of scalar multiplication of matrices. Therefore $\mathbb{M}_{2,3}$ satisfies the axioms of scalar multiplication.

In conclusion, $\mathbb{M}_{2,3}$ satisfies the required axioms and is a vector space.

While here we proved that the set of all 2×3 matrices is a vector space, there is nothing special about this choice of matrix size. In fact if we instead consider $\mathbb{M}_{m,n}$, the set of all $m \times n$ matrices, then $\mathbb{M}_{m,n}$ is a vector space under the operations of matrix addition and scalar multiplication.

Below is a video on finding the zero vector, the multiplicative inverse, and the additive inverse of the vector space M_{2x2}.



We now examine an example of a set that does not satisfy all of the above axioms, and is therefore *not* a vector space.

Example 9.1.4: Not a Vector Space

Let V denote the set of 2×3 matrices. Let addition in V be defined by $A+B=A$ for matrices A, B in V. Let scalar multiplication in V be the usual scalar multiplication of matrices. Show that V is not a vector space.

Solution

In order to show that V is not a vector space, it suffices to find only one axiom which is not satisfied. We will begin by examining the axioms for addition until one is found which does not hold. Let A, B be matrices in V.

- We first want to check if addition is closed. Consider $A+B$. By the definition of addition in the example, we have that $A+B=A$. Since A is a 2×3 matrix, it follows that the sum $A+B$ is in V, and V is closed under addition.
- We now wish to check if addition is commutative. That is, we want to check if $A+B=B+A$ **for all** choices of A and B in V. From the definition of addition, we have that $A+B=A$ and $B+A=B$. Therefore, we can find A, B in V such that these sums are not equal. One example is

$$A=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B=\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Using the operation defined by $A+B=A$, we have

$$\begin{aligned} A+B &= A \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ B+A &= B \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{aligned}$$

It follows that $A+B \neq B+A$. Therefore addition as defined for V is not commutative and V fails this axiom. Hence V is not a vector space.

Below is a video on determining if the vectors in R^2 form a vector space with refined addition.



Consider another example of a vector space.

Example 9.1.5: Vector Space of Functions

Let S be a nonempty set and define $\mathbb{F}_S$ to be the set of *real* functions defined on S. In other words, we write $\mathbb{F}_S : S \mapsto \mathbb{R}$. Letting a, b, c be scalars and f, g, h functions, the vector operations are defined as

$$\begin{aligned}(f+g)(x) &= f(x)+g(x)\\ (af)(x) &= a(f(x))\end{aligned}$$

Show that $\mathbb{F}_S$ is a vector space.

Solution

To verify that $\mathbb{F}_S$ is a vector space, we must prove the axioms beginning with those for addition. Let f, g, h be functions in $\mathbb{F}_S$.

- First we check that addition is closed. For functions f, g defined on the set S, their sum given by

  $$(f+g)(x) = f(x)+g(x)$$

  is again a function defined on S. Hence this sum is in $\mathbb{F}_S$ and $\mathbb{F}_S$ is closed under addition.
- Secondly, we check the commutative law of addition:

  $$(f+g)(x) = f(x)+g(x) = g(x)+f(x) = (g+f)(x)$$

  Since x is arbitrary, $f+g = g+f$.
- Next we check the associative law of addition:

  $$\begin{aligned}((f+g)+h)(x) &= (f+g)(x)+h(x) = (f(x)+g(x))+h(x)\\ &= f(x)+(g(x)+h(x)) = (f(x)+(g+h)(x)) = (f+(g+h))(x)\end{aligned}$$

  and so $(f+g)+h = f+(g+h)$.
- Next we check for an additive identity. Let 0 denote the function which is given by $0(x) = 0$. Then this is an additive identity because

  $$(f+0)(x) = f(x)+0(x) = f(x)$$

  and so $f+0 = f$.
- Finally, check for an additive inverse. Let $-f$ be the function which satisfies $(-f)(x) = -f(x)$. Then

  $$(f+(-f))(x) = f(x)+(-f)(x) = f(x)+-f(x) = 0$$

  Hence $f+(-f) = 0$.

Now, check the axioms for scalar multiplication.

- We first need to check that $\mathbb{F}_S$ is closed under scalar multiplication. For a function $f(x)$ in $\mathbb{F}_S$ and real number a, the function $(af)(x) = a(f(x))$ is again a function defined on the set S. Hence $a(f(x))$ is in $\mathbb{F}_S$ and $\mathbb{F}_S$ is closed under scalar multiplication.
- $$((a+b)f)(x) = (a+b)f(x) = af(x) + bf(x) = (af+bf)(x)$$
  and so $(a+b)f = af+bf$.
- $$(a(f+g))(x) = a(f+g)(x) = a(f(x)+g(x))$$
  $$= af(x) + bg(x) = (af+bg)(x)$$
  and so $a(f+g) = af+bg$.
- $$((ab)f)(x) = (ab)f(x) = a(bf(x)) = (a(bf))(x)$$
  so $(abf) = a(bf)$.
- Finally $(1f)(x) = 1f(x) = f(x)$ so $1f = f$.

It follows that V satisfies all the required axioms and is a vector space.

Consider the following important theorem.

Theorem 9.1.1: Uniqueness

In any vector space, the following are true:

1. $\vec{0}$, the additive identity, is unique
2. $-\vec{x}$, the additive inverse, is unique
3. $0\vec{x} = \vec{0}$ for all vectors $\vec{x}$
4. $(-1)\vec{x} = -\vec{x}$ for all vectors $\vec{x}$

Proof

1. When we say that the additive identity, $\vec{0}$, is unique, we mean that if a vector acts like the additive identity, then it **is** the additive identity. To prove this uniqueness, we want to show that another vector which acts like the additive identity is actually equal to $\vec{0}$. Suppose $\vec{0}'$ is also an additive identity. Then,
   $$\vec{0} + \vec{0}' = \vec{0}$$
   Now, for $\vec{0}$ the additive identity given above in the axioms, we have that
   $$\vec{0}' + \vec{0} = \vec{0}'$$
   So by the commutative property:
   $$0 = 0 + 0' = 0' + 0 = 0'$$
   This says that if a vector acts like an additive identity (such as $\vec{0}'$), it in fact equals $\vec{0}$. This proves the uniqueness of $\vec{0}$.
2. When we say that the additive inverse, $-\vec{x}$, is unique, we mean that if a vector acts like the additive inverse, then it **is** the additive inverse. Suppose that $\vec{y}$ acts like an additive inverse:
   $$\vec{x} + \vec{y} = \vec{0}$$
   Then the following holds:
   $$\vec{y} = \vec{0} + \vec{y} = (-\vec{x} + \vec{x}) + \vec{y} = -\vec{x} + (\vec{x} + \vec{y}) = -\vec{x} + \vec{0} = -\vec{x}$$
   Thus if $\vec{y}$ acts like the additive inverse, it is equal to the additive inverse $-\vec{x}$. This proves the uniqueness of $-\vec{x}$.
3. This statement claims that for all vectors $\vec{x}$, scalar multiplication by 0 equals the zero vector $\vec{0}$. Consider the following, using the fact that we can write $0 = 0 + 0$:

$$0\vec{x} = (0+0)\,\vec{x} = 0\vec{x} + 0\vec{x}$$

We use a small trick here: add $-0\vec{x}$ to both sides. This gives

$$\begin{aligned} 0\vec{x} + (-0\vec{x}) &= 0\vec{x} + 0\vec{x} + (-\vec{x}) \\ \vec{0} + 0 &= 0\vec{x} + 0 \\ \vec{0} &= 0\vec{x} \end{aligned}$$

This proves that scalar multiplication of any vector by 0 results in the zero vector $\vec{0}$.

4. Finally, we wish to show that scalar multiplication of -1 and any vector $\vec{x}$ results in the additive inverse of that vector, $-\vec{x}$. Recall from 2. above that the additive inverse is unique. Consider the following:

$$\begin{aligned} (-1)\,\vec{x} + \vec{x} &= (-1)\,\vec{x} + 1\vec{x} \\ &= (-1+1)\,\vec{x} \\ &= 0\vec{x} \\ &= \vec{0} \end{aligned}$$

By the uniqueness of the additive inverse shown earlier, any vector which acts like the additive inverse must be equal to the additive inverse. It follows that $(-1)\,\vec{x} = -\vec{x}$.

An important use of the additive inverse is the following theorem.

Theorem 9.1.2

Let V be a vector space. Then $\vec{v} + \vec{w} = \vec{v} + \vec{z}$ implies that $\vec{w} = \vec{z}$ for all $\vec{v}, \vec{w}, \vec{z} \in V$

Proof

The proof follows from the vector space axioms, in particular the existence of an additive inverse $(-\vec{u})$. The proof is left as an exercise to the reader.

This page titled 9.1: Algebraic Considerations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **9.1: Algebraic Considerations** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

9.2: Spanning Sets

Outcomes

A. Determine if a vector is within a given span.

In this section we will examine the concept of spanning introduced earlier in terms of $\mathbb{R}^n$. Here, we will discuss these concepts in terms of abstract vector spaces.

Consider the following definition.

Definition 9.2.1: Subset

Let X and Y be two sets. If all elements of X are also elements of Y then we say that X is a **subset** of Y and we write

$$X \subseteq Y$$

In particular, we often speak of subsets of a vector space, such as $X \subseteq V$. By this we mean that every element in the set X is contained in the vector space V.

Definition 9.2.2: Linear Combination

Let V be a vector space and let $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n \subseteq V$. A vector $\vec{v} \in V$ is called a linear combination of the $\vec{v}_i$ if there exist scalars $c_i \in \mathbb{R}$ such that

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$$

This definition leads to our next concept of span.

Definition 9.2.3: Span of Vectors

Let $\{\vec{v}_1, \cdots, \vec{v}_n\} \subseteq V$. Then

$$\text{span}\{\vec{v}_1, \cdots, \vec{v}_n\} = \left\{ \sum_{i=1}^{n} c_i\vec{v}_i : c_i \in \mathbb{R} \right\}$$

When we say that a vector $\vec{w}$ is in $\text{span}\{\vec{v}_1, \cdots, \vec{v}_n\}$ we mean that $\vec{w}$ can be written as a linear combination of the $\vec{v}_1$. We say that a collection of vectors $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a **spanning set** for V if $V = \text{span}\{\vec{v}_1, \cdots, \vec{v}_n\}$.

Consider the following example.

Example 9.2.1: Matrix Span

Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Determine if A and B are in

$$\text{span}\{M_1, M_2\} = \text{span}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

Solution

First consider A. We want to see if scalars s, t can be found such that $A = sM_1 + tM_2$.

$$\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = s \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + t \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

The solution to this equation is given by

$$1 = s$$
$$2 = t$$

and it follows that A is in span $\{M_1, M_2\}$.

Now consider B. Again we write $B = sM_1 + tM_2$ and see if a solution can be found for s, t.

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = s \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + t \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Clearly no values of s and t can be found such that this equation holds. Therefore B is not in span $\{M_1, M_2\}$.

Below is a video on linear combinations and the span of vectors.



Consider another example.

Example 9.2.2: Polynomial Span

Show that $p(x) = 7x^2 + 4x - 3$ is in span $\left\{4x^2 + x, x^2 - 2x + 3\right\}$.

Solution

To show that $p(x)$ is in the given span, we need to show that it can be written as a linear combination of polynomials in the span. Suppose scalars a, b existed such that

$$7x^2 + 4x - 3 = a(4x^2 + x) + b(x^2 - 2x + 3)$$

If this linear combination were to hold, the following would be true:

$$\begin{aligned} 4a + b &= 7 \\ a - 2b &= 4 \\ 3b &= -3 \end{aligned}$$

You can verify that $a = 2, b = -1$ satisfies this system of equations. This means that we can write $p(x)$ as follows:

$$7x^2 + 4x - 3 = 2(4x^2 + x) - (x^2 - 2x + 3)$$

Hence $p(x)$ is in the given span.

Consider the following example.

Example 9.2.3: Spanning Set

Let $S = \left\{x^2 + 1, x - 2, 2x^2 - x\right\}$. Show that S is a spanning set for $\mathbb{P}_2$, the set of all polynomials of degree at most 2.

Solution

Let $p(x) = ax^2 + bx + c$ be an arbitrary polynomial in $\mathbb{P}_2$. To show that S is a spanning set, it suffices to show that $p(x)$ can be written as a linear combination of the elements of S. In other words, can we find r, s, t such that:

$$p(x) = ax^2 + bx + c = r(x^2 + 1) + s(x - 2) + t(2x^2 - x)$$

If a solution r, s, t can be found, then this shows that for any such polynomial $p(x)$, it can be written as a linear combination of the above polynomials and S is a spanning set.

$$\begin{aligned} ax^2 + bx + c &= r(x^2 + 1) + s(x - 2) + t(2x^2 - x) \\ &= rx^2 + r + sx - 2s + 2tx^2 - tx \\ &= (r + 2t)x^2 + (s - t)x + (r - 2s) \end{aligned}$$

For this to be true, the following must hold:

$$\begin{aligned} a &= r + 2t \\ b &= s - t \\ c &= r - 2s \end{aligned}$$

To check that a solution exists, set up the augmented matrix and row reduce:

$$\left[\begin{array}{ccc|c} 1 & 0 & 2 & a \\ 0 & 1 & -1 & b \\ 1 & -2 & 0 & c \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2}a + 2b + \frac{1}{2}c \\ 0 & 1 & 0 & \frac{1}{4}a - \frac{1}{4}c \\ 0 & 0 & 1 & \frac{1}{4}a - b - \frac{1}{4}c \end{array}\right]$$

Clearly a solution exists for any choice of a, b, c. Hence S is a spanning set for $\mathbb{P}_2$.

Below is a video on determining if a vector is a linear combination of a pair of vectors (No).



Below is a video on determining if a vector is a linear combination of a pair of vectors (Yes).



Below is a video on finding a linear combination of a pair of vectors in 3D.



Below is a video on showing that a vector is in the span of two other vectors.



Below is a video on determining which vectors are in the span of a pair of vectors in R^3.



Below is a video on determining if a vector is in the span of a pair of vectors in R^3.



Below is a video on finding a linear combination and additive inverse of two vectors (polynomials).



This page titled 9.2: Spanning Sets is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **9.2: Spanning Sets** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

9.3: Linear Independence

Outcomes

A. Determine if a set is linearly independent.

In this section, we will again explore concepts introduced earlier in terms of $\mathbb{R}^n$ and extend them to apply to abstract vector spaces.

Definition 9.3.1: Linear Independence

Let V be a vector space. If $\{\vec{v}_1, \cdots, \vec{v}_n\} \subseteq V$, then it is linearly independent if

$$\sum_{i=1}^{n} a_i \vec{v}_i = \vec{0} \text{ implies } a_1 = \cdots = a_n = 0$$

where the a_i are real numbers.

The set of vectors is called linearly dependent if it is not linearly independent.

Below is a video on linear independent and dependent sets of vectors.



Example 9.3.1: Linear Independence

Let $S \subseteq \mathbb{P}_2$ be a set of polynomials given by

$$S = \left\{x^2 + 2x - 1, 2x^2 - x + 3\right\}$$

Determine if S is linearly independent.

Solution

To determine if this set S is linearly independent, we write

$$a(x^2 + 2x - 1) + b(2x^2 - x + 3) = 0x^2 + 0x + 0$$

If it is linearly independent, then $a = b = 0$ will be the only solution. We proceed as follows.

$$\begin{aligned} a(x^2 + 2x - 1) + b(2x^2 - x + 3) &= 0x^2 + 0x + 0 \\ ax^2 + 2ax - a + 2bx^2 - bx + 3b &= 0x^2 + 0x + 0 \\ (a + 2b)x^2 + (2a - b)x - a + 3b &= 0x^2 + 0x + 0 \end{aligned}$$

It follows that

$$\begin{aligned} a+2b &= 0 \\ 2a-b &= 0 \\ -a+3b &= 0 \end{aligned}$$

The augmented matrix and resulting reduced row-echelon form are given by

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & -1 & 0 \\ -1 & 3 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Hence the solution is $a = b = 0$ and the set is linearly independent.

The next example shows us what it means for a set to be dependent.

Example 9.3.2: Dependent Set

Determine if the set S given below is independent.

$$S = \left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \right\}$$

Solution

To determine if S is linearly independent, we look for solutions to

$$a \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Notice that this equation has nontrivial solutions, for example $a = 2$, $b = 3$ and $c = -1$. Therefore S is dependent.

The following is an important result regarding dependent sets.

Lemma 9.3.1: Dependent Sets

Let V be a vector space and suppose $W = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_k\}$ is a subset of V. Then W is dependent if and only if $\vec{v}_i$ can be written as a linear combination of $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_{i-1}, \vec{v}_{i+1}, \cdots, \vec{v}_k\}$ for some $i \leq k$.

Revisit Example 9.3.2 with this in mind. Notice that we can write one of the three vectors as a combination of the others.

$$\begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = 2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

By Lemma 9.3.1 this set is dependent.

Below is a video on determining if matrix columns are dependent or independent.



If we know that one particular set is linearly independent, we can use this information to determine if a related set is linearly independent. Consider the following example.

Example 9.3.3: Related Independent Sets

Let V be a vector space and suppose $S \subseteq V$ is a set of linearly independent vectors given by $S = \{\vec{u}, \vec{v}, \vec{w}\}$. Let $R \subseteq V$ be given by $R = \left\{2\vec{u} - \vec{w}, \vec{w} + \vec{v}, 3\vec{v} + \frac{1}{2}\vec{u}\right\}$. Show that R is also linearly independent.

Solution

To determine if R is linearly independent, we write

$$a(2\vec{u} - \vec{w}) + b(\vec{w} + \vec{v}) + c(3\vec{v} + \frac{1}{2}\vec{u}) = \vec{0}$$

If the set is linearly independent, the only solution will be $a = b = c = 0$. We proceed as follows.

$$\begin{aligned} a(2\vec{u} - \vec{w}) + b(\vec{w} + \vec{v}) + c(3\vec{v} + \frac{1}{2}\vec{u}) &= \vec{0} \\ 2a\vec{u} - a\vec{w} + b\vec{w} + b\vec{v} + 3c\vec{v} + \frac{1}{2}c\vec{u} &= \vec{0} \\ (2a + \frac{1}{2}c)\vec{u} + (b + 3c)\vec{v} + (-a + b)\vec{w} &= \vec{0} \end{aligned}$$

We know that the set $S = \{\vec{u}, \vec{v}, \vec{w}\}$ is linearly independent, which implies that the coefficients in the last line of this equation must all equal 0. In other words:

$$\begin{aligned} 2a + \frac{1}{2}c &= 0 \\ b + 3c &= 0 \\ -a + b &= 0 \end{aligned}$$

The augmented matrix and resulting reduced row-echelon form are given by:

$$\left[\begin{array}{ccc|c} 2 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 3 & 0 \\ -1 & 1 & 0 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

Hence the solution is $a = b = c = 0$ and the set is linearly independent.

The following theorem was discussed in terms in $\mathbb{R}^n$. We consider it here in the general case.

Theorem 9.3.1: Unique Representation

Let V be a vector space and let $U = \{\vec{v}_1, \cdots, \vec{v}_k\} \subseteq V$ be an independent set. If $\vec{v} \in \text{span } U$, then $\vec{v}$ can be written uniquely as a linear combination of the vectors in U.

Consider the span of a linearly independent set of vectors. Suppose we take a vector which is not in this span and add it to the set. The following lemma claims that the resulting set is still linearly independent.

Lemma 9.3.2: Adding to a Linearly Independent Set

Suppose $\vec{v} \notin \text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$ and $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is linearly independent. Then the set

$$\{\vec{u}_1, \cdots, \vec{u}_k, \vec{v}\}$$

is also linearly independent.

Proof

Suppose $\sum_{i=1}^{k} c_i \vec{u}_i + d\vec{v} = \vec{0}$. It is required to verify that each $c_i = 0$ and that $d = 0$. But if $d \neq 0$, then you can solve for $\vec{v}$ as a linear combination of the vectors, $\{\vec{u}_1, \cdots, \vec{u}_k\}$,

$$\vec{v} = -\sum_{i=1}^{k} \left(\frac{c_i}{d}\right) \vec{u}_i$$

contrary to the assumption that $\vec{v}$ is not in the span of the $\vec{u}_i$. Therefore, $d = 0$. But then $\sum_{i=1}^{k} c_i \vec{u}_i = \vec{0}$ and the linear independence of $\{\vec{u}_1, \cdots, \vec{u}_k\}$ implies each $c_i = 0$ also.

Below is a video on finding a third vector in R^3 that makes a set dependent (and then independent).



Consider the following example.

Example 9.3.4: Adding to a Linearly Independent Set

Let $S \subseteq M_{22}$ be a linearly independent set given by

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$$

Show that the set $R \subseteq M_{22}$ given by

$$R = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$

is also linearly independent.

Solution

Instead of writing a linear combination of the matrices which equals 0 and showing that the coefficients must equal 0, we can instead use Lemma 9.3.2.

To do so, we show that

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \notin \text{span}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$$

Write

$$\begin{aligned} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} &= a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \end{aligned}$$

Clearly there are no possible a, b to make this equation true. Hence the new matrix does not lie in the span of the matrices in S. By Lemma 9.3.2, R is also linearly independent.

This page titled 9.3: Linear Independence is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **9.3: Linear Independence** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

9.4: Subspaces and Basis

Outcomes

A. Utilize the subspace test to determine if a set is a subspace of a given vector space.
B. Extend a linearly independent set and shrink a spanning set to a basis of a given vector space.

In this section we will examine the concept of subspaces introduced earlier in terms of $\mathbb{R}^n$. Here, we will discuss these concepts in terms of abstract vector spaces.

Consider the definition of a subspace.

Definition 9.4.1: Subspace

Let V be a vector space. A subset $W \subseteq V$ is said to be a **subspace** of V if $a\vec{x} + b\vec{y} \in W$ whenever $a, b \in \mathbb{R}$ and $\vec{x}, \vec{y} \in W$.

Below is a video on proving that a null space is a subspace.



Below is a video on determining if the set of lower triangular matrices is a subspace.



Below is a video on determining if the set of degree 3 polynomials with $p(0) = 1$ is a subspace.



The span of a set of vectors as described in Definition 9.2.3 is an example of a subspace. The following fundamental result says that subspaces are subsets of a vector space which are themselves vector spaces.

Theorem 9.4.1: Subspaces are Vector Spaces

Let W be a nonempty collection of vectors in a vector space V. Then W is a subspace if and only if W satisfies the vector space axioms, using the same operations as those defined on V.

Proof

Suppose first that W is a subspace. It is obvious that all the algebraic laws hold on W because it is a subset of V and they hold on V. Thus $\vec{u}+\vec{v}=\vec{v}+\vec{u}$ along with the other axioms. Does W contain $\vec{0}$? Yes because it contains $0\vec{u}=\vec{0}$. See Theorem 9.1.1.

Are the operations of V defined on W? That is, when you add vectors of W do you get a vector in W? When you multiply a vector in W by a scalar, do you get a vector in W? Yes. This is contained in the definition. Does every vector in W have an additive inverse? Yes by Theorem 9.1.1 because $-\vec{v}=(-1)\vec{v}$ which is given to be in W provided $\vec{v}\in W$.

Next suppose W is a vector space. Then by definition, it is closed with respect to linear combinations. Hence it is a subspace.

Below is a video on determining if the subset of vectors <a,b,0> is a subspace in R^3.



Consider the following useful Corollary.

Corollary 9.4.1: Span is a Subspace

Let V be a vector space with $W \subseteq V$. If $W = \text{span}\{\vec{v}_1, \cdots, \vec{v}_n\}$ then W is a subspace of V.

When determining spanning sets the following theorem proves useful.

Theorem 9.4.2: Spanning Set

Let $W \subseteq V$ for a vector space V and suppose $W = \text{span}\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n\}$.

Let $U \subseteq V$ be a subspace such that $\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_n \in U$. Then it follows that $W \subseteq U$.

In other words, this theorem claims that any subspace that contains a set of vectors must also contain the span of these vectors.

The following example will show that two spans, described differently, can in fact be equal.

Example 9.4.1: Equal Span

Let $p(x), q(x)$ be polynomials and suppose $U = \text{span}\{2p(x) - q(x), p(x) + 3q(x)\}$ and $W = \text{span}\{p(x), q(x)\}$. Show that $U = W$.

Solution

We will use Theorem 9.4.2 to show that $U \subseteq W$ and $W \subseteq U$. It will then follow that $U = W$.

1. $U \subseteq W$ Notice that $2p(x) - q(x)$ and $p(x) + 3q(x)$ are both in $W = span\{p(x), q(x)\}$. Then by Theorem 9.4.2 W must contain the span of these polynomials and so $U \subseteq W$.
2. $W \subseteq U$ Notice that

$$\begin{aligned} p(x) &= \frac{3}{7}(2p(x) - q(x)) + \frac{2}{7}(p(x) + 3q(x)) \\ q(x) &= -\frac{1}{7}(2p(x) - q(x)) + \frac{2}{7}(p(x) + 3q(x)) \end{aligned}$$

Hence $p(x),\ q(x)$ are in $span\{2p(x) - q(x),\ p(x) + 3q(x)\}$. By Theorem 9.4.2 U must contain the span of these polynomials and so $W \subseteq U$.

To prove that a set is a vector space, one must verify each of the axioms given in Definition 9.1.2 and 9.1.3. This is a cumbersome task, and therefore a shorter procedure is used to verify a subspace.

Procedure 9.4.1: Subspace Test

Suppose W is a subset of a vector space V. To determine if W is a subspace of V, it is sufficient to determine if the following three conditions hold, using the operations of V:

1. The additive identity $\vec{0}$ of V is contained in W.
2. For any vectors $\vec{w}_1, \vec{w}_2$ in W, $\vec{w}_1 + \vec{w}_2$ is also in W.
3. For any vector $\vec{w}_1$ in W and scalar a, the product $a\vec{w}_1$ is also in W.

Therefore it suffices to prove these three steps to show that a set is a subspace.

Consider the following example.

Example 9.4.2: Improper Subspaces

Let V be an arbitrary vector space. Then V is a subspace of itself. Similarly, the set $\left\{\vec{0}\right\}$ containing only the zero vector is also a subspace.

Solution

Using the subspace test in Procedure 9.4.1 we can show that V and $\left\{\vec{0}\right\}$ are subspaces of V.

Since V satisfies the vector space axioms it also satisfies the three steps of the subspace test. Therefore V is a subspace.

Let's consider the set $\left\{\vec{0}\right\}$.

1. The vector $\vec{0}$ is clearly contained in $\left\{\vec{0}\right\}$, so the first condition is satisfied.
2. Let $\vec{w}_1, \vec{w}_2$ be in $\left\{\vec{0}\right\}$. Then $\vec{w}_1 = \vec{0}$ and $\vec{w}_2 = \vec{0}$ and so

$$\vec{w}_1 + \vec{w}_2 = \vec{0} + \vec{0} = \vec{0}$$

   It follows that the sum is contained in $\left\{\vec{0}\right\}$ and the second condition is satisfied.
3. Let $\vec{w}_1$ be in $\left\{\vec{0}\right\}$ and let a be an arbitrary scalar. Then

$$a\vec{w}_1 = a\vec{0} = \vec{0}$$

   Hence the product is contained in $\left\{\vec{0}\right\}$ and the third condition is satisfied.

It follows that $\left\{\vec{0}\right\}$ is a subspace of V.

The two subspaces described above are called **improper subspaces**. Any subspace of a vector space V which is not equal to V or $\left\{\vec{0}\right\}$ is called a **proper subspace**.

Below is a video on determining if the subset of vectors <a,b,1> is a subspace in R^3.



Consider another example.

Example 9.4.3: Subspace of Polynomials

Let $\mathbb{P}_2$ be the vector space of polynomials of degree two or less. Let $W \subseteq \mathbb{P}_2$ be all polynomials of degree two or less which have 1 as a root. Show that W is a subspace of $\mathbb{P}_2$.

Solution

First, express W as follows:

$$W = \left\{p(x) = ax^2 + bx + c, a, b, c, \in \mathbb{R} | p(1) = 0\right\}$$

We need to show that W satisfies the three conditions of Procedure 9.4.1.

1. The zero polynomial of $\mathbb{P}_2$ is given by $0(x) = 0x^2 + 0x + 0 = 0$. Clearly $0(1) = 0$ so $0(x)$ is contained in W.
2. Let $p(x), q(x)$ be polynomials in W. It follows that $p(1) = 0$ and $q(1) = 0$. Now consider $p(x) + q(x)$. Let $r(x)$ represent this sum.

$$\begin{aligned} r(1) &= p(1)+q(1) \\ &= 0+0 \\ &= 0 \end{aligned}$$

Therefore the sum is also in W and the second condition is satisfied.

3. Let $p(x)$ be a polynomial in W and let a be a scalar. It follows that $p(1)=0$. Consider the product $ap(x)$.

$$\begin{aligned} ap(1) &= a(0) \\ &= 0 \end{aligned}$$

Therefore the product is in W and the third condition is satisfied.

It follows that W is a subspace of $\mathbb{P}_2$.

Recall the definition of basis, considered now in the context of vector spaces.

Definition 9.4.2: Basis

Let V be a vector space. Then $\{\vec{v}_1,\cdots,\vec{v}_n\}$ is called a basis for V if the following conditions hold.

1. $\text{span}\{\vec{v}_1,\cdots,\vec{v}_n\}=V$
2. $\{\vec{v}_1,\cdots,\vec{v}_n\}$ is linearly independent

Consider the following example.

Example 9.4.4: Polynomials of Degree Two

Let $\mathbb{P}_2$ be the set polynomials of degree no more than 2. We can write $\mathbb{P}_2=\text{span}\left\{x^2,x,1\right\}$. Is $\left\{x^2,x,1\right\}$ a basis for $\mathbb{P}_2$?

Solution

It can be verified that $\mathbb{P}_2$ is a vector space defined under the usual addition and scalar multiplication of polynomials.

Now, since $\mathbb{P}_2=\text{span}\left\{x^2,x,1\right\}$, the set $\left\{x^2,x,1\right\}$ is a basis if it is linearly independent. Suppose then that

$$ax^2+bx+c=0x^2+0x+0$$

where a,b,c are real numbers. It is clear that this can only occur if $a=b=c=0$. Hence the set is linearly independent and forms a basis of $\mathbb{P}_2$.

The next theorem is an essential result in linear algebra and is called the exchange theorem.

Theorem 9.4.3: Exchange Theorem

Let $\{\vec{x}_1,\cdots,\vec{x}_r\}$ be a linearly independent set of vectors such that each $\vec{x}_i$ is contained in $\text{span}\{\vec{y}_1,\cdots,\vec{y}_s\}$. Then $r\leq s$.

Proof

The proof will proceed as follows. First, we set up the necessary steps for the proof. Next, we will assume that $r>s$ and show that this leads to a contradiction, thus requiring that $r\leq s$.

Define $\text{span}\{\vec{y}_1,\cdots,\vec{y}_s\}=V$. Since each $\vec{x}_i$ is in $\text{span}\{\vec{y}_1,\cdots,\vec{y}_s\}$, it follows there exist scalars $c_1,\cdots,c_s$ such that

$$\vec{x}_1=\sum_{i=1}^{s}c_i\vec{y}_i \tag{9.4.1}$$

Note that not all of these scalars c_i can equal zero. Suppose that all the $c_i=0$. Then it would follow that $\vec{x}_1=\vec{0}$ and so $\{\vec{x}_1,\cdots,\vec{x}_r\}$ would not be linearly independent. Indeed, if $\vec{x}_1=\vec{0}$, $1\vec{x}_1+\sum_{i=2}^{r}0\vec{x}_i=\vec{x}_1=\vec{0}$ and so there would exist a nontrivial linear combination of the vectors $\{\vec{x}_1,\cdots,\vec{x}_r\}$ which equals zero. Therefore at least one c_i is nonzero.

Say $c_k\neq 0$. Then solve (9.4.1) for $\vec{y}_k$ and obtain

$$\vec{y}_k \in \text{span}\left\{\vec{x}_1, \overbrace{\vec{y}_1, \cdots, \vec{y}_{k-1}, \vec{y}_{k+1}, \cdots, \vec{y}_s}^{\text{s-1 vectors here}}\right\}.$$

Define $\{\vec{z}_1, \cdots, \vec{z}_{s-1}\}$ to be

$$\{\vec{z}_1, \cdots, \vec{z}_{s-1}\} = \{\vec{y}_1, \cdots, \vec{y}_{k-1}, \vec{y}_{k+1}, \cdots, \vec{y}_s\}$$

Now we can write

$$\vec{y}_k \in \text{span}\{\vec{x}_1, \vec{z}_1, \cdots, \vec{z}_{s-1}\}$$

Therefore, $\text{span}\{\vec{x}_1, \vec{z}_1, \cdots, \vec{z}_{s-1}\} = V$. To see this, suppose $\vec{v} \in V$. Then there exist constants $c_1, \cdots, c_s$ such that

$$\vec{v} = \sum_{i=1}^{s-1} c_i \vec{z}_i + c_s \vec{y}_k.$$

Replace this $\vec{y}_k$ with a linear combination of the vectors $\{\vec{x}_1, \vec{z}_1, \cdots, \vec{z}_{s-1}\}$ to obtain $\vec{v} \in \text{span}\{\vec{x}_1, \vec{z}_1, \cdots, \vec{z}_{s-1}\}$. The vector $\vec{y}_k$, in the list $\{\vec{y}_1, \cdots, \vec{y}_s\}$, has now been replaced with the vector $\vec{x}_1$ and the resulting modified list of vectors has the same span as the original list of vectors, $\{\vec{y}_1, \cdots, \vec{y}_s\}$.

We are now ready to move on to the proof. Suppose that $r > s$ and that

$$\text{span}\{\vec{x}_1, \cdots, \vec{x}_l, \vec{z}_1, \cdots, \vec{z}_p\} = V$$

where the process established above has continued. In other words, the vectors $\vec{z}_1, \cdots, \vec{z}_p$ are each taken from the set $\{\vec{y}_1, \cdots, \vec{y}_s\}$ and $l + p = s$. This was done for $l = 1$ above. Then since $r > s$, it follows that $l \leq s < r$ and so $l + 1 \leq r$. Therefore, $\vec{x}_{l+1}$ is a vector not in the list, $\{\vec{x}_1, \cdots, \vec{x}_l\}$ and since

$$\text{span}\{\vec{x}_1, \cdots, \vec{x}_l, \vec{z}_1, \cdots, \vec{z}_p\} = V$$

there exist scalars, c_i and d_j such that

$$\vec{x}_{l+1} = \sum_{i=1}^{l} c_i \vec{x}_i + \sum_{j=1}^{p} d_j \vec{z}_j. \tag{9.4.2}$$

Not all the d_j can equal zero because if this were so, it would follow that $\{\vec{x}_1, \cdots, \vec{x}_r\}$ would be a linearly dependent set because one of the vectors would equal a linear combination of the others. Therefore, (9.4.2) can be solved for one of the $\vec{z}_i$, say $\vec{z}_k$, in terms of $\vec{x}_{l+1}$ and the other $\vec{z}_i$ and just as in the above argument, replace that $\vec{z}_i$ with $\vec{x}_{l+1}$ to obtain

$$\text{span}\left\{\vec{x}_1, \cdots \vec{x}_l, \vec{x}_{l+1}, \overbrace{\vec{z}_1, \cdots \vec{z}_{k-1}, \vec{z}_{k+1}, \cdots, \vec{z}_p}^{\text{p-1 vectors here}}\right\} = V$$

Continue this way, eventually obtaining

$$\text{span}\{\vec{x}_1, \cdots, \vec{x}_s\} = V.$$

But then $\vec{x}_r \in \text{span}\{\vec{x}_1, \cdots, \vec{x}_s\}$ contrary to the assumption that $\{\vec{x}_1, \cdots, \vec{x}_r\}$ is linearly independent. Therefore, $r \leq s$ as claimed.

The following corollary follows from the exchange theorem.

Corollary 9.4.2: Two Bases of the Same Length

Let B_1, B_2 be two bases of a vector space V. Suppose B_1 contains m vectors and B_2 contains n vectors. Then $m = n$.

Proof

By Theorem 9.4.3, $m \leq n$ and $n \leq m$. Therefore $m = n$.

This corollary is very important so we provide another proof independent of the exchange theorem above.

Proof

Suppose $n > m$. Then since the vectors $\{\vec{u}_1, \cdots, \vec{u}_m\}$ span V, there exist scalars c_{ij} such that

$$\sum_{i=1}^{m} c_{ij}\vec{u}_i = \vec{v}_j.$$

Therefore,

$$\sum_{j=1}^{n} d_j\vec{v}_j = \vec{0} \text{ if and only if } \sum_{j=1}^{n}\sum_{i=1}^{m} c_{ij}d_j\vec{u}_i = \vec{0}$$

if and only if

$$\sum_{i=1}^{m}\left(\sum_{j=1}^{n} c_{ij}d_j\right)\vec{u}_i = \vec{0}$$

Now since $\{\vec{u}_1, \cdots, \vec{u}_n\}$ is independent, this happens if and only if

$$\sum_{j=1}^{n} c_{ij}d_j = 0,\ i = 1, 2, \cdots, m.$$

However, this is a system of m equations in n variables, $d_1, \cdots, d_n$ and $m < n$. Therefore, there exists a solution to this system of equations in which not all the d_j are equal to zero. Recall why this is so. The augmented matrix for the system is of the form $\left[\begin{array}{c|c} C & \vec{0} \end{array}\right]$ where C is a matrix which has more columns than rows. Therefore, there are free variables and hence nonzero solutions to the system of equations. However, this contradicts the linear independence of $\{\vec{u}_1, \cdots, \vec{u}_m\}$. Similarly it cannot happen that $m > n$.

Given the result of the previous corollary, the following definition follows.

Definition 9.4.3: Dimension

A vector space V is of dimension n if it has a basis consisting of n vectors.

Notice that the dimension is well defined by Corollary 9.4.2. It is assumed here that $n < \infty$ and therefore such a vector space is said to be **finite dimensional**.

Example 9.4.5: Dimension of a Vector Space

Let $\mathbb{P}_2$ be the set of all polynomials of degree at most 2. Find the dimension of $\mathbb{P}_2$.

Solution

If we can find a basis of $\mathbb{P}_2$ then the number of vectors in the basis will give the dimension. Recall from Example 9.4.4 that a basis of $\mathbb{P}_2$ is given by

$$S = \left\{x^2, x, 1\right\}$$

There are three polynomials in S and hence the dimension of $\mathbb{P}_2$ is three.

It is important to note that a basis for a vector space is not unique. A vector space can have many bases. Consider the following example.

Example 9.4.6: A Different Basis for Polynomials of Degree Two

Let $\mathbb{P}_2$ be the polynomials of degree no more than 2. Is $\left\{x^2+x+1, 2x+1, 3x^2+1\right\}$ a basis for $\mathbb{P}_2$?

Solution

Suppose these vectors are linearly independent but do not form a spanning set for $\mathbb{P}_2$. Then by Lemma 9.3.2, we could find a fourth polynomial in $\mathbb{P}_2$ to create a new linearly independent set containing four polynomials. However this would imply that we could find a basis of $\mathbb{P}_2$ of more than three polynomials. This contradicts the result of Example 9.4.5 in which we determined the dimension of $\mathbb{P}_2$ is three. Therefore if these vectors are linearly independent they must also form a spanning set and thus a basis for $\mathbb{P}_2$.

Suppose then that

$$\begin{aligned} a\left(x^2+x+1\right)+b\left(2x+1\right)+c\left(3x^2+1\right) &= 0 \\ \left(a+3c\right)x^2+\left(a+2b\right)x+\left(a+b+c\right) &= 0 \end{aligned}$$

We know that $\left\{x^2, x, 1\right\}$ is linearly independent, and so it follows that

$$\begin{aligned} a+3c &= 0 \\ a+2b &= 0 \\ a+b+c &= 0 \end{aligned}$$

and there is only one solution to this system of equations, $a=b=c=0$. Therefore, these are linearly independent and form a basis for $\mathbb{P}_2$.

Below is a video on determining which set of polynomials form a basis for P_2.



Consider the following theorem.

Theorem 9.4.4: Every Subspace has a Basis

Let W be a nonzero subspace of a finite dimensional vector space V. Suppose V has dimension n. Then W has a basis with no more than n vectors.

Proof

Let $\vec{v}_1 \in V$ where $\vec{v}_1 \neq 0$. If $\text{span}\left\{\vec{v}_1\right\} = V$, then it follows that $\left\{\vec{v}_1\right\}$ is a basis for V. Otherwise, there exists $\vec{v}_2 \in V$ which is not in $\text{span}\left\{\vec{v}_1\right\}$. By Lemma 9.3.2 $\left\{\vec{v}_1, \vec{v}_2\right\}$ is a linearly independent set of vectors. Then $\left\{\vec{v}_1, \vec{v}_2\right\}$ is a basis for V and we are done. If $\text{span}\left\{\vec{v}_1, \vec{v}_2\right\} \neq V$, then there exists $\vec{v}_3 \notin \text{span}\left\{\vec{v}_1, \vec{v}_2\right\}$ and $\left\{\vec{v}_1, \vec{v}_2, \vec{v}_3\right\}$ is a larger linearly independent set of vectors. Continuing this way, the process must stop before $n+1$ steps because if not, it would be possible to obtain $n+1$ linearly independent vectors contrary to the exchange theorem, Theorem 9.4.3

If in fact W has n vectors, then it follows that $W = V$.

Theorem 9.4.5: Subspace of Same Dimension

Let V be a vector space of dimension n and let W be a subspace. Then $W = V$ if and only if the dimension of W is also n.

Proof

First suppose $W = V$. Then obviously the dimension of $W = n$.

Now suppose that the dimension of W is n. Let a basis for W be $\{\vec{w}_1, \cdots, \vec{w}_n\}$. If W is not equal to V, then let $\vec{v}$ be a vector of V which is not contained in W. Thus $\vec{v}$ is not in $\text{span}\{\vec{w}_1, \cdots, \vec{w}_n\}$ and by Lemma 9.7.2, $\{\vec{w}_1, \cdots, \vec{w}_n, \vec{v}\}$ is linearly independent which contradicts Theorem 9.4.3because it would be an independent set of $n+1$ vectors even though each of these vectors is in a spanning set of n vectors, a basis of V.

Consider the following example.

Example 9.4.7: Basis of a Subspace

Let $U = \left\{ A \in \mathbb{M}_{22} \left| A \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} A \right. \right\}$. Then U is a subspace of $\mathbb{M}_{22}$ Find a basis of U, and hence $\dim(U)$.

Solution

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{M}_{22}$. Then

$$A \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} a+b & -b \\ c+d & -d \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} A = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ -c & -d \end{bmatrix}.$$

If $A \in U$, then $\begin{bmatrix} a+b & -b \\ c+d & -d \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ -c & -d \end{bmatrix}$.

Equating entries leads to a system of four equations in the four variables a, b, c and d.

$$\begin{aligned} a+b &= a+c \\ -b &= b+d \\ c+d &= -c \\ -d &= -d \end{aligned}$$

or

$$\begin{aligned} b-c &= 0 \\ -2b-d &= 0 \\ 2c+d &= 0 \end{aligned}.$$

The solution to this system is $a = s$, $b = -\frac{1}{2}t$, $c = -\frac{1}{2}t$, $d = t$ for any $s, t \in \mathbb{R}$, and thus

$$A = \begin{bmatrix} s & \frac{t}{2} \\ -\frac{t}{2} & t \end{bmatrix} = s \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + t \begin{bmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}.$$

Let

$$B = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \right\}.$$

Then $\text{span}(B) = U$, and it is routine to verify that B is an independent subset of $\mathbb{M}_{22}$. Therefore B is a basis of U, and $\dim(U) = 2$.

Below is a video on determining which set of polynomials form a basis for P_3.



The following theorem claims that a spanning set of a vector space V can be shrunk down to a basis of V. Similarly, a linearly independent set within V can be enlarged to create a basis of V.

Theorem 9.4.6: Basis of V

If $V = \text{span}\{\vec{u}_1, \cdots, \vec{u}_n\}$ is a vector space, then some subset of $\{\vec{u}_1, \cdots, \vec{u}_n\}$ is a basis for V. Also, if $\{\vec{u}_1, \cdots, \vec{u}_k\} \subseteq V$ is linearly independent and the vector space is finite dimensional, then the set $\{\vec{u}_1, \cdots, \vec{u}_k\}$, can be enlarged to obtain a basis of V.

Proof

Let

$$S = \{E \subseteq \{\vec{u}_1, \cdots, \vec{u}_n\} \text{ such that span}\{E\} = V\}.$$

For $E \in S$, let $|E|$ denote the number of elements of E. Let

$$m = \min\{|E| \text{ such that } E \in S\}.$$

Thus there exist vectors

$$\{\vec{v}_1, \cdots, \vec{v}_m\} \subseteq \{\vec{u}_1, \cdots, \vec{u}_n\}$$

such that

$$\text{span}\{\vec{v}_1, \cdots, \vec{v}_m\} = V$$

and m is as small as possible for this to happen. If this set is linearly independent, it follows it is a basis for V and the theorem is proved. On the other hand, if the set is not linearly independent, then there exist scalars, $c_1, \cdots, c_m$ such that

$$\vec{0} = \sum_{i=1}^{m} c_i \vec{v}_i$$

and not all the c_i are equal to zero. Suppose $c_k \neq 0$. Then solve for the vector $\vec{v}_k$ in terms of the other vectors. Consequently,

$$V = \text{span}\{\vec{v}_1, \cdots, \vec{v}_{k-1}, \vec{v}_{k+1}, \cdots, \vec{v}_m\}$$

contradicting the definition of m. This proves the first part of the theorem.

To obtain the second part, begin with $\{\vec{u}_1, \cdots, \vec{u}_k\}$ and suppose a basis for V is

$$\{\vec{v}_1, \cdots, \vec{v}_n\}$$

If

$$\text{span}\{\vec{u}_1, \cdots, \vec{u}_k\} = V,$$

then $k = n$. If not, there exists a vector

$$\vec{u}_{k+1} \notin \text{span}\{\vec{u}_1, \cdots, \vec{u}_k\}$$

Then from Lemma 9.3.2, $\{\vec{u}_1, \cdots, \vec{u}_k, \vec{u}_{k+1}\}$ is also linearly independent. Continue adding vectors in this way until n linearly independent vectors have been obtained. Then

$$\text{span}\{\vec{u}_1, \cdots, \vec{u}_n\} = V$$

because if it did not do so, there would exist $\vec{u}_{n+1}$ as just described and $\{\vec{u}_1, \cdots, \vec{u}_{n+1}\}$ would be a linearly independent set of vectors having $n+1$ elements. This contradicts the fact that $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis. In turn this would contradict Theorem 9.4.3. Therefore, this list is a basis.

Recall Example 9.3.4 in which we added a matrix to a linearly independent set to create a larger linearly independent set. By Theorem 9.4.6 we can extend a linearly independent set to a basis.

Example 9.4.8: Adding to a Linearly Independent Set

Let $S \subseteq M_{22}$ be a linearly independent set given by

$$S = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$$

Enlarge S to a basis of M_{22}.

Solution

Recall from the solution of Example 9.3.4 that the set $R \subseteq M_{22}$ given by

$$R = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$$

is also linearly independent. However this set is still not a basis for M_{22} as it is not a spanning set. In particular, $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ is not in span$R$. Therefore, this matrix can be added to the set by Lemma 9.3.2 to obtain a new linearly independent set given by

$$T = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

This set is linearly independent and now spans M_{22}. Hence T is a basis.

Next we consider the case where you have a spanning set and you want a subset which is a basis. The above discussion involved adding vectors to a set. The next theorem involves removing vectors.

Theorem 9.4.7: Basis from a Spanning Set

Let V be a vector space and let W be a subspace. Also suppose that $W = \text{span}\{\vec{w}_1, \cdots, \vec{w}_m\}$. Then there exists a subset of $\{\vec{w}_1, \cdots, \vec{w}_m\}$ which is a basis for W.

Proof

Let S denote the set of positive integers such that for $k \in S$, there exists a subset of $\{\vec{w}_1, \cdots, \vec{w}_m\}$ consisting of exactly k vectors which is a spanning set for W. Thus $m \in S$. Pick the smallest positive integer in S. Call it k. Then there exists $\{\vec{u}_1, \cdots, \vec{u}_k\} \subseteq \{\vec{w}_1, \cdots, \vec{w}_m\}$ such that $span\{\vec{u}_1, \cdots, \vec{u}_k\} = W$. If

$$\sum_{i=1}^{k} c_i \vec{w}_i = \vec{0}$$

and not all of the $c_i = 0$, then you could pick $c_j \neq 0$, divide by it and solve for $\vec{u}_j$ in terms of the others.

$$\vec{w}_j = \sum_{i \neq j} \left(-\frac{c_i}{c_j} \right) \vec{w}_i$$

Then you could delete $\vec{w}_j$ from the list and have the same span. In any linear combination involving $\vec{w}_j$, the linear combination would equal one in which $\vec{w}_j$ is replaced with the above sum, showing that it could have been obtained as a linear combination of $\vec{w}_i$ for $i \neq j$. Thus $k-1 \in S$ contrary to the choice of k. Hence each $c_i = 0$ and so $\{\vec{u}_1, \cdots, \vec{u}_k\}$ is a basis for W consisting of vectors of $\{\vec{w}_1, \cdots, \vec{w}_m\}$.

Consider the following example of this concept.

Example 9.4.9: Basis from a Spanning Set

Let V be the vector space of polynomials of degree no more than 3, denoted earlier as $\mathbb{P}_3$. Consider the following vectors in V.

$$2x^2+x+1, x^3+4x^2+2x+2, 2x^3+2x^2+2x+1,$$
$$x^3+4x^2-3x+2, x^3+3x^2+2x+1$$

Then, as mentioned above, V has dimension 4 and so clearly these vectors are not linearly independent. A basis for V is $\left\{1, x, x^2, x^3\right\}$. Determine a linearly independent subset of these which has the same span. Determine whether this subset is a basis for V.

Solution

Consider an isomorphism which maps $\mathbb{R}^4$ to V in the obvious way. Thus

$$\begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}$$

corresponds to $2x^2+x+1$ through the use of this isomorphism. Then corresponding to the above vectors in V we would have the following vectors in $\mathbb{R}^4$.

$$\begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}$$

Now if we obtain a subset of these which has the same span but which is linearly independent, then the corresponding vectors from V will also be linearly independent. If there are four in the list, then the resulting vectors from V must be a basis for V. The reduced row-echelon form for the matrix which has the above vectors as columns is

$$\begin{bmatrix} 1 & 0 & 0 & -15 & 0 \\ 0 & 1 & 0 & 11 & 0 \\ 0 & 0 & 1 & -5 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Therefore, a basis for V consists of the vectors

$$2x^2+x+1, x^3+4x^2+2x+2, 2x^3+2x^2+2x+1,$$
$$x^3+3x^2+2x+1.$$

Note how this is a subset of the original set of vectors. If there had been only three pivot columns in this matrix, then we would not have had a basis for V but we would at least have obtained a linearly independent subset of the original set of vectors in this way.

Note also that, since all linear relations are preserved by an isomorphism,

$$\begin{aligned} &-15\left(2x^2+x+1\right)+11\left(x^3+4x^2+2x+2\right)+(-5)\left(2x^3+2x^2+2x+1\right) \\ &=x^3+4x^2-3x+2 \end{aligned}$$

Consider the following example.

Example 9.4.10: Shrinking a Spanning Set

Consider the set $S \subseteq \mathbb{P}_2$ given by

$$S=\left\{1, x, x^2, x^2+1\right\}$$

Show that S spans $\mathbb{P}_2$, then remove vectors from S until it creates a basis.

Solution

First we need to show that S spans $\mathbb{P}_2$. Let ax^2+bx+c be an arbitrary polynomial in $\mathbb{P}_2$. Write

$$ax^2+bx+c=r(1)+s(x)+t(x^2)+u(x^2+1)$$

Then,

$$\begin{aligned} ax^2+bx+c &=r(1)+s(x)+t(x^2)+u(x^2+1) \\ &=(t+u)x^2+s(x)+(r+u) \end{aligned}$$

It follows that

$$\begin{aligned} a &=t+u \\ b &=s \\ c &=r+u \end{aligned}$$

Clearly a solution exists for all a, b, c and so S is a spanning set for $\mathbb{P}_2$. By Theorem 9.4.6, some subset of S is a basis for $\mathbb{P}_2$.

Recall that a basis must be both a spanning set and a linearly independent set. Therefore we must remove a vector from S keeping this in mind. Suppose we remove x from S. The resulting set would be $\left\{1, x^2, x^2+1\right\}$. This set is clearly linearly dependent (and also does not span $\mathbb{P}_2$) and so is not a basis.

Suppose we remove x^2+1 from S. The resulting set is $\left\{1, x, x^2\right\}$ which is both linearly independent and spans $\mathbb{P}_2$. Hence this is a basis for $\mathbb{P}_2$. Note that removing any one of $1, x^2$, or x^2+1 will result in a basis.

Now the following is a fundamental result about subspaces.

Theorem 9.4.8: Basis of a Vector Space

Let V be a finite dimensional vector space and let W be a non-zero subspace. Then W has a basis. That is, there exists a linearly independent set of vectors $\{\vec{w}_1, \cdots, \vec{w}_r\}$ such that

$$\{\vec{w}_1, \cdots, \vec{w}_r\}=W$$

Also if $\{\vec{w}_1, \cdots, \vec{w}_s\}$ is a linearly independent set of vectors, then W has a basis of the form $\{\vec{w}_1, \cdots, \vec{w}_s, \cdots, \vec{w}_r\}$ for $r \geq s$.

Proof

Let the dimension of V be n. Pick $\vec{w}_1 \in W$ where $\vec{w}_1 \neq \vec{0}$. If $\vec{w}_1, \cdots, \vec{w}_s$ have been chosen such that $\{\vec{w}_1, \cdots, \vec{w}_s\}$ is linearly independent, if $span\{\vec{w}_1, \cdots, \vec{w}_r\}=W$, stop. You have the desired basis. Otherwise, there exists $\vec{w}_{s+1} \notin span\{\vec{w}_1, \cdots, \vec{w}_s\}$ and $\{\vec{w}_1, \cdots, \vec{w}_s, \vec{w}_{s+1}\}$ is linearly independent. Continue this way until the process stops. It must stop since otherwise, you could obtain a linearly independent set of vectors having more than n vectors which is impossible.

The last claim is proved by following the above procedure starting with $\{\vec{w}_1, \cdots, \vec{w}_s\}$ as above.

This also proves the following corollary. Let V play the role of W in the above theorem and begin with a basis for W, enlarging it to form a basis for V as discussed above.

Corollary 9.4.3: Basis Extension

Let W be any non-zero subspace of a vector space V. Then every basis of W can be extended to a basis for V.

Consider the following example.

Example 9.4.11: Basis Extension

Let $V = \mathbb{R}^4$ and let

$$W = \text{span}\left\{\begin{bmatrix}1\\0\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\\1\end{bmatrix}\right\}$$

Extend this basis of W to a basis of V.

Solution

An easy way to do this is to take the reduced row-echelon form of the matrix

$$\begin{bmatrix}1&0&1&0&0&0\\0&1&0&1&0&0\\1&0&0&0&1&0\\1&1&0&0&0&1\end{bmatrix} \tag{9.4.3}$$

Note how the given vectors were placed as the first two and then the matrix was extended in such a way that it is clear that the span of the columns of this matrix yield all of $\mathbb{R}^4$. Now determine the pivot columns. The reduced row-echelon form is

$$\begin{bmatrix}1&0&0&0&1&0\\0&1&0&0&-1&1\\0&0&1&0&-1&0\\0&0&0&1&1&-1\end{bmatrix} \tag{9.4.4}$$

These are

$$\begin{bmatrix}1\\0\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}$$

and now this is an extension of the given basis for W to a basis for $\mathbb{R}^4$.

Why does this work? The columns of (9.4.3) obviously span $\mathbb{R}^4$ the span of the first four is the same as the span of all six.

This page titled 9.4: Subspaces and Basis is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **9.4: Subspaces and Basis** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

9.5: Sums and Intersections

Outcomes

A. Show that the sum of two subspaces is a subspace.
B. Show that the intersection of two subspaces is a subspace.

We begin this section with a definition.

Definition 9.5.1: Sum and Intersection

Let V be a vector space, and let U and W be subspaces of V. Then

1. $U+W=\{\vec{u}+\vec{w} \mid \vec{u}\in U \text{ and } \vec{w}\in W\}$ and is called the sum of U and W.
2. $U\cap W=\{\vec{v} \mid \vec{v}\in U \text{ and } \vec{v}\in W\}$ and is called the intersection of U and W.

Therefore the intersection of two subspaces is all the vectors shared by both. If there are no vectors shared by both subspaces, meaning that $U\cap W=\left\{\vec{0}\right\}$, the sum $U+W$ takes on a special name.

Definition 9.5.2: Direct Sum

Let V be a vector space and suppose U and W are subspaces of V such that $U\cap W=\left\{\vec{0}\right\}$. Then the sum of U and W is called the direct sum and is denoted $U\oplus W$.

An interesting result is that both the sum $U+W$ and the intersection $U\cap W$ are subspaces of V.

Example 9.5.1: Intersection is a Subspace

Let V be a vector space and suppose U and W are subspaces. Then the intersection $U\cap W$ is a subspace of V.

Solution

By the subspace test, we must show three things:

1. $\vec{0}\in U\cap W$
2. For vectors $\vec{v}_1,\vec{v}_2\in U\cap W, \vec{v}_1+\vec{v}_2\in U\cap W$
3. For scalar a and vector $\vec{v}\in U\cap W, a\vec{v}\in U\cap W$

We proceed to show each of these three conditions hold.

1. Since U and W are subspaces of V, they each contain $\vec{0}$. By definition of the intersection, $\vec{0}\in U\cap W$.
2. Let $\vec{v}_1,\vec{v}_2\in U\cap W,$. Then in particular, $\vec{v}_1,\vec{v}_2\in U$. Since U is a subspace, it follows that $\vec{v}_1+\vec{v}_2\in U$. The same argument holds for W. Therefore $\vec{v}_1+\vec{v}_2$ is in both U and W and by definition is also in $U\cap W$.
3. Let a be a scalar and $\vec{v}\in U\cap W$. Then in particular, $\vec{v}\in U$. Since U is a subspace, it follows that $a\vec{v}\in U$. The same argument holds for W so $a\vec{v}$ is in both U and W. By definition, it is in $U\cap W$.

Therefore $U\cap W$ is a subspace of V.

It can also be shown that $U+W$ is a subspace of V.

We conclude this section with an important theorem on dimension.

Theorem 9.5.1: Dimension of Sum

Let V be a vector space with subspaces U and W. Suppose U and W each have finite dimension. Then $U+W$ also has finite dimension which is given by

$$\dim(U+W)=\dim(U)+\dim(W)-\dim(U\cap W)$$

Notice that when $U \cap W = \left\{\vec{0}\right\}$, the sum becomes the direct sum and the above equation becomes

$$\dim(U \oplus W) = \dim(U) + \dim(W)$$

Below is a video on intersections and sums of subspaces.



This page titled 9.5: Sums and Intersections is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **9.5: Sums and Intersections** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

9.6: Linear Transformations

Outcomes

A. Understand the definition of a linear transformation in the context of vector spaces.

Recall that a function is simply a transformation of a vector to result in a new vector. Consider the following definition.

Definition 9.6.1: Linear Transformation

Let V and W be vector spaces. Suppose $T : V \mapsto W$ is a function, where for each $\vec{x} \in V, T(\vec{x}) \in W$. Then T is a **linear transformation** if whenever k, p are scalars and $\vec{v}_1$ and $\vec{v}_2$ are vectors in V

$$T(k\vec{v}_1 + p\vec{v}_2) = kT(\vec{v}_1) + pT(\vec{v}_2)$$

Several important examples of linear transformations include the zero transformation, the identity transformation, and the scalar transformation.

Example 9.6.1: Linear Transformations

Let V and W be vector spaces.

1. **The zero transformation**
   $0 : V \to W$ is defined by $0(\vec{v}) = \vec{0}$ for all $\vec{v} \in V$.
2. **The identity transformation**
   $1_V : V \to V$ is defined by $1_V(\vec{v}) = \vec{v}$ for all $\vec{v} \in V$.
3. **The scalar transformation** Let $a \in \mathbb{R}$.
   $s_a : V \to V$ is defined by $s_a(\vec{v}) = a\vec{v}$ for all $\vec{v} \in V$.

Solution

We will show that the scalar transformation s_a is linear, the rest are left as an exercise.

By Definition 9.6.1 we must show that for all scalars k, p and vectors $\vec{v}_1$ and $\vec{v}_2$ in V, $s_a(k\vec{v}_1 + p\vec{v}_2) = ks_a(\vec{v}_1) + ps_a(\vec{v}_2)$. Assume that a is also a scalar.

$$\begin{aligned} s_a(k\vec{v}_1 + p\vec{v}_2) &= a(k\vec{v}_1 + p\vec{v}_2) \\ &= ak\vec{v}_1 + ap\vec{v}_2 \\ &= k(a\vec{v}_1) + p(a\vec{v}_2) \\ &= ks_a(\vec{v}_1) + ps_a(\vec{v}_2) \end{aligned}$$

Therefore s_a is a linear transformation.

Below is a video on finding a linear transformation given $T(a+bt)$ and $T(c+dt)$: P^1 to M_{2x2}.



Consider the following important theorem.

Theorem 9.6.1: Properties of Linear Transformations

Let V and W be vector spaces, and $T : V \mapsto W$ a linear transformation. Then

1. T preserves the zero vector.
$$T(\vec{0}) = \vec{0}$$
2. T preserves additive inverses. For all $\vec{v} \in V$,
$$T(-\vec{v}) = -T(\vec{v})$$
3. T preserves linear combinations. For all $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m \in V$ and all $k_1, k_2, \ldots, k_m \in \mathbb{R}$,
$$T(k_1\vec{v}_1 + k_2\vec{v}_2 + \cdots + k_m\vec{v}_m) = k_1T(\vec{v}_1) + k_2T(\vec{v}_2) + \cdots + k_mT(\vec{v}_m).$$

Proof

1. Let $\vec{0}_V$ denote the zero vector of V and let $\vec{0}_W$ denote the zero vector of W. We want to prove that $T(\vec{0}_V) = \vec{0}_W$. Let $\vec{v} \in V$. Then $0\vec{v} = \vec{0}_V$ and
$$T(\vec{0}_V) = T(0\vec{v}) = 0T(\vec{v}) = \vec{0}_W.$$
2. Let $\vec{v} \in V$; then $-\vec{v} \in V$ is the additive inverse of $\vec{v}$, so $\vec{v} + (-\vec{v}) = \vec{0}_V$. Thus
$$\begin{aligned} T(\vec{v} + (-\vec{v})) &= T(\vec{0}_V) \\ T(\vec{v}) + T(-\vec{v})) &= \vec{0}_W \\ T(-\vec{v}) &= \vec{0}_W - T(\vec{v}) = -T(\vec{v}). \end{aligned}$$
3. This result follows from preservation of addition and preservation of scalar multiplication. A formal proof would be by induction on m.

Consider the following example using the above theorem.

Example 9.6.2: Linear Combination

Let $T : \mathbb{P}_2 \to \mathbb{R}$ be a linear transformation such that
$$T(x^2 + x) = -1; T(x^2 - x) = 1; T(x^2 + 1) = 3.$$
Find $T(4x^2 + 5x - 3)$.

We provide two solutions to this problem.

Solution 1:

Suppose $a(x^2+x)+b(x^2-x)+c(x^2+1)=4x^2+5x-3$. Then

$$(a+b+c)x^2+(a-b)x+c=4x^2+5x-3.$$

Solving for a, b, and c results in the unique solution $a=6$, $b=1$, $c=-3$. Thus

$$\begin{aligned} T(4x^2+5x-3) &= T(6(x^2+x)+(x^2-x)-3(x^2+1)) \\ &= 6T(x^2+x)+T(x^2-x)-3T(x^2+1) \\ &= 6(-1)+1-3(3)=-14. \end{aligned}$$

Solution 2:

Notice that $S=\{x^2+x, x^2-x, x^2+1\}$ is a basis of $\mathbb{P}_2$, and thus x^2, x, and 1 can each be written as a linear combination of elements of S.

$$\begin{aligned} x^2 &= \tfrac{1}{2}(x^2+x)+\tfrac{1}{2}(x^2-x) \\ x &= \tfrac{1}{2}(x^2+x)-\tfrac{1}{2}(x^2-x) \\ 1 &= (x^2+1)-\tfrac{1}{2}(x^2+x)-\tfrac{1}{2}(x^2-x). \end{aligned}$$

Then

$$\begin{aligned} T(x^2) &= T\left(\tfrac{1}{2}(x^2+x)+\tfrac{1}{2}(x^2-x)\right)=\tfrac{1}{2}T(x^2+x)+\tfrac{1}{2}T(x^2-x) \\ &= \tfrac{1}{2}(-1)+\tfrac{1}{2}(1)=0. \\ T(x) &= T\left(\tfrac{1}{2}(x^2+x)-\tfrac{1}{2}(x^2-x)\right)=\tfrac{1}{2}T(x^2+x)-\tfrac{1}{2}T(x^2-x) \\ &= \tfrac{1}{2}(-1)-\tfrac{1}{2}(1)=-1. \\ T(1) &= T\left((x^2+1)-\tfrac{1}{2}(x^2+x)-\tfrac{1}{2}(x^2-x)\right) \\ &= T(x^2+1)-\tfrac{1}{2}T(x^2+x)-\tfrac{1}{2}T(x^2-x) \\ &= 3-\tfrac{1}{2}(-1)-\tfrac{1}{2}(1)=3. \end{aligned}$$

Therefore,

$$\begin{aligned} T(4x^2+5x-3) &= 4T(x^2)+5T(x)-3T(1) \\ &= 4(0)+5(-1)-3(3)=-14. \end{aligned}$$

The advantage of **Solution 2** over **Solution 1** is that if you were now asked to find $T(-6x^2-13x+9)$, it is easy to use $T(x^2)=0$, $T(x)=-1$ and $T(1)=3$:

$$\begin{aligned} T(-6x^2-13x+9) &= -6T(x^2)-13T(x)+9T(1) \\ &= -6(0)-13(-1)+9(3)=13+27=40. \end{aligned}$$

More generally,

$$\begin{aligned} T(ax^2+bx+c) &= aT(x^2)+bT(x)+cT(1) \\ &= a(0)+b(-1)+c(3)=-b+3c. \end{aligned}$$

Below is a video on matching 2x2 matrices with R^2 transformations.



Below is a video on writing a matrix for the integration linear transformation.



Below is a video on writing a matrix for the horizontal shift linear transformation.



Below is a video on writing a matrix for a derivative linear transformation.



Suppose two linear transformations act in the same way on $\vec{v}$ for all vectors. Then we say that these transformations are equal.

Definition 9.6.2: Equal Transformations

Let S and T be linear transformations from V to W. Then $S = T$ if and only if for every $\vec{v} \in V$,

$$S(\vec{v}) = T(\vec{v})$$

The definition above requires that two transformations have the same action on every vector in order for them to be equal. The next theorem argues that it is only necessary to check the action of the transformations on basis vectors.

Theorem 9.6.2: Transformation of a Spanning Set

Let V and W be vector spaces and suppose that S and T are linear transformations from V to W. Then in order for S and T to be equal, it suffices that $S(\vec{v}_i) = T(\vec{v}_i)$ where $V = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$.

This theorem tells us that a linear transformation is completely determined by its actions on a spanning set. We can also examine the effect of a linear transformation on a basis.

Theorem 9.6.3: Transformation of a Basis

Suppose V and W are vector spaces and let $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n\}$ be any given vectors in W that may not be distinct. Then there exists a basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ of V and a unique linear transformation $T : V \mapsto W$ with $T(\vec{v}_i) = \vec{w}_i$.

Furthermore, if

$$\vec{v} = k_1\vec{v}_1 + k_2\vec{v}_2 + \cdots + k_n\vec{v}_n$$

is a vector of V, then

$$T(\vec{v}) = k_1\vec{w}_1 + k_2\vec{w}_2 + \cdots + k_n\vec{w}_n.$$

Below is a video on determining which sequences of linear transformations are valid (composition of linear transformations).



This page titled 9.6: Linear Transformations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **9.6: Linear Transformations** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

9.7: Isomorphisms

Outcomes

A. Apply the concepts of one to one and onto to transformations of vector spaces.
B. Determine if a linear transformation of vector spaces is an isomorphism.
C. Determine if two vector spaces are isomorphic.

One to One and Onto Transformations

Recall the following definitions, given here in terms of vector spaces.

Definition 9.7.1: One to One Transformation

Let V, W be vector spaces with $\vec{v}_1, \vec{v}_2$ vectors in V. Then a linear transformation $T : V \mapsto W$ is called **one to one** if whenever $\vec{v}_1 \neq \vec{v}_2$ it follows that

$$T(\vec{v}_1) \neq T(\vec{v}_2)$$

Definition 9.7.2: Onto Transformation

Let V, W be vector spaces. Then a linear transformation $T : V \mapsto W$ is called **onto** if for all $\vec{w} \in \vec{W}$ there exists $\vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$.

Recall that every linear transformation T has the property that $T(\vec{0}) = \vec{0}$. This will be necessary to prove the following useful lemma.

Lemma 9.7.1: One to One

The assertion that a linear transformation T is one to one is equivalent to saying that if $T(\vec{v}) = \vec{0}$, then $\vec{v} = 0$.

Proof

Suppose first that T is one to one.

$$T(\vec{0}) = T\left(\vec{0} + \vec{0}\right) = T(\vec{0}) + T(\vec{0})$$

and so, adding the additive inverse of $T(\vec{0})$ to both sides, one sees that $T(\vec{0}) = \vec{0}$. Therefore, if $T(\vec{v}) = \vec{0}$, it must be the case that $\vec{v} = \vec{0}$ because it was just shown that $T(\vec{0}) = \vec{0}$.

Now suppose that if $T(\vec{v}) = \vec{0}$, then $\vec{v} = 0$. If $T(\vec{v}) = T(\vec{u})$, then $T(\vec{v}) - T(\vec{u}) = T(\vec{v} - \vec{u}) = \vec{0}$ which shows that $\vec{v} - \vec{u} = 0$ or in other words, $\vec{v} = \vec{u}$.

Consider the following example.

Example 9.7.1: One to One Transformation

Let $S : \mathbb{P}_2 \to \mathbb{M}_{22}$ be a linear transformation defined by

$$S(ax^2 + bx + c) = \begin{bmatrix} a+b & a+c \\ b-c & b+c \end{bmatrix}$$

for all $ax^2 + bx + c \in \mathbb{P}_2$.

Prove that S is one to one but not onto.

Solution

By definition,

$$\ker(S) = \{ax^2 + bx + c \in \mathbb{P}_2 \mid a+b=0, a+c=0, b-c=0, b+c=0\}.$$

Suppose $p(x) = ax^2 + bx + c \in \ker(S)$. This leads to a homogeneous system of four equations in three variables. Putting the augmented matrix in reduced row-echelon form:

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The solution is $a = b = c = 0$. This tells us that if $S(p(x)) = 0$, then $p(x) = ax^2 + bx + c = 0x^2 + 0x + 0 = 0$. Therefore it is one to one.

To show that S is **not** onto, find a matrix $A \in \mathbb{M}_{22}$ such that for every $p(x) \in \mathbb{P}_2$, $S(p(x)) \neq A$. Let

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix},$$

and suppose $p(x) = ax^2 + bx + c \in \mathbb{P}_2$ is such that $S(p(x)) = A$. Then

$$\begin{array}{ll} a+b=0 & a+c=1 \\ b-c=0 & b+c=2 \end{array}$$

Solving this system

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 2 \end{array}\right].$$

Since the system is inconsistent, there is no $p(x) \in \mathbb{P}_2$ so that $S(p(x)) = A$, and therefore S is not onto.

Example 9.7.2: An Onto Transformation

Let $T : \mathbb{M}_{22} \to \mathbb{R}^2$ be a linear transformation defined by

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+d \\ b+c \end{bmatrix} \text{ for all } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{M}_{22}.$$

Prove that T is onto but not one to one.

Solution

Let $\begin{bmatrix} x \\ y \end{bmatrix}$ be an arbitrary vector in $\mathbb{R}^2$. Since $T\begin{bmatrix} x & y \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$, T is onto.

By Lemma 9.7.1 T is one to one if and only if $T(A) = \vec{0}$ implies that $A = 0$ the zero matrix. Observe that

$$T\left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right) = \begin{bmatrix} 1+-1 \\ 0+0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

There exists a nonzero matrix A such that $T(A) = \vec{0}$. It follows that T is not one to one.

The following example demonstrates that a one to one transformation preserves linear independence.

Example 9.7.3: One to One and Independence

Let V and W be vector spaces and $T : V \mapsto W$ a linear transformation. Prove that if T is one to one and $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is an independent subset of V, then $\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_k)\}$ is an independent subset of W.

Solution

Let $\vec{0}_V$ and $\vec{0}_W$ denote the zero vectors of V and W, respectively. Suppose that

$$a_1 T(\vec{v}_1) + a_2 T(\vec{v}_2) + \cdots + a_k T(\vec{v}_k) = \vec{0}_W$$

for some $a_1, a_2, \ldots, a_k \in \mathbb{R}$. Since linear transformations preserve linear combinations (addition and scalar multiplication),

$$T(a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k) = \vec{0}_W.$$

Now, since T is one to one, $\ker(T) = \{\vec{0}_V\}$, and thus

$$a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k = \vec{0}_V.$$

However, $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is independent so $a_1 = a_2 = \cdots = a_k = 0$. Therefore, $\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_k)\}$ is independent.

A similar claim can be made regarding onto transformations. In this case, an onto transformation preserves a spanning set.

Example 9.7.4: Onto and Spanning

Let V and W be vector spaces and $T : V \to W$ a linear transformation. Prove that if T is onto and $V = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, then

$$W = span\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_k)\}.$$

Solution

Suppose that T is onto and let $\vec{w} \in W$. Then there exists $\vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$. Since $V = span\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, there exist $a_1, a_2, \ldots a_k \in \mathbb{R}$ such that $\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k$. Using the fact that T is a linear transformation,

$$\begin{aligned} \vec{w} = T(\vec{v}) &= T(a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_k\vec{v}_k) \\ &= a_1 T(\vec{v}_1) + a_2 T(\vec{v}_2) + \cdots + a_k T(\vec{v}_k), \end{aligned}$$

i.e., $\vec{w} \in span\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_k)\}$, and thus

$$W \subseteq span\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_k)\}.$$

Since $T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_k) \in W$, it follows from that $span\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_k)\} \subseteq W$, and therefore $W = span\{T(\vec{v}_1), T(\vec{v}_2), \ldots, T(\vec{v}_k)\}$.

Below is a video on determining if a linear transformation is one to one and/or onto: R^2 to P_2.



Isomorphisms

The focus of this section is on linear transformations which are both one to one and onto. When this is the case, we call the transformation an isomorphism.

Definition 9.7.3: Isomorphism

Let V and W be two vector spaces and let $T : V \mapsto W$ be a linear transformation. Then T is called an **isomorphism** if the following two conditions are satisfied.

- T is one to one.
- T is onto.

Definition 9.7.4: Isomorphic

Let V and W be two vector spaces and let $T : V \mapsto W$ be a linear transformation. Then if T is an isomorphism, we say that V and W are **isomorphic.**

Consider the following example of an isomorphism.

Example 9.7.5: Isomorphism

Let $T : \mathbb{M}_{22} \to \mathbb{R}^4$ be defined by

$$T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \text{ for all } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{M}_{22}.$$

Show that T is an isomorphism.

Solution

Notice that if we can prove T is an isomorphism, it will mean that $\mathbb{M}_{22}$ and $\mathbb{R}^4$ are isomorphic. It remains to prove that

1. T is a linear transformation;
2. T is one-to-one;
3. T is onto.

***T* is linear:** Let k, p be scalars.

$$
\begin{aligned}
T\left(k\begin{bmatrix}a_1 & b_1\\ c_1 & d_1\end{bmatrix}+p\begin{bmatrix}a_2 & b_2\\ c_2 & d_2\end{bmatrix}\right) &= T\left(\begin{bmatrix}ka_1 & kb_1\\ kc_1 & kd_1\end{bmatrix}+\begin{bmatrix}pa_2 & pb_2\\ pc_2 & pd_2\end{bmatrix}\right)\\
&= T\left(\begin{bmatrix}ka_1+pa_2 & kb_1+pb_2\\ kc_1+pc_2 & kd_1+pd_2\end{bmatrix}\right)\\
&= \begin{bmatrix}ka_1+pa_2\\ kb_1+pb_2\\ kc_1+pc_2\\ kd_1+pd_2\end{bmatrix}\\
&= \begin{bmatrix}ka_1\\ kb_1\\ kc_1\\ kd_1\end{bmatrix}+\begin{bmatrix}pa_2\\ pb_2\\ pc_2\\ pd_2\end{bmatrix}\\
&= k\begin{bmatrix}a_1\\ b_1\\ c_1\\ d_1\end{bmatrix}+p\begin{bmatrix}a_2\\ b_2\\ c_2\\ d_2\end{bmatrix}\\
&= kT\left(\begin{bmatrix}a_1 & b_1\\ c_1 & d_1\end{bmatrix}\right)+pT\left(\begin{bmatrix}a_2 & b_2\\ c_2 & d_2\end{bmatrix}\right)
\end{aligned}
$$

Therefore T is linear.

***T* is one-to-one:** By Lemma 9.7.1 we need to show that if $T(A)=0$ then $A=0$ for some matrix $A\in\mathbb{M}_{22}$.

$$
T\begin{bmatrix}a & b\\ c & d\end{bmatrix}=\begin{bmatrix}a\\ b\\ c\\ d\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\\ 0\end{bmatrix}
$$

This clearly only occurs when $a=b=c=d=0$ which means that

$$
A=\begin{bmatrix}a & b\\ c & d\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}=0
$$

Hence T is one-to-one.

***T* is onto:** Let

$$
\vec{x}=\begin{bmatrix}x_1\\ x_2\\ x_3\\ x_4\end{bmatrix}\in\mathbb{R}^4,
$$

and define matrix $A\in\mathbb{M}_{22}$ as follows:

$$
A=\begin{bmatrix}x_1 & x_2\\ x_3 & x_4\end{bmatrix}.
$$

Then $T(A)=\vec{x}$, and therefore T is onto.

Since T is a linear transformation which is one-to-one and onto, T is an isomorphism. Hence $\mathbb{M}_{22}$ and $\mathbb{R}^4$ are isomorphic.

Below is a video on determining if a linear transformation is one to one and/or onto: P_2 to R^2.



An important property of isomorphisms is that the inverse of an isomorphism is itself an isomorphism and the composition of isomorphisms is an isomorphism. We first recall the definition of composition.

Definition 9.7.5: Composition of Transformations

Let V, W, Z be vector spaces and suppose $T : V \mapsto W$ and $S : W \mapsto Z$ are linear transformations. Then the composite of S and T is

$$S \circ T : V \mapsto Z$$

and is defined by

$$(S \circ T)(\vec{v}) = S(T(\vec{v})) \text{ for all } \vec{v} \in V$$

Consider now the following proposition.

Proposition 9.7.1: Composite and Inverse Isomorphism

Let $T : V \to W$ be an isomorphism. Then $T^{-1} : W \to V$ is also an isomorphism. Also if $T : V \to W$ is an isomorphism and if $S : W \to Z$ is an isomorphism for the vector spaces V, W, Z, then $S \circ T$ defined by $(S \circ T)(v) = S(T(v))$ is also an isomorphism.

Proof

Consider the first claim. Since T is onto, a typical vector in W is of the form $T(\vec{v})$ where $\vec{v} \in V$. Consider then for a, b scalars,

$$T^{-1}\left(aT(\vec{v}_1) + bT(\vec{v}_2)\right)$$

where $\vec{v}_1, \vec{v}_2 \in V$. Consider if this is equal to

$$aT^{-1}\left(T(\vec{v}_1)\right) + bT^{-1}\left(T(\vec{v}_2)\right) = a\vec{v}_1 + b\vec{v}_2?$$

Since T is one to one, this will be so if

$$T\left(a\vec{v}_1 + b\vec{v}_2\right) = T\left(T^{-1}\left(aT(\vec{v}_1) + bT(\vec{v}_2)\right)\right) = aT(\vec{v}_1) + bT(\vec{v}_2)$$

However, the above statement is just the condition that T is a linear map. Thus T^{-1} is indeed a linear map. If $\vec{v} \in V$ is given, then $\vec{v} = T^{-1}\left(T(\vec{v})\right)$ and so T^{-1} is onto. If $T^{-1}(\vec{v}) = \vec{0}$, then

$$\vec{v} = T\left(T^{-1}(\vec{v})\right) = T(\vec{0}) = \vec{0}$$

and so T^{-1} is one to one.

Next suppose T and S are as described. Why is $S \circ T$ a linear map? Let for a, b scalars,

$$\begin{aligned} S\circ T\left(a\vec{v}_1+b\vec{v}_2\right) &\equiv S\left(T\left(a\vec{v}_1+b\vec{v}_2\right)\right)=S\left(aT(\vec{v}_1)+bT(\vec{v}_2)\right) \\ &= aS\left(T(\vec{v}_1)\right)+bS\left(T(\vec{v}_2)\right)\equiv a\left(S\circ T\right)\left(\vec{v}_1\right)+b\left(S\circ T\right)\left(\vec{v}_2\right) \end{aligned}$$

Hence $S\circ T$ is a linear map. If $\left(S\circ T\right)\left(\vec{v}\right)=0$, then $S\left(T\left(\vec{v}\right)\right)=\vec{0}$ and it follows that $T(\vec{v})=\vec{0}$ and hence by this lemma again, $\vec{v}=\vec{0}$. Thus $S\circ T$ is one to one. It remains to verify that it is onto. Let $\vec{z}\in Z$. Then since S is onto, there exists $\vec{w}\in W$ such that $S(\vec{w})=\vec{z}$. Also, since T is onto, there exists $\vec{v}\in V$ such that $T(\vec{v})=\vec{w}$. It follows that $S\left(T\left(\vec{v}\right)\right)=\vec{z}$ and so $S\circ T$ is also onto.

Suppose we say that two vector spaces V and W are related if there exists an isomorphism of one to the other, written as $V\sim W$. Then the above proposition suggests that $\sim$ is an equivalence relation. That is: $\sim$ satisfies the following conditions:

- $V\sim V$
- If $V\sim W$, it follows that $W\sim V$
- If $V\sim W$ and $W\sim Z$, then $V\sim Z$

We leave the proof of these to the reader.

The following fundamental lemma describes the relation between bases and isomorphisms.

Lemma 9.7.2: Bases and Isomorphisms

Let $T:V\to W$ be a linear map where V,W are vector spaces. Then a linear transformation T which is one to one has the property that if $\{\vec{u}_1,\cdots,\vec{u}_k\}$ is linearly independent, then so is $\{T(\vec{u}_1),\cdots,T(\vec{u}_k)\}$. More generally, T is an isomorphism if and only if whenever $\{\vec{v}_1,\cdots,\vec{v}_n\}$ is a basis for V, it follows that $\{T(\vec{v}_1),\cdots,T(\vec{v}_n)\}$ is a basis for W.

Proof

First suppose that T is a linear map and is one to one and $\{\vec{u}_1,\cdots,\vec{u}_k\}$ is linearly independent. It is required to show that $\{T(\vec{u}_1),\cdots,T(\vec{u}_k)\}$ is also linearly independent. Suppose then that

$$\sum_{i=1}^{k}c_iT(\vec{u}_i)=\vec{0}$$

Then, since T is linear,

$$T\left(\sum_{i=1}^{n}c_i\vec{u}_i\right)=\vec{0}$$

Since T is one to one, it follows that

$$\sum_{i=1}^{n}c_i\vec{u}_i=0$$

Now the fact that $\{\vec{u}_1,\cdots,\vec{u}_n\}$ is linearly independent implies that each $c_i=0$. Hence $\{T(\vec{u}_1),\cdots,T(\vec{u}_n)\}$ is linearly independent.

Now suppose that T is an isomorphism and $\{\vec{v}_1,\cdots,\vec{v}_n\}$ is a basis for V. It was just shown that $\{T(\vec{v}_1),\cdots,T(\vec{v}_n)\}$ is linearly independent. It remains to verify that the span of $\{T(\vec{v}_1),\cdots,T(\vec{v}_n)\}$ is all of W. This is where T is onto is used. If $\vec{w}\in W$, there exists $\vec{v}\in V$ such that $T(\vec{v})=\vec{w}$. Since $\{\vec{v}_1,\cdots,\vec{v}_n\}$ is a basis, it follows that there exists scalars $\{c_i\}_{i=1}^{n}$ such that

$$\sum_{i=1}^{n}c_i\vec{v}_i=\vec{v}.$$

Hence,

$$\vec{w}=T(\vec{v})=T\left(\sum_{i=1}^{n}c_i\vec{v}_i\right)=\sum_{i=1}^{n}c_iT\vec{v}_i$$

which shows that the span of these vectors $\{T(\vec{v}_1),\cdots,T(\vec{v}_n)\}$ is all of W showing that this set of vectors is a basis for W.

Next suppose that T is a linear map which takes a basis to a basis. Then for $\{\vec{v}_1,\cdots,\vec{v}_n\}$ a basis for V, it follows $\{T(\vec{v}_1),\cdots,T(\vec{v}_n)\}$ is a basis for W. Then if $w\in W$, there exist scalars c_i such that $w=\sum_{i=1}^n c_iT(\vec{v}_i)=T\left(\sum_{i=1}^n c_i\vec{v}_i\right)$ showing that T is onto. If $T\left(\sum_{i=1}^n c_i\vec{v}_i\right)=0$ then $\sum_{i=1}^n c_iT(\vec{v}_i)=\vec{0}$ and since the vectors $\{T(\vec{v}_1),\cdots,T(\vec{v}_n)\}$ are linearly independent, it follows that each $c_i=0$. Since $\sum_{i=1}^n c_i\vec{v}_i$ is a typical vector in V, this has shown that if $T(\vec{v})=0$ then $\vec{v}=\vec{0}$ and so T is also one to one. Thus T is an isomorphism.

The following theorem illustrates a very useful idea for defining an isomorphism. Basically, if you know what it does to a basis, then you can construct the isomorphism.

Theorem 9.7.1: Isomorphic Vector Spaces

Suppose V and W are two vector spaces. Then the two vector spaces are isomorphic if and only if they have the same dimension. In the case that the two vector spaces have the same dimension, then for a linear transformation $T:V\to W$, the following are equivalent.

1. T is one to one.
2. T is onto.
3. T is an isomorphism.

Proof

Suppose first these two vector spaces have the same dimension. Let a basis for V be $\{\vec{v}_1,\cdots,\vec{v}_n\}$ and let a basis for W be $\{\vec{w}_1,\cdots,\vec{w}_n\}$. Now define T as follows.

$$T(\vec{v}_i)=\vec{w}_i$$

for $\sum_{i=1}^n c_i\vec{v}_i$ an arbitrary vector of V,

$$T\left(\sum_{i=1}^n c_i\vec{v}_i\right)=\sum_{i=1}^n c_iT(\vec{v}_i)=\sum_{i=1}^n c_i\vec{w}_i.$$

It is necessary to verify that this is well defined. Suppose then that

$$\sum_{i=1}^n c_i\vec{v}_i=\sum_{i=1}^n \hat{c}_i\vec{v}_i$$

Then

$$\sum_{i=1}^n (c_i-\hat{c}_i)\,\vec{v}_i=0$$

and since $\{\vec{v}_1,\cdots,\vec{v}_n\}$ is a basis, $c_i=\hat{c}_i$ for each i. Hence

$$\sum_{i=1}^n c_i\vec{w}_i=\sum_{i=1}^n \hat{c}_i\vec{w}_i$$

and so the mapping is well defined. Also if a,b are scalars,

$$\begin{aligned}T\left(a\sum_{i=1}^n c_i\vec{v}_i+b\sum_{i=1}^n \hat{c}_i\vec{v}_i\right) &= T\left(\sum_{i=1}^n (ac_i+b\hat{c}_i)\,\vec{v}_i\right)=\sum_{i=1}^n (ac_i+b\hat{c}_i)\,\vec{w}_i\\ &= a\sum_{i=1}^n c_i\vec{w}_i+b\sum_{i=1}^n \hat{c}_i\vec{w}_i\\ &= aT\left(\sum_{i=1}^n c_i\vec{v}_i\right)+bT\left(\sum_{i=1}^n \hat{c}_i\vec{v}_i\right)\end{aligned}$$

Thus T is a linear map.

Now if

$$T\left(\sum_{i=1}^{n} c_i \vec{v}_i\right) = \sum_{i=1}^{n} c_i \vec{w}_i = \vec{0},$$

then since the $\{\vec{w}_1, \cdots, \vec{w}_n\}$ are independent, each $c_i = 0$ and so $\sum_{i=1}^{n} c_i \vec{v}_i = \vec{0}$ also. Hence T is one to one. If $\sum_{i=1}^{n} c_i \vec{w}_i$ is a vector in W, then it equals

$$\sum_{i=1}^{n} c_i T\vec{v}_i = T\left(\sum_{i=1}^{n} c_i \vec{v}_i\right)$$

showing that T is also onto. Hence T is an isomorphism and so V and W are isomorphic.

Next suppose these two vector spaces are isomorphic. Let T be the name of the isomorphism. Then for $\{\vec{v}_1, \cdots, \vec{v}_n\}$ a basis for V, it follows that a basis for W is $\{T\vec{v}_1, \cdots, T\vec{v}_n\}$ showing that the two vector spaces have the same dimension.

Now suppose the two vector spaces have the same dimension.

First consider the claim that $1.) \Rightarrow 2.)$. If T is one to one, then if $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for V, then $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is linearly independent. If it is not a basis, then it must fail to span W. But then there would exist $\vec{w} \notin span\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ and it follows that $\{T(\vec{v}_1), \cdots, T(\vec{v}_n), \vec{w}\}$ would be linearly independent which is impossible because there exists a basis for W of n vectors. Hence

$$span\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\} = W$$

and so $\{T(\vec{v}_1), \cdots, T(\vec{v}_n)\}$ is a basis. Hence, if $\vec{w} \in W$, there exist scalars c_i such that

$$\vec{w} = \sum_{i=1}^{n} c_i T(\vec{v}_i) = T\left(\sum_{i=1}^{n} c_i \vec{v}_i\right)$$

showing that T is onto. This shows that $1.) \Rightarrow 2.)$.

Next consider the claim that $2.) \Rightarrow 3.)$. Since $2.)$ holds, it follows that T is onto. It remains to verify that T is one to one. Since T is onto, there exists a basis of the form $\{T(\vec{v}_i), \cdots, T(\vec{v}_n)\}$. If $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is linearly independent, then this set of vectors must also be a basis for V because if not, there would exist $\vec{u} \notin span\{\vec{v}_1, \cdots, \vec{v}_n\}$ so $\{\vec{v}_1, \cdots, \vec{v}_n, \vec{u}\}$ would be a linearly independent set which is impossible because by assumption, there exists a basis which has n vectors. So why is $\{\vec{v}_1, \cdots, \vec{v}_n\}$ linearly independent? Suppose

$$\sum_{i=1}^{n} c_i \vec{v}_i = \vec{0}$$

Then

$$\sum_{i=1}^{n} c_i T\vec{v}_i = \vec{0}$$

Hence each $c_i = 0$ and so, as just discussed, $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a basis for V. Now it follows that a typical vector in V is of the form $\sum_{i=1}^{n} c_i \vec{v}_i$. If $T\left(\sum_{i=1}^{n} c_i \vec{v}_i\right) = \vec{0}$, it follows that

$$\sum_{i=1}^{n} c_i T(\vec{v}_i) = \vec{0}$$

and so, since $\{T(\vec{v}_i), \cdots, T(\vec{v}_n)\}$ is independent, it follows each $c_i = 0$ and hence $\sum_{i=1}^{n} c_i \vec{v}_i = \vec{0}$. Thus T is one to one as well as onto and so it is an isomorphism.

If T is an isomorphism, it is both one to one and onto by definition so $3.)$ implies both $1.)$ and $2.)$.

Note the interesting way of defining a linear transformation in the first part of the argument by describing what it does to a basis and then “extending it linearly”.

Consider the following example.

Example 9.7.6

Let $V = \mathbb{R}^3$ and let W denote the polynomials of degree at most 2. Show that these two vector spaces are isomorphic.

Solution

First, observe that a basis for W is $\{1, x, x^2\}$ and a basis for V is $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. Since these two have the same dimension, the two are isomorphic. An example of an isomorphism is this:

$$T(\vec{e}_1) = 1, T(\vec{e}_2) = x, T(\vec{e}_3) = x^2$$

and extend T linearly as in the above proof. Thus

$$T(a, b, c) = a + bx + cx^2$$

This page titled 9.7: Isomorphisms is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **9.7: Isomorphisms** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

9.8: The Kernel and Image of a Linear Map

Outcomes

A. Describe the kernel and image of a linear transformation.
B. Use the kernel and image to determine if a linear transformation is one to one or onto.

Here we consider the case where the linear map is not necessarily an isomorphism. First here is a definition of what is meant by the image and kernel of a linear transformation.

Definition 9.8.1: Kernel and Image

Let V and W be vector spaces and let $T : V \to W$ be a linear transformation. Then the image of T denoted as $\text{im}(T)$ is defined to be the set

$$\{T(\vec{v}) : \vec{v} \in V\}$$

In words, it consists of all vectors in W which equal $T(\vec{v})$ for some $\vec{v} \in V$. The kernel, $\ker(T)$, consists of all $\vec{v} \in V$ such that $T(\vec{v}) = \vec{0}$. That is,

$$\ker(T) = \left\{\vec{v} \in V : T(\vec{v}) = \vec{0}\right\}$$

Then in fact, both $\text{im}(T)$ and $\ker(T)$ are subspaces of W and V respectively.

Proposition 9.8.1: Kernel and Image as Subspaces

Let V, W be vector spaces and let $T : V \to W$ be a linear transformation. Then $\ker(T) \subseteq V$ and $\text{im}(T) \subseteq W$. In fact, they are both subspaces.

Proof

First consider $\ker(T)$. It is necessary to show that if $\vec{v}_1, \vec{v}_2$ are vectors in $\ker(T)$ and if a, b are scalars, then $a\vec{v}_1 + b\vec{v}_2$ is also in $\ker(T)$. But

$$T\left(a\vec{v}_1 + b\vec{v}_2\right) = aT(\vec{v}_1) + bT(\vec{v}_2) = a\vec{0} + b\vec{0} = \vec{0}$$

Thus $\ker(T)$ is a subspace of V.

Next suppose $T(\vec{v}_1), T(\vec{v}_2)$ are two vectors in $\text{im}(T)$. Then if a, b are scalars,

$$aT(\vec{v}_2) + bT(\vec{v}_2) = T\left(a\vec{v}_1 + b\vec{v}_2\right)$$

and this last vector is in $\text{im}(T)$ by definition.

Consider the following example.

Example 9.8.1: Kernel and Image of a Transformation

Let $T : \mathbb{P}_1 \to \mathbb{R}$ be the linear transformation defined by

$$T(p(x)) = p(1) \text{ for all } p(x) \in \mathbb{P}_1.$$

Find the kernel and image of T.

Solution

We will first find the kernel of T. It consists of all polynomials in $\mathbb{P}_1$ that have 1 for a root.

$$\begin{aligned}\ker(T) &= \{p(x) \in \mathbb{P}_1 \mid p(1) = 0\} \\ &= \{ax + b \mid a, b \in \mathbb{R} \text{ and } a + b = 0\} \\ &= \{ax - a \mid a \in \mathbb{R}\}\end{aligned}$$

Therefore a basis for $\ker(T)$ is

$$\{x - 1\}$$

Notice that this is a subspace of $\mathbb{P}_1$.

Now consider the image. It consists of all numbers which can be obtained by evaluating all polynomials in $\mathbb{P}_1$ at 1.

$$\begin{aligned}\text{im}(T) &= \{p(1) \mid p(x) \in \mathbb{P}_1\} \\ &= \{a + b \mid ax + b \in \mathbb{P}_1\} \\ &= \{a + b \mid a, b \in \mathbb{R}\} \\ &= \mathbb{R}\end{aligned}$$

Therefore a basis for $\text{im}(T)$ is

$$\{1\}$$

Notice that this is a subspace of $\mathbb{R}$, and in fact is the space $\mathbb{R}$ itself.

Below is a video on finding a nontrivial matrix for the kernel of a linear transformation: M_{2x2} to M_{2x2}.



Example 9.8.2: Kernel and Image of a Linear Transformation

Let $T : \mathbb{M}_{22} \mapsto \mathbb{R}^2$ be defined by

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a - b \\ c + d \end{bmatrix}$$

Then T is a linear transformation. Find a basis for $\ker(T)$ and $\text{im}(T)$.

Solution

You can verify that T represents a linear transformation.

Now we want to find a way to describe all matrices A such that $T(A) = \vec{0}$, that is the matrices in $\ker(T)$. Suppose $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is such a matrix. Then

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a - b \\ c + d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The values of a, b, c, d that make this true are given by solutions to the system

$$a-b=0$$
$$c+d=0$$

The solution is $a=s, b=s, c=t, d=-t$ where s,t are scalars. We can describe $\ker(T)$ as follows.

$$\ker(T)=\left\{\begin{bmatrix} s & s \\ t & -t \end{bmatrix}\right\}=\text{span}\left\{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}\right\}$$

It is clear that this set is linearly independent and therefore forms a basis for $\ker(T)$.

We now wish to find a basis for $\text{im}(T)$. We can write the image of T as

$$\text{im}(T)=\left\{\begin{bmatrix} a-b \\ c+d \end{bmatrix}\right\}$$

Notice that this can be written as

$$\text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix},\begin{bmatrix} -1 \\ 0 \end{bmatrix},\begin{bmatrix} 0 \\ 1 \end{bmatrix},\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$$

However this is clearly not linearly independent. By removing vectors from the set to create an independent set gives a basis of $\text{im}(T)$.

$$\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix},\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$$

Notice that these vectors have the same span as the set above but are now linearly independent.

Below is a video on finding a nontrivial matrix for the kernel of a linear transformation: P_2 to R^2.



A major result is the relation between the dimension of the kernel and dimension of the image of a linear transformation. A special case was done earlier in the context of matrices. Recall that for an $m\times n$ matrix it was the case that the dimension of the kernel of A added to the rank of A equals n.

Theorem 9.8.1: Dimension of Kernel + Image

Let $T:V\to W$ be a linear transformation where V,W are vector spaces. Suppose the dimension of V is n. Then $n=\dim(\ker(T))+\dim(\text{im}\,(T))$.

Proof

From Proposition 9.8.1, $\text{im}\,(T)$ is a subspace of W. By Theorem 9.4.8, there exists a basis for $\text{im}\,(T)$, $\{T(\vec{v}_1),\cdots,T(\vec{v}_r)\}$. Similarly, there is a basis for $\ker(T)$, $\{\vec{u}_1,\cdots,\vec{u}_s\}$. Then if $\vec{v}\in V$, there exist scalars c_i such that

$$T(\vec{v}) = \sum_{i=1}^{r} c_i T(\vec{v}_i)$$

Hence $T\left(\vec{v} - \sum_{i=1}^{r} c_i \vec{v}_i\right) = 0$. It follows that $\vec{v} - \sum_{i=1}^{r} c_i \vec{v}_i$ is in $\ker(T)$. Hence there are scalars a_i such that

$$\vec{v} - \sum_{i=1}^{r} c_i \vec{v}_i = \sum_{j=1}^{s} a_j \vec{u}_j$$

Hence $\vec{v} = \sum_{i=1}^{r} c_i \vec{v}_i + \sum_{j=1}^{s} a_j \vec{u}_j$. Since $\vec{v}$ is arbitrary, it follows that

$$V = \text{span}\left\{\vec{u}_1, \cdots, \vec{u}_s, \vec{v}_1, \cdots, \vec{v}_r\right\}$$

If the vectors $\left\{\vec{u}_1, \cdots, \vec{u}_s, \vec{v}_1, \cdots, \vec{v}_r\right\}$ are linearly independent, then it will follow that this set is a basis. Suppose then that

$$\sum_{i=1}^{r} c_i \vec{v}_i + \sum_{j=1}^{s} a_j \vec{u}_j = 0$$

Apply T to both sides to obtain

$$\sum_{i=1}^{r} c_i T(\vec{v}_i) + \sum_{j=1}^{s} a_j T(\vec{u}_j) = \sum_{i=1}^{r} c_i T(\vec{v}_i) = \vec{0}$$

Since $\{T(\vec{v}_1), \cdots, T(\vec{v}_r)\}$ is linearly independent, it follows that each $c_i = 0$. Hence $\sum_{j=1}^{s} a_j \vec{u}_j = 0$ and so, since the $\left\{\vec{u}_1, \cdots, \vec{u}_s\right\}$ are linearly independent, it follows that each $a_j = 0$ also. It follows that $\left\{\vec{u}_1, \cdots, \vec{u}_s, \vec{v}_1, \cdots, \vec{v}_r\right\}$ is a basis for V and so

$$n = s + r = \dim(\ker(T)) + \dim(\text{im}\,(T))$$

Below is a video on finding a basis for the image and kernel of a linear transformation: R^3 to P_3.



Below is a video on finding a basis for the image and kernel of a linear transformation: M_{2x2} to R^3.



Consider the following definition.

Definition 9.8.2: Rank of Linear Transformation

Let $T : V \to W$ be a linear transformation and suppose V, W are finite dimensional vector spaces. Then the rank of T denoted as $\text{rank}\,(T)$ is defined as the dimension of $\text{im}\,(T)$. The nullity of T is the dimension of $\ker(T)$. Thus the above theorem says that $\text{rank}\,(T) + \dim(\ker(T)) = \dim(V)$.

Recall the following important result.

Theorem 9.8.2: Subspace of Same Dimension

Let V be a vector space of dimension n and let W be a subspace. Then $W = V$ if and only if the dimension of W is also n.

From this theorem follows the next corollary.

Corollary 9.8.1: One to One and Onto Characterization

Let $T : V \to W$ be a linear map where the dimension of V is n and the dimension of W is m. Then T is one to one if and only if $\ker(T) = \left\{\vec{0}\right\}$ and T is onto if and only if $\text{rank}\,(T) = m$.

Proof

The statement $\ker(T) = \left\{\vec{0}\right\}$ is equivalent to saying if $T\left(\vec{v}\right) = \vec{0},$ it follows that $\vec{v} = \vec{0}$. Thus by Lemma 9.7.1 T is one to one. If T is onto, then $\text{im}\,(T) = W$ and so $\text{rank}\,(T)$ which is defined as the dimension of $\text{im}\,(T)$ is m. If $\text{rank}\,(T) = m$, then by Theorem 9.8.2 since $\text{im}\,(T)$ is a subspace of W, it follows that $\text{im}\,(T) = W$.

Example 9.8.3: One to One Transformation

Let $S : \mathbb{P}_2 \to \mathbb{M}_{22}$ be a linear transformation defined by

$$S(ax^2 + bx + c) = \begin{bmatrix} a+b & a+c \\ b-c & b+c \end{bmatrix} \text{ for all } ax^2 + bx + c \in \mathbb{P}_2.$$

Prove that S is one to one but not onto.

Solution

You may recall this example from earlier in Example 9.7.1. Here we will determine that S is one to one, but not onto, using the method provided in Corollary 9.8.1.

By definition,

$$\ker(S) = \{ax^2 + bx + c \in \mathbb{P}_2 \mid a+b=0, a+c=0, b-c=0, b+c=0\}.$$

Suppose $p(x) = ax^2 + bx + c \in \ker(S)$. This leads to a homogeneous system of four equations in three variables. Putting the augmented matrix in reduced row-echelon form:

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Since the unique solution is $a = b = c = 0$, $\ker(S) = \{\vec{0}\}$, and thus S is one-to-one by Corollary 9.8.1.

Similarly, by Corollary 9.8.1, if S is onto it will have $\text{rank}(S) = \dim(\mathbb{M}_{22}) = 4$. The image of S is given by

$$\text{im}(S) = \left\{ \begin{bmatrix} a+b & a+c \\ b-c & b+c \end{bmatrix} \right\} = \text{span} \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \right\}$$

These matrices are linearly independent which means this set forms a basis for $\text{im}(S)$. Therefore the dimension of $\text{im}(S)$, also called $\text{rank}(S)$, is equal to 3. It follows that S is not onto.

This page titled 9.8: The Kernel and Image of a Linear Map is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **9.8: The Kernel and Image of a Linear Map** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

9.9: The Matrix of a Linear Transformation

Outcomes

A. Find the matrix of a linear transformation with respect to general bases in vector spaces.

You may recall from $\mathbb{R}^n$ that the matrix of a linear transformation depends on the bases chosen. This concept is explored in this section, where the linear transformation now maps from one arbitrary vector space to another.

Let $T : V \mapsto W$ be an isomorphism where V and W are vector spaces. Recall from Lemma 9.7.2 that T maps a basis in V to a basis in W. When discussing this Lemma, we were not specific on what this basis looked like. In this section we will make such a distinction.

Consider now an important definition.

Definition 9.9.1: Coordinate Isomorphism

Let V be a vector space with $\dim(V) = n$, let $B = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ be a fixed basis of V, and let $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ denote the standard basis of $\mathbb{R}^n$. We define a transformation $C_B : V \to \mathbb{R}^n$ by

$$C_B(a_1\vec{b}_1 + a_2\vec{b}_2 + \cdots + a_n\vec{b}_n) = a_1\vec{e}_1 + a_2\vec{e}_2 + \cdots + a_n\vec{e}_n = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Then C_B is a linear transformation such that $C_B(\vec{b}_i) = \vec{e}_i$, $1 \leq i \leq n$.

C_B is an isomorphism, called the coordinate isomorphism corresponding to B.

We continue with another related definition.

Definition 9.9.2: Coordinate Vector

Let V be a finite dimensional vector space with $\dim(V) = n$, and let $B = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ be an ordered basis of V (meaning that the order that the vectors are listed is taken into account). The coordinate vector of $\vec{v}$ with respect to B is defined as $C_B(\vec{v})$.

coordinatevector Consider the following example.

Example 9.9.1: Coordinate Vector

Let $V = \mathbb{P}_2$ and $\vec{x} = -x^2 - 2x + 4$. Find $C_B(\vec{x})$ for the following bases B:

1. $B = \{1, x, x^2\}$
2. $B = \{x^2, x, 1\}$
3. $B = \{x + x^2, x, 4\}$

Solution

1. First, note the order of the basis is important. Now we need to find a_1, a_2, a_3 such that $\vec{x} = a_1(1) + a_2(x) + a_3(x^2)$, that is:

$$-x^2 - 2x + 4 = a_1(1) + a_2(x) + a_3(x^2)$$

Clearly the solution is

$$\begin{aligned} a_1 &= 4 \\ a_2 &= -2 \\ a_3 &= -1 \end{aligned}$$

Therefore the coordinate vector is

$$C_B(\vec{x}) = \begin{bmatrix} 4 \\ -2 \\ -1 \end{bmatrix}$$

2. Again remember that the order of B is important. We proceed as above. We need to find a_1, a_2, a_3 such that $\vec{x} = a_1(x^2) + a_2(x) + a_3(1)$, that is:

$$-x^2 - 2x + 4 = a_1(x^2) + a_2(x) + a_3(1)$$

Here the solution is

$$\begin{aligned} a_1 &= -1 \\ a_2 &= -2 \\ a_3 &= 4 \end{aligned}$$

Therefore the coordinate vector is

$$C_B(\vec{x}) = \begin{bmatrix} -1 \\ -2 \\ 4 \end{bmatrix}$$

3. Now we need to find a_1, a_2, a_3 such that $\vec{x} = a_1(x + x^2) + a_2(x) + a_3(4)$, that is:

$$\begin{aligned} -x^2 - 2x + 4 &= a_1(x + x^2) + a_2(x) + a_3(4) \\ &= a_1(x^2) + (a_1 + a_2)(x) + a_3(4) \end{aligned}$$

The solution is

$$\begin{aligned} a_1 &= -1 \\ a_2 &= -1 \\ a_3 &= 1 \end{aligned}$$

and the coordinate vector is

$$C_B(\vec{x}) = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

Given that the coordinate transformation $C_B : V \to \mathbb{R}^n$ is an isomorphism, its inverse exists.

Theorem 9.9.1: Inverse of the Coordinate Isomorphism

Let V be a finite dimensional vector space with dimension n and ordered basis $B = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$. Then $C_B : V \to \mathbb{R}^n$ is an isomorphism whose inverse,

$$C_B^{-1} : \mathbb{R}^n \to V$$

is given by

$$C_B^{-1} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a_1\vec{b}_1 + a_2\vec{b}_2 + \cdots + a_n\vec{b}_n \text{ for all } \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n.$$

We now discuss the main result of this section, that is how to represent a linear transformation with respect to different bases.

Let V and W be finite dimensional vector spaces, and suppose

- $\dim(V) = n$ and $B_1 = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ is an ordered basis of V;
- $\dim(W) = m$ and B_2 is an ordered basis of W.

Let $T : V \to W$ be a linear transformation. If $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$, then we can find a matrix A so that $T_A = T$. For arbitrary vector spaces V and W, our goal is to represent T as a matrix., i.e., find a matrix A so that $T_A : \mathbb{R}^n \to \mathbb{R}^m$ and $T_A = C_{B_2} T C_{B_1}^{-1}$.

To find the matrix A:

$$T_A = C_{B_2} T C_{B_1}^{-1} \text{ implies that } T_A C_{B_1} = C_{B_2} T,$$

and thus for any $\vec{v} \in V$,

$$C_{B_2}[T(\vec{v})] = T_A[C_{B_1}(\vec{v})] = AC_{B_1}(\vec{v}).$$

Since $C_{B_1}(\vec{b}_j) = \vec{e}_j$ for each $\vec{b}_j \in B_1$, $AC_{B_1}(\vec{b}_j) = A\vec{e}_j$, which is simply the jth column of A. Therefore, the jth column of A is equal to $C_{B_2}[T(\vec{b}_j)]$.

The matrix of T corresponding to the ordered bases B_1 and B_2 is denoted $M_{B_2B_1}(T)$ and is given by

$$M_{B_2B_1}(T) = \begin{bmatrix} C_{B_2}[T(\vec{b}_1)] & C_{B_2}[T(\vec{b}_2)] & \cdots & C_{B_2}[T(\vec{b}_n)] \end{bmatrix}.$$

This result is given in the following theorem.

Theorem 9.9.2

Let V and W be vectors spaces of dimension n and m respectively, with $B_1 = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\}$ an ordered basis of V and B_2 an ordered basis of W. Suppose $T : V \to W$ is a linear transformation. Then the unique matrix $M_{B_2B_1}(T)$ of T corresponding to B_1 and B_2 is given by

$$M_{B_2B_1}(T) = \begin{bmatrix} C_{B_2}[T(\vec{b}_1)] & C_{B_2}[T(\vec{b}_2)] & \cdots & C_{B_2}[T(\vec{b}_n)] \end{bmatrix}.$$

This matrix satisfies $C_{B_2}[T(\vec{v})] = M_{B_2B_1}(T)C_{B_1}(\vec{v})$ for all $\vec{v} \in V$.

We demonstrate this content in the following examples.

Example 9.9.2: Matrix of a Linear Transformation

Let $T : \mathbb{P}_3 \mapsto \mathbb{R}^4$ be an isomorphism defined by

$$T(ax^3 + bx^2 + cx + d) = \begin{bmatrix} a+b \\ b-c \\ c+d \\ d+a \end{bmatrix}$$

Suppose $B_1 = \{x^3, x^2, x, 1\}$ is an ordered basis of $\mathbb{P}_3$ and

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

be an ordered basis of $\mathbb{R}^4$. Find the matrix $M_{B_2B_1}(T)$.

Solution

To find $M_{B_2B_1}(T)$, we use the following definition.

$$M_{B_2B_1}(T) = \begin{bmatrix} C_{B_2}[T(x^3)] & C_{B_2}[T(x^2)] & C_{B_2}[T(x)] & C_{B_2}[T(x^2)] \end{bmatrix}$$

First we find the result of applying T to the basis B_1.

$$T(x^3)=\begin{bmatrix}1\\0\\0\\1\end{bmatrix}, T(x^2)=\begin{bmatrix}1\\1\\0\\0\end{bmatrix}, T(x)=\begin{bmatrix}0\\-1\\1\\0\end{bmatrix}, T(1)=\begin{bmatrix}0\\0\\1\\1\end{bmatrix}$$

Next we apply the coordinate isomorphism C_{B_2} to each of these vectors. We will show the first in detail.

$$C_{B_2}\left(\begin{bmatrix}1\\0\\0\\1\end{bmatrix}\right)=a_1\begin{bmatrix}1\\0\\1\\0\end{bmatrix}+a_2\begin{bmatrix}0\\1\\0\\0\end{bmatrix}+a_3\begin{bmatrix}0\\0\\-1\\0\end{bmatrix}+a_4\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

This implies that

$$\begin{aligned}a_1&=1\\a_2&=0\\a_1-a_3&=0\\a_4&=1\end{aligned}$$

which has a solution given by

$$\begin{aligned}a_1&=1\\a_2&=0\\a_3&=1\\a_4&=1\end{aligned}$$

Therefore $C_{B_2}[T(x^3)]=\begin{bmatrix}1\\0\\1\\1\end{bmatrix}$.

You can verify that the following are true.

$$C_{B_2}[T(x^2)]=\begin{bmatrix}1\\1\\1\\0\end{bmatrix}, C_{B_2}[T(x)]=\begin{bmatrix}0\\-1\\-1\\0\end{bmatrix}, C_{B_2}[T(1)]=\begin{bmatrix}0\\0\\-1\\1\end{bmatrix}$$

Using these vectors as the columns of $M_{B_2B_1}(T)$ we have

$$M_{B_2B_1}(T)=\begin{bmatrix}1&1&0&0\\0&1&-1&0\\1&1&-1&-1\\1&0&0&1\end{bmatrix}$$

Below is a video on finding the matrix for the linear transformation that corresponds to integrating polynomials.



Below is a video on finding the matrix for the linear transformation that corresponds to a horizontal shift: P_2 to P_2.



The next example demonstrates that this method can be used to solve different types of problems. We will examine the above example and see if we can work backwards to determine the action of T from the matrix $M_{B_2B_1}(T)$.

Example 9.9.3: Finding the Action of a Linear Transformation

Let $T : \mathbb{P}_3 \mapsto \mathbb{R}^4$ be an isomorphism with

$$M_{B_2B_1}(T) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 1 & -1 & -1 \\ 1 & 0 & 0 & 1 \end{bmatrix},$$

where $B_1 = \left\{x^3, x^2, x, 1\right\}$ is an ordered basis of $\mathbb{P}_3$ and

$$B_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is an ordered basis of $\mathbb{R}^4$. If $p(x) = ax^3 + bx^2 + cx + d$, find $T(p(x))$.

Solution

Recall that $C_{B_2}[T(p(x))] = M_{B_2B_1}(T)C_{B_1}(p(x))$. Then we have

$$C_{B_2}[T(p(x))] = M_{B_2B_1}(T)C_{B_1}(p(x))$$
$$= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 1 & -1 & -1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$
$$= \begin{bmatrix} a+b \\ b-c \\ a+b-c-d \\ a+d \end{bmatrix}$$

Therefore

$$T(p(x)) = C_D^{-1} \begin{bmatrix} a+b \\ b-c \\ a+b-c-d \\ a+d \end{bmatrix}$$
$$= (a+b)\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + (b-c)\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + (a+b-c-d)\begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix} + (a+d)\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} a+b \\ b-c \\ c+d \\ a+d \end{bmatrix}$$

You can verify that this was the definition of $T(p(x))$ given in the previous example.

We can also find the matrix of the composite of multiple transformations.

Theorem 9.9.3: Matrix of Composition

Let V, W and U be finite dimensional vector spaces, and suppose $T : V \mapsto W$, $S : W \mapsto U$ are linear transformations. Suppose V, W and U have ordered bases of B_1, B_2 and B_3 respectively. Then the matrix of the composite transformation $S \circ T$ (or ST) is given by

$$M_{B_3B_1}(ST) = M_{B_3B_2}(S)M_{B_2B_1}(T).$$

The next important theorem gives a condition on when T is an isomorphism.

Theorem 9.9.4: Isomorphism

Let V and W be vector spaces such that both have dimension n and let $T : V \mapsto W$ be a linear transformation. Suppose B_1 is an ordered basis of V and B_2 is an ordered basis of W.

Then the conditions that $M_{B_2B_1}(T)$ is invertible for **all** B_1 and B_2, and that $M_{B_2B_1}(T)$ is invertible for **some** B_1 and B_2 are equivalent. In fact, these occur if and only if T is an isomorphism.

If T is an isomorphism, the matrix $M_{B_2B_1}(T)$ is invertible and its inverse is given by $[M_{B_2B_1}(T)]^{-1} = M_{B_1B_2}(T^{-1})$.

Consider the following example.

Example 9.9.4

Suppose $T : \mathbb{P}_3 \to \mathbb{M}_{22}$ is a linear transformation defined by

$$T(ax^3+bx^2+cx+d) = \begin{bmatrix} a+d & b-c \\ b+c & a-d \end{bmatrix}$$

for all $ax^3+bx^2+cx+d \in \mathbb{P}_3$. Let $B_1 = \{x^3, x^2, x, 1\}$ and

$$B_2 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

be ordered bases of $\mathbb{P}_3$ and $\mathbb{M}_{22}$, respectively.

1. Find $M_{B_2B_1}(T)$.
2. Verify that T is an isomorphism by proving that $M_{B_2B_1}(T)$ is invertible.
3. Find $M_{B_1B_2}(T^{-1})$, and verify that $M_{B_1B_2}(T^{-1}) = [M_{B_2B_1}(T)]^{-1}$.
4. Use $M_{B_1B_2}(T^{-1})$ to find T^{-1}.

Solution

1. $$\begin{aligned} M_{B_2B_1}(T) &= [\, C_{B_2}[T(1)] \quad C_{B_2}[T(x)] \quad C_{B_2}[T(x^2)] \quad C_{B_2}[T(x^3)]\,] \\ &= \left[C_{B_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad C_{B_2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad C_{B_2} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad C_{B_2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right] \\ &= \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix} \end{aligned}$$

2. $\det(M_{B_2B_1}(T)) = 4$, so the matrix is invertible, and hence T is an isomorphism.

3. $$T^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1, T^{-1} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = x, T^{-1} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = x^2, T^{-1} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = x^3,$$

   so

   $$T^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \frac{1+x^3}{2}, T^{-1} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \frac{x-x^2}{2},$$
   $$T^{-1} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \frac{x+x^2}{2}, T^{-1} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \frac{1-x^3}{2}.$$

   Therefore,

   $$M_{B_1B_2}(T^{-1}) = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{bmatrix}$$

   You should verify that $M_{B_2B_1}(T)M_{B_1B_2}(T^{-1}) = I_4$. From this it follows that $[M_{B_2B_1}(T)]^{-1} = M_{B_1B_2}(T^{-1})$.

4.

$$
\begin{aligned}
C_{B_1}\left(T^{-1}\begin{bmatrix}p & q\\ r & s\end{bmatrix}\right) &= M_{B_1B_2}(T^{-1})C_{B_2}\left(\begin{bmatrix}p & q\\ r & s\end{bmatrix}\right)\\
T^{-1}\begin{bmatrix}p & q\\ r & s\end{bmatrix} &= C_{B_1}^{-1}\left(M_{B_1B_2}(T^{-1})C_{B_2}\left(\begin{bmatrix}p & q\\ r & s\end{bmatrix}\right)\right)\\
&= C_{B_1}^{-1}\left(\frac{1}{2}\begin{bmatrix}1 & 0 & 0 & 1\\ 0 & 1 & 1 & 0\\ 0 & -1 & 1 & 0\\ 1 & 0 & 0 & -1\end{bmatrix}\begin{bmatrix}p\\ q\\ r\\ s\end{bmatrix}\right)\\
&= C_{B_1}^{-1}\left(\frac{1}{2}\begin{bmatrix}p+s\\ q+r\\ r-q\\ p-s\end{bmatrix}\right)\\
&= \frac{1}{2}(p+s)x^3+\frac{1}{2}(q+r)x^2+\frac{1}{2}(r-q)x+\frac{1}{2}(p-s).
\end{aligned}
$$

Below is a video on finding the matrix for the linear transformation that corresponds to multiplying $x+a$ by the derivative of the polynomial: P_2 to P_2.



This page titled 9.9: The Matrix of a Linear Transformation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **9.9: The Matrix of a Linear Transformation** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

9.E: Exercises

Exercise 9.E.1

Suppose you have $\mathbb{R}^2$ and the $+$ operation is as follows:

$$(a,b)+(c,d)=(a+d,b+c).$$

Scalar multiplication is defined in the usual way. Is this a vector space? Explain why or why not.

Exercise 9.E.2

Suppose you have $\mathbb{R}^2$ and the $+$ operation is as follows:

$$(a,b)+(c,d)=(0,b+d)$$

Scalar multiplication is defined in the usual way. Is this a vector space? Explain why or why not.

Exercise 9.E.3

Suppose you have $\mathbb{R}^2$ and scalar multiplication is defined as $c(a,b)=(a,cb)$ while vector addition is defined as usual. Is this a vector space? Explain why or why not.

Exercise 9.E.4

Suppose you have $\mathbb{R}^2$ and the $+$ operation is defined as follows.

$$(a,b)+(c,d)=(a-c,b-d)$$

Scalar multiplication is same as usual. Is this a vector space? Explain why or why not.

Exercise 9.E.5

Consider all the functions defined on a non empty set which have values in $\mathbb{R}$. Is this a vector space? Explain. The operations are defined as follows. Here f,g signify functions and a is a scalar

$$\begin{aligned}(f+g)(x) &= f(x)+g(x)\\ (af)(x) &= a(f(x))\end{aligned}$$

Exercise 9.E.6

Denote by $\mathbb{R}^{\mathbb{N}}$ the set of real valued sequences. For $\vec{a}\equiv\{a_n\}_{n=1}^{\infty}$, $\vec{b}\equiv\{b_n\}_{n=1}^{\infty}$ two of these, define their sum to be given by

$$\vec{a}+\vec{b}=\{a_n+b_n\}_{n=1}^{\infty}$$

and define scalar multiplication by

$$c\vec{a}=\{ca_n\}_{n=1}^{\infty} \text{ where } \vec{a}=\{a+n\}_{n=1}^{\infty}$$

Is this a special case of Exercise 9.E.5? Is this a vector space?

Exercise 9.E.7

Let $\mathbb{C}^2$ be the set of ordered pairs of complex numbers. Define addition and scalar multiplication in the usual way.

$$(z,w)+(\hat{z},\hat{w})=(z+\hat{z},w+\hat{w}),u(z,w)\equiv(uz,uw)$$

Here the scalars are from $\mathbb{C}$. Show this is a vector space.

Exercise 9.E.8

Let V be the set of functions defined on a nonempty set which have values in a vector space W. Is this a vector space? Explain.

Exercise 9.E.9

Consider the space of $m \times n$ matrices with operation of addition and scalar multiplication defined the usual way. That is, if A, B are two $m \times n$ matrices and c a scalar,

$$(A+B)_{ij} = A_{ij} + B_{ij},\ (cA)_{ij} \equiv c(A_{ij})$$

Exercise 9.E.10

Consider the set of $n \times n$ symmetric matrices. That is, $A = A^T$. In other words, $A_{ij} = A_{ji}$. Show that this set of symmetric matrices is a vector space and a subspace of the vector space of $n \times n$ matrices.

Exercise 9.E.11

Consider the set of all vectors in $\mathbb{R}^2, (x, y)$ such that $x + y \geq 0$. Let the vector space operations be the usual ones. Is this a vector space? Is it a subspace of $\mathbb{R}^2$?

Exercise 9.E.12

Consider the vectors in $\mathbb{R}^2, (x, y)$ such that $xy = 0$. Is this a subspace of $\mathbb{R}^2$? Is it a vector space? The addition and scalar multiplication are the usual operations.

Exercise 9.E.13

Define the operation of vector addition on $\mathbb{R}^2$ by $(x, y) + (u, v) = (x + u, y + v + 1)$. Let scalar multiplication be the usual operation. Is this a vector space with these operations? Explain.

Exercise 9.E.14

Let the vectors be real numbers. Define vector space operations in the usual way. That is $x + y$ means to add the two numbers and xy means to multiply them. Is $\mathbb{R}$ with these operations a vector space? Explain.

Exercise 9.E.15

Let the scalars be the rational numbers and let the vectors be real numbers which are the form $a + b\sqrt{2}$ for a, b rational numbers. Show that with the usual operations, this is a vector space.

Exercise 9.E.16

Let $\mathbb{P}_2$ be the set of all polynomials of degree 2 or less. That is, these are of the form $a + bx + cx^2$. Addition is defined as

$$(a + bx + cx^2) + (\hat{d} + \hat{b}x + \hat{c}x^2) = (a + \hat{a}) + (b + \hat{b})x + (c + \hat{c})x^2$$

and scalar multiplication is defined as

$$d(a + bx + cx^2) = da + dbx + cdx^2$$

Show that, with this definition of the vector space operations that $\mathbb{P}_2$ is a vector space. Now let V denote those polynomials $a + bx + cx^2$ such that $a + b + c = 0$. Is V a subspace of $\mathbb{P}_2$? Explain.

Exercise 9.E.17

Let M, N be subspaces of a vector space V and consider $M+N$ defined as the set of all $m+n$ where $m \in M$ and $n \in N$. Show that $M+N$ is a subspace of V.

Exercise 9.E.18

Let M, N be subspaces of a vector space V. Then $M \cap N$ consists of all vectors which are in both M and N. Show that $M \cap N$ is a subspace of V.

Exercise 9.E.19

Let M, N be subspaces of a vector space $\mathbb{R}^2$. Then $N \cup M$ consists of all vectors which are in either M or N. Show that $N \cup M$ is not necessarily a subspace of $\mathbb{R}^2$ by giving an example where $N \cup M$ fails to be a subspace.

Exercise 9.E.20

Let X consist of the real valued functions which are defined on an interval $[a, b]$. For $f, g \in X, f+g$ is the name of the function which satisfies $(f+g)(x) = f(x)+g(x)$. For s a real number, $(sf)(x) = s(f(x))$. Show this is a vector space.

Answer

The axioms of a vector space all hold because they hold for a vector space. The only thing left to verify is the assertions about the things which are supposed to exist. 0 would be the zero function which sends everything to 0. This is an additive identity. Now if f is a function, $-f(x) \equiv (-f(x))$. Then

$$(f+(-f))(x) \equiv f(x)+(-f)(x) \equiv f(x)+(-f(x)) = 0$$

Hence $f+-f = 0$. For each $x \in [a, b]$, let $f_x(x) = 1$ and $f_x(y) = 0$ if $y \neq x$. Then these vectors are obviously linearly independent.

Exercise 9.E.21

Consider functions defined on $\{1, 2, \cdots, n\}$ having values in $\mathbb{R}$. Explain how, if V is the set of all such functions, V can be considered as $\mathbb{R}^n$.

Answer

Let $f(i)$ be the ith component of a vector $\vec{x} \in \mathbb{R}^n$. Thus a typical element in $\mathbb{R}^n$ is $(f(1), \cdots, f(n))$.

Exercise 9.E.22

Let the vectors be polynomials of degree no more than 3. Show that with the usual definitions of scalar multiplication and addition wherein, for $p(x)$ a polynomial, $(ap)(x) = ap(x)$ and for p, q polynomials $(p+q)(x) = p(x)+q(x)$, this is a vector space.

Answer

This is just a subspace of the vector space of functions because it is closed with respect to vector addition and scalar multiplication. Hence this is a vector space.

Exercise 9.E. 23

Let V be a vector space and suppose $\{\vec{x}_1, \cdots, \vec{x}_l\}$ is a set of vectors in V. Show that $\vec{0}$ is in $span\{\vec{x}_1, \cdots, \vec{x}_k\}$.

Answer

$$\sum_{i=1}^{k} 0\vec{x}_k = \vec{0}$$

Exercise 9.E. 24

Determine if $p(x) = 4x^2 - x$ is in the span given by

$$span\{x^2 + x,\ x^2 - 1,\ -x + 2\}$$

Exercise 9.E. 25

Determine if $p(x) = -x^2 + x + 2$ is in the span given by

$$span\{x^2 + x + 1,\ 2x^2 + x\}$$

Exercise 9.E. 26

Determine if $A = \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$ is in the span given by

$$span\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\right\}$$

Exercise 9.E. 27

Show that the spanning set in Exercise 9.E. 26 is a spanning set for M_{22}, the vector space of all 2×2 matrices.

Exercise 9.E. 28

Consider the vector space of polynomials of degree at most 2, $\mathbb{P}_2$. Determine whether the following is a basis for $\mathbb{P}_2$.

$$\{x^2 + x + 1,\ 2x^2 + 2x + 1,\ x + 1\}$$

Hint: There is a isomorphism from $\mathbb{R}^3$ to $\mathbb{P}_2$. It is defined as follows:

$$T\vec{e}_1 = 1,\ T\vec{e}_2 = x,\ T\vec{e}_3 = x^2$$

Then extend T linearly. Thus

$$T\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = x^2 + x + 1,\ T\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = 2x^2 + 2x + 1,\ T\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 1 + x$$

It follows that if

$$\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$$

is a basis for $\mathbb{R}^3$, then the polynomials will be a basis for $\mathbb{P}_2$ because they will be independent. Recall that an isomorphism takes a linearly independent set to a linearly independent set. Also, since T is an isomorphism, it preserves all linear relations.

Exercise 9.E.29

Find a basis in $\mathbb{P}_2$ for the subspace

$$span\{1+x+x^2,\ 1+2x,\ 1+5x-3x^2\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others. Hint: This is the situation in which you have a spanning set and you want to cut it down to form a linearly independent set which is also a spanning set. Use the same isomorphism above. Since T is an isomorphism, it preserves all linear relations so if such can be found in $\mathbb{R}^3$, the same linear relations will be present in $\mathbb{P}_2$.

Exercise 9.E.30

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{1+x-x^2+x^3,\ 1+2x+3x^3,\ -1+3x+5x^2+7x^3,\ 1+6x+4x^2+11x^3\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.31

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{1+x-x^2+x^3,\ 1+2x+3x^3,\ -1+3x+5x^2+7x^3,\ 1+6x+4x^2+11x^3\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.32

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-2x^2+x+2,\ 3x^3-x^2+2x+2,\ 7x^3+x^2+4x+2,\ 5x^3+3x+2\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.33

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3+2x^2+x-2,\ 3x^3+3x^2+2x-2,\ 3x^3+x+2,\ 3x^3+x+2\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.34

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-5x^2+x+5,\ 3x^3-4x^2+2x+5,\ 5x^3+8x^2+2x-5,\ 11x^3+6x+5\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.35

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-3x^2+x+3,\ 3x^3-2x^2+2x+3,\ 7x^3+7x^2+3x-3,\ 7x^3+4x+3\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.36

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-x^2+x+1,\ 3x^3+2x+1,\ 4x^3+x^2+2x+1,\ 3x^3+2x-1\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.37

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-x^2+x+1,\ 3x^3+2x+1,\ 13x^3+x^2+8x+4,\ 3x^3+2x-1\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.38

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-3x^2+x+3,\ 3x^3-2x^2+2x+3,\ -5x^3+5x^2-4x-6,\ 7x^3+4x-3\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.39

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-2x^2+x+2,\ 3x^3-x^2+2x+2,\ 7x^3-x^2+4x+4,\ 5x^3+3x-2\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.40

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-2x^2+x+2,\ 3x^3-x^2+2x+2,\ 3x^3+4x^2+x-2,\ 7x^3-x^2+4x+4\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.41

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-4x^2+x+4,\ 3x^3-3x^2+2x+4,\ -3x^3+3x^2-2x-4,\ -2x^3+4x^2-2x-4\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.42

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3+2x^2+x-2,\ 3x^3+3x^2+2x-2,\ 5x^3+x^2+2x+2,\ 10x^3+10x^2+6x-6\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.43

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3+x^2+x-1,\ 3x^3+2x^2+2x-1,\ x^3+1,\ 4x^3+3x^2+2x-1\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.44

Find a basis in $\mathbb{P}_3$ for the subspace

$$span\{x^3-x^2+x+1,\ 3x^3+2x+1,\ x^3+2x^2-1,\ 4x^3+x^2+2x+1\}$$

If the above three vectors do not yield a basis, exhibit one of them as a linear combination of the others.

Exercise 9.E.45

Here are some vectors.

$$\{x^3+x^2-x-1,\ 3x^3+2x^2+2x-1\}$$

If these are linearly independent, extend to a basis for all of $\mathbb{P}_3$.

Exercise 9.E.46

Here are some vectors.

$$\{x^3-2x^2-x+2,\ 3x^3-x^2+2x+2\}$$

If these are linearly independent, extend to a basis for all of $\mathbb{P}_3$.

Exercise 9.E.47

Here are some vectors.

$$\{x^3-3x^2-x+3,\ 3x^3-2x^2+2x+3\}$$

If these are linearly independent, extend to a basis for all of $\mathbb{P}_3$.

Exercise 9.E.48

Here are some vectors.

$$\{x^3-2x^2-3x+2,\ 3x^3-x^2-6x+2,\ -8x^3+18x+10\}$$

If these are linearly independent, extend to a basis for all of $\mathbb{P}_3$.

Exercise 9.E.49

Here are some vectors.

$$\{x^3-3x^2-3x+3,\ 3x^3-2x^2-6x+3,\ -8x^3+18x+40\}$$

If these are linearly independent, extend to a basis for all of $\mathbb{P}_3$.

Exercise 9.E.50

Here are some vectors.

$$\{x^3 - x^2 + x + 1,\ 3x^3 + 2x + 1,\ 4x^3 + 2x + 2\}$$

If these are linearly independent, extend to a basis for all of $\mathbb{P}_3$.

Exercise 9.E.51

Here are some vectors.

$$\{x^3 + x^2 + 2x - 1,\ 3x^3 + 2x^2 + 4x - 1,\ 7x^3 + 8x + 23\}$$

If these are linearly independent, extend to a basis for all of $\mathbb{P}_3$.

Exercise 9.E.52

Determine if the following set is linearly independent. If it is linearly dependent, write one vector as a linear combination of the other vectors in the set.

$$\{x + 1,\ x^2 + 2,\ x^2 - x - 3\}$$

Exercise 9.E.53

Determine if the following set is linearly independent. If it is linearly dependent, write one vector as a linear combination of the other vectors in the set.

$$\{x^2 + x,\ -2x^2 - 4x - 6,\ 2x - 2\}$$

Exercise 9.E.54

Determine if the following set is linearly independent. If it is linearly dependent, write one vector as a linear combination of the other vectors in the set.

$$\left\{\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -7 & 2 \\ -2 & -3 \end{bmatrix}, \begin{bmatrix} 4 & 0 \\ 1 & 2 \end{bmatrix}\right\}$$

Exercise 9.E.55

Determine if the following set is linearly independent. If it is linearly dependent, write one vector as a linear combination of the other vectors in the set.

$$\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\right\}$$

Exercise 9.E.56

If you have 5 vectors in $\mathbb{R}^5$ and the vectors are linearly independent, can it always be concluded they span $\mathbb{R}^5$?

Answer

Yes. If not, there would exist a vector not in the span. But then you could add in this vector and obtain a linearly independent set of vectors with more vectors than a basis.

Exercise $9.E.57$

If you have 6 vectors in $\mathbb{R}^5$, is it possible they are linearly independent? Explain.

Answer

No. They can't be.

Exercise $9.E.58$

Let $\mathbb{P}_3$ be the polynomials of degree no more than 3. Determine which of the following are bases for this vector space.

a. $\{x+1,\ x^3+x^2+2x,\ x^2+x,\ x^3+x^2+x\}$
b. $\{x^3+1,\ x^2+x,\ 2x^3+x^2,\ 2x^3-x^2-3x+1\}$

Answer

a.

b. Suppose

$$c_1(x^3+1)+c_2(x^2+x)+c_3(2x^3+x^2)+c_4(2x^3-x^2-3x+1)=0$$

Then combine the terms according to power of x.

$$(c_1+2c_3+2c_4)x^3+(c_2+c_3-c_4)x^2+(c_2-3c_4)x+(c_1+c_4)=0$$

Is there a non zero solution to the system

$$\begin{aligned} c_1+2c_3+2c_4 &= 0 \\ c_2+c_3-c_4 &= 0 \\ c_2-3c_4 &= 0 \\ c_1+c_4 &= 0 \end{aligned}$$

, Solution is:

$$[c_1=0,\ c_2=0,\ c_3=0,\ c_4=0]$$

Therefore, these are linearly independent.

Exercise $9.E.59$

In the context of the above problem, consider polynomials

$$\{a_i x^3+b_i x^2+c_i x+d_i,\ i=1,2,3,4\}$$

Show that this collection of polynomials is linearly independent on an interval $[s,t]$ if and only if

$$\begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{bmatrix}$$

is an invertible matrix.

Answer

Let $p_i(x)$ denote the ith of these polynomials. Suppose $\sum_i C_i p_i(x)=0$. Then collecting terms according to the exponent of x, you need to have

$$\begin{aligned} C_1a_1 + C_2a_2 + C_3a_3 + C_4a_4 &= 0 \\ C_1b_1 + C_2b_2 + C_3b_3 + C_4b_4 &= 0 \\ C_1c_1 + C_2c_2 + C_3c_3 + C_4c_4 &= 0 \\ C_1d_1 + C_2d_2 + C_3d_3 + C_4d_4 &= 0 \end{aligned}$$

The matrix of coefficients is just the transpose of the above matrix. There exists a non trivial solution if and only if the determinant of this matrix equals 0.

? Exercise 9.E.60

Let the field of scalars be $\mathbb{Q}$, the rational numbers and let the vectors be of the form $a+b\sqrt{2}$ where a, b are rational numbers. Show that this collection of vectors is a vector space with field of scalars $\mathbb{Q}$ and give a basis for this vector space.

Answer

When you add two of these you get one and when you multiply one of these by a scalar, you get another one. A basis is $\{1, \sqrt{2}\}$. By definition, the span of these gives the collection of vectors. Are they independent? Say $a+b\sqrt{2}=0$ where a, b are rational numbers. If $a \neq 0$, then $b\sqrt{2}=-a$ which can't happen since a is rational. If $b \neq 0$, then $-a=b\sqrt{2}$ which again can't happen because on the left is a rational number and on the right is an irrational. Hence both $a, b=0$ and so this is a basis.

? Exercise 9.E.61

Suppose V is a finite dimensional vector space. Based on the exchange theorem above, it was shown that any two bases have the same number of vectors in them. Give a different proof of this fact using the earlier material in the book. Hint: Suppose $\{\vec{x}_1, \cdots, \vec{x}_n\}$ and $\{\vec{y}_1, \cdots, \vec{y}_m\}$ are two bases with $m < n$. Then define

$$\varphi : \mathbb{R}^n \mapsto V, \ \psi : \mathbb{R}^m \mapsto V$$

by

$$\varphi(\vec{a}) = \sum_{k=1}^{n} a_k \vec{x}_k, \ \psi(\vec{b}) = \sum_{j=1}^{m} b_j \vec{y}_j$$

Consider the linear transformation, $\psi^{-1} \circ \varphi$. Argue it is a one to one and onto mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. Now consider a matrix of this linear transformation and its reduced row-echelon form.

Answer

This is obvious because when you add two of these you get one and when you multiply one of these by a scalar, you get another one. A basis is $\{1, \sqrt{2}\}$. By definition, the span of these gives the collection of vectors. Are they independent? Say $a+b\sqrt{2}=0$ where a, b are rational numbers. If $a \neq 0$, then $b\sqrt{2}=-a$ which can't happen since a is rational. If $b \neq 0$, then $-a=b\sqrt{2}$ which again can't happen because on the left is a rational number and on the right is an irrational. Hence both $a, b=0$ and so this is a basis.

? Exercise 9.E.62

Let $M = \{\vec{u} = (u_1, \ u_2, \ u_3, \ u_4) \in \mathbb{R}^4 \ : \ |u_1| \leq 4\}$. Is M a subspace of $\mathbb{R}^4$?

Answer

This is not a subspace. $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ is in it, but $20 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ is not.

Exercise 9.E.63

Let $M = \{\vec{u} = (u_1,\ u_2,\ u_3,\ u_4) \in \mathbb{R}^4\ :\ \sin(u_1) = 1\}$. Is M a subspace of $\mathbb{R}^4$?

Answer

This is not a subspace.

Exercise 9.E.64

Let W be a subset of M_{22} given by

$$W = \{A | A \in M_{22}, A^T = A\}$$

In words, W is the set of all symmetric 2×2 matrices. Is W a subspace of M_{22}?

Exercise 9.E.65

Let W be a subset of M_{22} given by

$$W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \ | a, b, c, d \in \mathbb{R},\ a + b = c + d \right\}$$

Is W a subspace of M_{22}?

Exercise 9.E.66

Let W be a subset of P_3 given by

$$W = \{ax^3 + bx^2 + cx + d |\ a, b, c, d \in \mathbb{R}, d = 0\}$$

Is W a subspace of P_3?

Exercise 9.E.67

Let W be a subset of P_3 given by

$$W = \{p(x) = ax^3 + bx^2 + cx + d |\ a, b, c, d \in \mathbb{R}, p(2) = 1\}$$

Is W a subspace of P_3?

Exercise 9.E.68

Let T: $\mathbb{P}_2 \to \mathbb{R}$ be a linear transformation such that

$$T(x^2) = 1;\ T(x^2 + x) = 5;\ T(x^2 + x + 1) = -1.$$

Find $T(ax^2 + bx + c)$.

Answer

By linearity we have $T(x^2) = 1$, $T(x) = T(x^2 + x - x^2) = T(x^2 + x) - T(x^2) = 5 - 1 = 5$, and $T(1) = T(x^2 + x + 1 - (x^2 + x)) = T(x^2 + x + 1) - T(x^2 + x)) = -1 - 5 = -6$. Thus $T(ax^2 + bx + c) = aT(x^2) + bT(x) + cT(1) = a + 5b - 6c$.

Exercise $9.E.69$

Consider the following functions T: $\mathbb{R}^3 \to \mathbb{R}^2$. Explain why each of these functions T is not linear.

a. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+3z+1 \\ 2y-3x+z \end{bmatrix}$

b. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y^2+3z \\ 2y+3z+z \end{bmatrix}$

c. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \sin x+2y+3z \\ 2y+3z+z \end{bmatrix}$

d. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+3z \\ 2y+3z-\ln z \end{bmatrix}$

Exercise $9.E.70$

Suppose T is a linear transformation such that

$$T\begin{bmatrix} 1 \\ 1 \\ -7 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$
$$T\begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$
$$T\begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x}) = A\vec{x}$.

Answer

$$\begin{bmatrix} 3 & 1 & 1 \\ 3 & 2 & 3 \\ 3 & 3 & -1 \end{bmatrix}\begin{bmatrix} 6 & 2 & 1 \\ 5 & 2 & 1 \\ 6 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 29 & 9 & 5 \\ 46 & 13 & 8 \\ 27 & 11 & 5 \end{bmatrix}$$

Exercise $9.E.71$

Suppose T is a linear transformation such that

$$T\begin{bmatrix} 1 \\ 2 \\ -18 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ 5 \end{bmatrix}$$
$$T\begin{bmatrix} -1 \\ -1 \\ 15 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 5 \end{bmatrix}$$
$$T\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ -2 \end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x}) = A\vec{x}$.

Answer

$$\begin{bmatrix} 5 & 3 & 2 \\ 2 & 3 & 5 \\ 5 & 5 & -2 \end{bmatrix} \begin{bmatrix} 11 & 4 & 1 \\ 10 & 4 & 1 \\ 12 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 109 & 38 & 10 \\ 112 & 35 & 10 \\ 81 & 34 & 8 \end{bmatrix}$$

Exercise $9.E.72$

Consider the following functions $T\colon \mathbb{R}^3 \to \mathbb{R}^2$. Show that each is a linear transformation and determine for each the matrix A such that $T(\vec{x}) = A\vec{x}$.

a. $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+2y+3z \\ 2y-3x+z \end{bmatrix}$

b. $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 7x+2y+z \\ 3x-11y+2z \end{bmatrix}$

c. $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3x+2y+z \\ x+2y+6z \end{bmatrix}$

d. $T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2y-5x+z \\ x+y+z \end{bmatrix}$

Exercise $9.E.73$

Suppose

$$[A_1 \cdots A_n]^{-1}$$

exists where each $A_j \in \mathbb{R}^n$ and let vectors $\{B_1, \cdots, B_n\}$ in $\mathbb{R}^m$ be given. Show that there **always** exists a linear transformation T such that $T(A_i) = B_i$.

Exercise $9.E.74$

Let V and W be subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and let $T\colon V \to W$ be a linear transformation. Suppose that $\{T\vec{v}_1, \cdots, T\vec{v}_r\}$ is linearly independent. Show that it must be the case that $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is also linearly independent.

Answer

If $\sum_i^r a_i \vec{v}_r = 0$, then using linearity properties of T we get

$$0 = T(0) = T\left(\sum_i^r a_i \vec{v}_r\right) = \sum_i^r a_i T(\vec{v}_r).$$

Since we assume that $\{T\vec{v}_1, \cdots, T\vec{v}_r\}$ is linearly independent, we must have all $a_i = 0$, and therefore we conclude that $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is also linearly independent.

Exercise $9.E.75$

Let

$$V = span\left\{\begin{bmatrix}1\\1\\2\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\\1\end{bmatrix}\right\}$$

Let $T\vec{x} = A\vec{x}$ where A is the matrix

$$\begin{bmatrix}1 & 1 & 1 & 1\\0 & 1 & 1 & 0\\0 & 1 & 2 & 1\\1 & 1 & 1 & 2\end{bmatrix}$$

Give a basis for $Im(T)$.

? Exercise $9.E.76$

Let

$$V = span\left\{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\4\\4\\1\end{bmatrix}\right\}$$

Let $T\vec{x} = A\vec{x}$ where A is the matrix

$$\begin{bmatrix}1 & 1 & 1 & 1\\0 & 1 & 1 & 0\\0 & 1 & 2 & 1\\1 & 1 & 1 & 2\end{bmatrix}$$

Find a basis for $Im(T)$. In this case, the original vectors do not form an independent set.

Answer

Since the third vector is a linear combinations of the first two, then the image of the third vector will also be a linear combinations of the image of the first two. However the image of the first two vectors are linearly independent (check!), and hence form a basis of the image. Thus a basis for $Im(T)$ is:

$$V = span\left\{\begin{bmatrix}2\\0\\1\\3\end{bmatrix}, \begin{bmatrix}4\\2\\4\\5\end{bmatrix}\right\}$$

? Exercise $9.E.77$

If $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is linearly independent and T is a one to one linear transformation, show that $\{T\vec{v}_1, \cdots, T\vec{v}_r\}$ is also linearly independent. Give an example which shows that if T is only linear, it can happen that, although $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is linearly independent, $\{T\vec{v}_1, \cdots, T\vec{v}_r\}$ is not. In fact, show that it can happen that each of the $T\vec{v}_j$ equals 0.

? Exercise $9.E.78$

Let V and W be subspaces of $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively and let $T: V \to W$ be a linear transformation. Show that if T is onto W and if $\{\vec{v}_1, \cdots, \vec{v}_r\}$ is a basis for V, then $span\{T\vec{v}_1, \cdots, T\vec{v}_r\} = W$.

Exercise 9.E.79

Define T: $\mathbb{R}^4 \to \mathbb{R}^3$ as follows.

$$T\vec{x} = \begin{bmatrix} 3 & 2 & 1 & 8 \\ 2 & 2 & -2 & 6 \\ 1 & 1 & -1 & 3 \end{bmatrix} \vec{x}$$

Find a basis for $Im(T)$. Also find a basis for $\ker(T)$.

Exercise 9.E.80

Define T: $\mathbb{R}^4 \to \mathbb{R}^3$ as follows.

$$T\vec{x} = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \vec{x}$$

where on the right, it is just matrix multiplication of the vector $\vec{x}$ which is meant. Explain why T is an isomorphism of $\mathbb{R}^3$ to $\mathbb{R}^3$.

Exercise 9.E.81

Suppose T: $\mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation given by

$$T\vec{x} = A\vec{x}$$

where A is a 3×3 matrix. Show that T is an isomorphism if and only if A is invertible.

Exercise 9.E.82

Suppose T: $\mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation given by

$$T\vec{x} = A\vec{x}$$

where A is a $m \times n$ matrix. Show that T is never an isomorphism if $m \neq n$. In particular, show that if $m > n$, T cannot be onto and if $m < n$, then T cannot be one to one.

Exercise 9.E.83

Define T: $\mathbb{R}^2 \to \mathbb{R}^3$ as follows.

$$T\vec{x} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \vec{x}$$

where on the right, it is just matrix multiplication of the vector $\vec{x}$ which is meant. Show that T is one to one. Next let $W = Im(T)$. Show that T is an isomorphism of $\mathbb{R}^2$ and $Im(T)$.

Exercise 9.E.84

In the above problem, find a 2×3 matrix A such that the restriction of A to $Im(T)$ gives the same result as T^{-1} on $Im(T)$. Hint: You might let A be such that

$$A \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; A \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

now find another vector $\vec{v} \in \mathbb{R}^3$ such that

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \vec{v} \right\}$$

is a basis. You could pick

$$\vec{v} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

for example. Explain why this one works or one of your choice works. Then you could define $A\vec{v}$ to equal some vector in $\mathbb{R}^2$. Explain why there will be more than one such matrix A which will deliver the inverse isomorphism T^{-1} on $Im(T)$.

? Exercise $9.E.85$

Now let V equal $span \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$ and let $T: V \to W$ be a linear transformation where

$$W = span \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

and

$$T \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, T \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Explain why T is an isomorphism. Determine a matrix A which, when multiplied on the left gives the same result as T on V and a matrix B which delivers T^{-1} on W. Hint: You need to have

$$A \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Now enlarge $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ to obtain a basis for $\mathbb{R}^3$. You could add in $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ for example, and then pick another vector in $\mathbb{R}^4$ and let $A \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ equal this other vector. Then you would have

$$A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

This would involve picking for the new vector in $\mathbb{R}^4$ the vector $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$. Then you could find A. You can do something similar to find a matrix for T^{-1} denoted as B.

Exercise 9.E.86

Let $V = \mathbb{R}^3$ and let

$$W = span(S), \text{ where } S = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} \right\}$$

Find a basis of W consisting of vectors in S.

Answer

In this case $\dim(W) = 1$ and a basis for W consisting of vectors in S can be obtained by taking any (nonzero) vector from S.

Exercise 9.E.87

Let T be a linear transformation given by

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Find a basis for $\ker(T)$ and $Im(T)$.

Answer

A basis for $\ker(T)$ is $\left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ and a basis for $Im(T)$ is $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$. There are many other possibilities for the specific bases, but in this case $\dim(\ker(T)) = 1$ and $\dim(Im(T)) = 1$.

Exercise 9.E.88

Let T be a linear transformation given by

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Find a basis for $\ker(T)$ and $Im(T)$.

Answer

In this case $\ker(T) = \{0\}$ and $Im(T) = \mathbb{R}^2$ (pick any basis of $\mathbb{R}^2$).

Exercise 9.E.89

Let $V = \mathbb{R}^3$ and let

$$W = span \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} \right\}$$

Extend this basis of W to a basis of V.

Answer

There are many possible such extensions, one is (how do we know?):

$$\left\{\begin{bmatrix}1\\1\\1\end{bmatrix},\begin{bmatrix}-1\\2\\-1\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$$

Exercise 9.E.90

Let T be a linear transformation given by

$$T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}1&1&1\\1&1&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}$$

What is $\dim(\ker(T))$?

Answer

We can easily see that $\dim(Im(T))=1$, and thus $\dim(\ker(T))=3-\dim(Im(T))=3-1=2$.

Exercise 9.E.91

Consider the following functions which map $\mathbb{R}^n$ to $\mathbb{R}^n$.

a. T multiplies the jth component of $\vec{x}$ by a nonzero number b.
b. T replaces the ith component of $\vec{x}$ with b times the jth component added to the ith component.
c. T switches the ith and jth components.

Show these functions are linear transformations and describe their matrices A such that $T(\vec{x})=A\vec{x}$.

Answer

a. The matrix of T is the elementary matrix which multiplies the jth diagonal entry of the identity matrix by b.
b. The matrix of T is the elementary matrix which takes b times the jth row and adds to the ith row.
c. The matrix of T is the elementary matrix which switches the ith and the jth rows where the two components are in the ith and jth positions.

Exercise 9.E.92

You are given a linear transformation $T\colon \mathbb{R}^n\to\mathbb{R}^m$ and you know that

$$T(A_i)=B_i$$

where $\begin{bmatrix}A_1&\cdots&A_n\end{bmatrix}^{-1}$ exists. Show that the matrix of T is of the form

$$\begin{bmatrix}B_1&\cdots&B_n\end{bmatrix}\begin{bmatrix}A_1&\cdots&A_n\end{bmatrix}^{-1}$$

Answer

Suppose

$$\begin{bmatrix}\vec{c}_1^{\,T}\\\vdots\\\vec{c}_n^{\,T}\end{bmatrix}=\begin{bmatrix}\vec{a}_1&\cdots&\vec{a}_n\end{bmatrix}^{-1}$$

Thus $\vec{c}_i^{\,T}\vec{a}_j=\delta_{ij}$. Therefore

$$\begin{aligned}\begin{bmatrix}\vec{b}_1 & \cdots & \vec{b}_n\end{bmatrix}\begin{bmatrix}\vec{a}_1 & \cdots & \vec{a}_n\end{bmatrix}^{-1}\vec{a}_i &= \begin{bmatrix}\vec{b}_1 & \cdots & \vec{b}_n\end{bmatrix}\begin{bmatrix}\vec{c}_1^{\,T}\\ \vdots \\ \vec{c}_n^{\,T}\end{bmatrix}\vec{a}_i \\ &= \begin{bmatrix}\vec{b}_1 & \cdots & \vec{b}_n\end{bmatrix}\vec{e}_i \\ &= \vec{b}_i\end{aligned}$$

Thus $T\vec{a}_i = \begin{bmatrix}\vec{b}_1 & \cdots & \vec{b}_n\end{bmatrix}\begin{bmatrix}\vec{a}_1 & \cdots & \vec{a}_n\end{bmatrix}^{-1}\vec{a}_i = A\vec{a}_i$. If $\vec{x}$ is arbitrary, then since the matrix $\begin{bmatrix}\vec{a}_1 & \cdots & \vec{a}_n\end{bmatrix}$ is invertible, there exists a unique $\vec{y}$ such that $\begin{bmatrix}\vec{a}_1 & \cdots & \vec{a}_n\end{bmatrix}\vec{y} = \vec{x}$ Hence

$$T\vec{x} = T\left(\sum_{i=1}^{n} y_i\vec{a}_i\right) = \sum_{i=1}^{n} y_i T\vec{a}_i = \sum_{i=1}^{n} y_1 A\vec{a}_i = A\left(\sum_{i=1}^{n} y_i\vec{a}_i\right) = A\vec{x}$$

Exercise $9.E.93$

Suppose T is a linear transformation such that

$$T\begin{bmatrix}1\\2\\-6\end{bmatrix} = \begin{bmatrix}5\\1\\3\end{bmatrix}$$
$$T\begin{bmatrix}-1\\-1\\5\end{bmatrix} = \begin{bmatrix}1\\1\\5\end{bmatrix}$$
$$T\begin{bmatrix}0\\-1\\2\end{bmatrix} = \begin{bmatrix}5\\3\\-2\end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x}) = A\vec{x}$.

Answer

$$\begin{bmatrix}5 & 1 & 5\\1 & 1 & 3\\3 & 5 & -2\end{bmatrix}\begin{bmatrix}3 & 2 & 1\\2 & 2 & 1\\4 & 1 & 1\end{bmatrix} = \begin{bmatrix}37 & 17 & 11\\17 & 7 & 5\\11 & 14 & 6\end{bmatrix}$$

Exercise $9.E.94$

Suppose T is a linear transformation such that

$$T\begin{bmatrix}1\\1\\-8\end{bmatrix} = \begin{bmatrix}1\\3\\1\end{bmatrix}$$
$$T\begin{bmatrix}-1\\0\\6\end{bmatrix} = \begin{bmatrix}2\\4\\1\end{bmatrix}$$
$$T\begin{bmatrix}0\\-1\\3\end{bmatrix} = \begin{bmatrix}6\\1\\-1\end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x}) = A\vec{x}$.

Answer

$$\begin{bmatrix} 1 & 2 & 6 \\ 3 & 4 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 6 & 3 & 1 \\ 5 & 3 & 1 \\ 6 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 52 & 21 & 9 \\ 44 & 23 & 8 \\ 5 & 4 & 1 \end{bmatrix}$$

Exercise $9.E.95$

Suppose T is a linear transformation such that

$$T \begin{bmatrix} 1 \\ 3 \\ -7 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \\ 3 \end{bmatrix}$$
$$T \begin{bmatrix} -1 \\ -2 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -3 \end{bmatrix}$$
$$T \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ -3 \end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x}) = A\vec{x}$.

Answer

$$\begin{bmatrix} -3 & 1 & 5 \\ 1 & 3 & 3 \\ 3 & -3 & -3 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 \\ 1 & 2 & 1 \\ 4 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 15 & 1 & 3 \\ 17 & 11 & 7 \\ -9 & -3 & -3 \end{bmatrix}$$

Exercise $9.E.96$

Suppose T is a linear transformation such that

$$T \begin{bmatrix} 1 \\ 1 \\ -7 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$
$$T \begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$
$$T \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x}) = A\vec{x}$.

Answer

$$\begin{bmatrix} 3 & 1 & 1 \\ 3 & 2 & 3 \\ 3 & 3 & -1 \end{bmatrix} \begin{bmatrix} 6 & 2 & 1 \\ 5 & 2 & 1 \\ 6 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 29 & 9 & 5 \\ 46 & 13 & 8 \\ 27 & 11 & 5 \end{bmatrix}$$

Exercise $9.E.97$

Suppose T is a linear transformation such that

$$T\begin{bmatrix}1\\2\\-18\end{bmatrix}=\begin{bmatrix}5\\2\\5\end{bmatrix}$$
$$T\begin{bmatrix}-1\\-1\\15\end{bmatrix}=\begin{bmatrix}3\\3\\5\end{bmatrix}$$
$$T\begin{bmatrix}0\\-1\\4\end{bmatrix}=\begin{bmatrix}2\\5\\-2\end{bmatrix}$$

Find the matrix of T. That is find A such that $T(\vec{x})=A\vec{x}$.

Answer

$$\begin{bmatrix}5&3&2\\2&3&5\\5&5&-2\end{bmatrix}\begin{bmatrix}11&4&1\\10&4&1\\12&3&1\end{bmatrix}=\begin{bmatrix}109&38&10\\112&35&10\\81&34&8\end{bmatrix}$$

Exercise $9.E.98$

Consider the following functions $T\colon \mathbb{R}^3\to\mathbb{R}^2$. Show that each is a linear transformation and determine for each the matrix A such that $T(\vec{x})=A\vec{x}$.

a. $T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x+2y+3z\\2y-3x+z\end{bmatrix}$

b. $T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}7x+2y+z\\3x-11y+2z\end{bmatrix}$

c. $T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}3x+2y+z\\x+2y+6z\end{bmatrix}$

d. $T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}2y-5x+z\\x+y+z\end{bmatrix}$

Exercise $9.E.99$

Consider the following functions $T\colon \mathbb{R}^3\to\mathbb{R}^2$. Explain why each of these functions T is not linear.

a. $T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x+2y+3z+1\\2y-3x+z\end{bmatrix}$

b. $T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x+2y^2+3z\\2y+3x+z\end{bmatrix}$

c. $T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}\sin x+2y+3z\\2y+3x+z\end{bmatrix}$

d. $T\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}x+2y+3z\\2y+3x-\ln z\end{bmatrix}$

Exercise 9.E.100

Suppose

$$\begin{bmatrix} A_1 & \cdots & A_n \end{bmatrix}^{-1}$$

exists where each $A_j \in \mathbb{R}^n$ and let vectors $\{B_1, \cdots, B_n\}$ in $\mathbb{R}^m$ be given. Show that there always exists a linear transformation T such that $T(A_i) = B_i$.

Exercise 9.E.101

Find the matrix for $T(\vec{w}) = \text{proj}_{\vec{v}}(\vec{w})$ where $\vec{v} = \begin{bmatrix} 1 & -2 & 3 \end{bmatrix}^T$.

Answer

Recall that $\text{proj}_{\vec{u}}(\vec{v}) = \frac{\vec{v} \bullet \vec{u}}{||\vec{u}||^2} \vec{u}$ and so the desired matrix has ith column equal to $\text{proj}_{\vec{u}}(\vec{e}_i)$. Therefore, the matrix desired is

$$\frac{1}{14} \begin{bmatrix} 1 & -2 & 3 \\ -2 & 4 & -6 \\ 3 & -6 & 9 \end{bmatrix}$$

Exercise 9.E.102

Find the matrix for $T(\vec{w}) = \text{proj}_{\vec{v}}(\vec{w})$ where $\vec{v} = \begin{bmatrix} 1 & 5 & 3 \end{bmatrix}^T$.

Answer

$$\frac{1}{35} \begin{bmatrix} 1 & 5 & 3 \\ 5 & 25 & 15 \\ 3 & 15 & 9 \end{bmatrix}$$

Exercise 9.E.103

Find the matrix for $T(\vec{w}) = \text{proj}_{\vec{v}}(\vec{w})$ where $\vec{v} = \begin{bmatrix} 1 & 0 & 3 \end{bmatrix}^T$.

Answer

$$\frac{1}{10} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 0 & 0 \\ 3 & 0 & 9 \end{bmatrix}$$

Exercise 9.E.104

Let $B = \left\{ \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\}$ be a basis of $\mathbb{R}^2$ and let $\vec{x} = \begin{bmatrix} 5 \\ -7 \end{bmatrix}$ be a vector in $\mathbb{R}^2$. Find $C_B(\vec{x})$.

Exercise 9.E.105

Let $B = \left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix} \right\}$ be a basis of $\mathbb{R}^3$ and let $\vec{x} = \begin{bmatrix} 5 \\ -1 \\ 4 \end{bmatrix}$ be a vector in $\mathbb{R}^2$. Find $C_B(\vec{x})$.

Answer

$$C_B(\vec{x}) = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}.$$

? Exercise 9.E. 106

Let T: $\mathbb{R}^2 \mapsto \mathbb{R}^2$ be a linear transformation defined by $T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} a+b \\ a-b \end{bmatrix}$.

Consider the two bases

$$B_1 = \{\vec{v}_1, \vec{v}_2\} = \left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$$

and

$$B_2 = \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}$$

Find the matrix M_{B_2,B_1} of T with respect to the bases B_1 and B_2.

Answer

$$M_{B_2B_1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

9.E: Exercises is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

10: Some Prerequisite Topics

The topics presented in this section are important concepts in mathematics and therefore should be examined.

10.1: Sets and Set Notation
10.2: Well Ordering and Induction

This page titled 10: Some Prerequisite Topics is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

10.1: Sets and Set Notation

A set is a collection of things called **elements**. For example $\{1,2,3,8\}$would be a set consisting of the elements 1,2,3, and 8. To indicate that 3 is an element of $\{1,2,3,8\}$,it is customary to write $3 \in \{1,2,3,8\}$ .We can also indicate when an element is not in a set, by writing $9 \notin \{1,2,3,8\}$ which says that 9 is not an element of $\{1,2,3,8\}$. Sometimes a rule specifies a set. For example you could specify a set as all integers larger than 2. This would be written as

$$S = \{x \in \mathbb{Z} : x > 2\}.$$

This notation says: S is the set of all integers, x, such that $x > 2$.

Suppose A and B are sets with the property that every element of A is an element of B. Then we say that A is a subset of B. For example, $\{1,2,3,8\}$is a subset of $\{1,2,3,4,5,8\}$ .In symbols, we write

$$\{1,2,3,8\} \subseteq \{1,2,3,4,5,8\}.$$

It is sometimes said that "A is contained in B" or even "B contains A". The same statement about the two sets may also be written as

$$\{1,2,3,4,5,8\} \supseteq \{1,2,3,8\}.$$

We can also talk about the **union** of two sets, which we write as $A \cup B$. This is the set consisting of everything which is an element of at least one of the sets, A or B. As an example of the union of two sets, consider

$$\{1,2,3,8\} \cup \{3,4,7,8\} = \{1,2,3,4,7,8\}.$$

This set is made up of the numbers which are in at least one of the two sets.

In general

$$A \cup B = \{x : x \in A \text{ or } x \in B\}$$

Notice that an element which is in *both* A and B is also in the union, as well as elements which are in only one of A or B.

Another important set is the intersection of two sets A and B, written $A \cap B$. This set consists of everything which is in *both* of the sets. Thus $\{1,2,3,8\} \cap \{3,4,7,8\} = \{3,8\}$because 3 and 8 are those elements the two sets have in common. In general,

$$A \cap B = \{x : x \in A \text{ and } x \in B\}$$

If A and B are two sets, $A \setminus B$ denotes the set of things which are in A but not in B. Thus

$$A \setminus B = \{x \in A : x \notin B\}$$

For example, if $A = \{1,2,3,8\}$ and $B = \{3,4,7,8\}$, then $A \setminus B = \{1,2,3,8\} \setminus \{3,4,7,8\} = \{1,2\}$.

A special set which is very important in mathematics is the empty set denoted by $\emptyset$, which is defined as the set which has no elements in it. It follows that the empty set is a subset of every set. This is true because if it were not so, there would have to exist a set A, such that $\emptyset$ has something in it which is not in A. However, $\emptyset$ has nothing in it and so it must be that $\emptyset \subseteq A$.

We can also use brackets to denote sets which are intervals of numbers. Let a and b be real numbers. Then

- $[a,b] = \{x \in \mathbb{R} : a \leq x \leq b\}$
- $[a,b) = \{x \in \mathbb{R} : a \leq x < b\}$
- $(a,b) = \{x \in \mathbb{R} : a < x < b\}$
- $(a,b] = \{x \in \mathbb{R} : a < x \leq b\}$
- $[a,\infty) = \{x \in \mathbb{R} : x \geq a\}$
- $(-\infty,a] = \{x \in \mathbb{R} : x \leq a\}$

These sorts of sets of real numbers are called intervals. The two points a and b are called endpoints, or bounds, of the interval. In particular, a is the *lower bound* while b is the *upper bound* of the above intervals, where applicable. Other intervals such as $(-\infty,b)$ are defined by analogy to what was just explained. In general, the curved parenthesis, $($, indicates the end point is not included in the interval, while the square parenthesis, $[$, indicates this end point is included. The reason that there will always be a

curved parenthesis next to ∞ or $-\infty$ is that these are not real numbers and cannot be included in the interval in the way a real number can.

To illustrate the use of this notation relative to intervals consider three examples of inequalities. Their solutions will be written in the interval notation just described.

Example 10.1.1: Solving an Inequality

Solve the inequality $2x+4 \leq x-8$.

Solution

We need to find x such that $2x+4 \leq x-8$. Solving for x, we see that $x \leq -12$ is the answer. This is written in terms of an interval as $(-\infty, -12]$.

Consider the following example.

Example 10.1.2: Solving an Inequality

Solve the inequality $(x+1)(2x-3) \geq 0$.

Solution

We need to find x such that $(x+1)(2x-3) \geq 0$. The solution is given by $x \leq -1$ or $x \geq \frac{3}{2}$. Therefore, x which fit into either of these intervals gives a solution. In terms of set notation this is denoted by $(-\infty, -1] \cup [\frac{3}{2}, \infty)$.

Consider one last example.

Example 10.1.3: Solving an Inequality

Solve the inequality $x(x+2) \geq -4$.

Solution

This inequality is true for any value of x where x is a real number. We can write the solution as $\mathbb{R}$ or $(-\infty, \infty)$.

In the next section, we examine another important mathematical concept.

This page titled 10.1: Sets and Set Notation is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **10.1: Sets and Set Notation** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

10.2: Well Ordering and Induction

We begin this section with some important notation. Summation notation, written $\sum_{i=1}^{j} i$, represents a sum. Here, i is called the index of the sum, and we add iterations until $i = j$. For example,

$$\sum_{i=1}^{j} i = 1 + 2 + \cdots + j$$

Another example:

$$a_{11} + a_{12} + a_{13} = \sum_{i=1}^{3} a_{1i}$$

The following notation is a specific use of summation notation.

Notation 10.2.1: Summation Notation

Let a_{ij} be real numbers, and suppose $1 \leq i \leq r$ while $1 \leq j \leq s$. These numbers can be listed in a rectangular array as given by

$$\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1s} \\ a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rs} \end{array}$$

Then $\sum_{j=1}^{s} \sum_{i=1}^{r} a_{ij}$ means to first sum the numbers in each column (using i as the index) and then to add the sums which result (using j as the index). Similarly, $\sum_{i=1}^{r} \sum_{j=1}^{s} a_{ij}$ means to sum the vectors in each row (using j as the index) and then to add the sums which result (using i as the index).

Notice that since addition is commutative,

$$\sum_{j=1}^{s} \sum_{i=1}^{r} a_{ij} = \sum_{i=1}^{r} \sum_{j=1}^{s} a_{ij}.$$

We now consider the main concept of this section. Mathematical induction and well ordering are two extremely important principles in math. They are often used to prove significant things which would be hard to prove otherwise.

Definition 10.2.1: Well Ordered

A set is **well ordered** if every nonempty subset S, contains a smallest element z having the property that $z \leq x$ for all $x \in S$.

In particular, the set of natural numbers defined as

$$\mathbb{N} = \{1, 2, \cdots\}$$

is well ordered.

Consider the following proposition.

Proposition 10.2.1: Well Ordered Sets

Any set of integers larger than a given number is well ordered.

This proposition claims that if a set has a lower bound which is a real number, then this set is well ordered.

Further, this proposition implies the principle of mathematical induction. The symbol $\mathbb{Z}$ denotes the set of all integers. Note that if a is an integer, then there are no integers between a and $a + 1$.

Theorem 10.2.1: Mathematical Induction

A set $S \subseteq \mathbb{Z}$, having the property that $a \in S$ and $n+1 \in S$ whenever $n \in S$, contains all integers $x \in \mathbb{Z}$ such that $x \geq a$.

Proof

Let T consist of all integers larger than or equal to a which are not in S. The theorem will be proved if $T = \emptyset$. If $T \neq \emptyset$ then by the well ordering principle, there would have to exist a smallest element of T, denoted as b. It must be the case that $b > a$ since by definition, $a \notin T$. Thus $b \geq a+1$, and so $b-1 \geq a$ and $b-1 \notin S$ because if $b-1 \in S$, then $b-1+1 = b \in S$ by the assumed property of S. Therefore, $b-1 \in T$ which contradicts the choice of b as the smallest element of T. ($b-1$ is smaller.) Since a contradiction is obtained by assuming $T \neq \emptyset$, it must be the case that $T = \emptyset$ and this says that every integer at least as large as a is also in S.

Mathematical induction is a very useful device for proving theorems about the integers. The procedure is as follows.

Procedure 10.2.1: Proof by Mathematical Induction

Suppose S_n is a statement which is a function of the number n, for $n = 1, 2, \cdots$, and we wish to show that S_n is true for all $n \geq 1$. To do so using mathematical induction, use the following steps.

1. **Base Case:** Show S_1 is true.
2. Assume S_n is true for some n, which is the **induction hypothesis**. Then, using this assumption, show that S_{n+1} is true.

Proving these two steps shows that S_n is true for all $n = 1, 2, \cdots$.

We can use this procedure to solve the following examples.

Example 10.2.1: Proving by Induction

Prove by induction that $\sum_{k=1}^{n} k^2 = \dfrac{n(n+1)(2n+1)}{6}$.

Solution

By Procedure 10.2.1, we first need to show that this statement is true for $n = 1$. When $n = 1$, the statement says that

$$\begin{aligned} \sum_{k=1}^{1} k^2 &= \frac{1(1+1)(2(1)+1)}{6} \\ &= \frac{6}{6} \\ &= 1 \end{aligned}$$

The sum on the left hand side also equals 1, so this equation is true for $n = 1$.

Now suppose this formula is valid for some $n \geq 1$ where n is an integer. Hence, the following equation is true.

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \tag{10.2.1}$$

We want to show that this is true for $n+1$.

Suppose we add $(n+1)^2$ to both sides of Equation (10.2.1).

$$\begin{aligned} \sum_{k=1}^{n+1} k^2 &= \sum_{k=1}^{n} k^2 + (n+1)^2 \\ &= \frac{n(n+1)(2n+1)}{6} + (n+1)^2 \end{aligned}$$

The step going from the first to the second line is based on the assumption that the formula is true for n. Now simplify the expression in the second line,

$$\frac{n(n+1)(2n+1)}{6}+(n+1)^2$$

This equals

$$(n+1)\left(\frac{n(2n+1)}{6}+(n+1)\right)$$

and

$$\frac{n(2n+1)}{6}+(n+1)=\frac{6(n+1)+2n^2+n}{6}=\frac{(n+2)(2n+3)}{6}$$

Therefore,

$$\sum_{k=1}^{n+1}k^2=\frac{(n+1)(n+2)(2n+3)}{6}=\frac{(n+1)((n+1)+1)(2(n+1)+1)}{6}$$

showing the formula holds for $n+1$ whenever it holds for n. This proves the formula by mathematical induction. In other words, this formula is true for all $n=1,2,\cdots$.

Consider another example.

Example 10.2.2: Proving an Inequality by Induction

Show that for all $n\in\mathbb{N}$, $\frac{1}{2}\cdot\frac{3}{4}\cdots\frac{2n-1}{2n}<\frac{1}{\sqrt{2n+1}}$.

Solution

Again we will use the procedure given in Procedure 10.2.1 to prove that this statement is true for all n. Suppose $n=1$. Then the statement says

$$\frac{1}{2}<\frac{1}{\sqrt{3}}$$

which is true.

Suppose then that the inequality holds for n. In other words,

$$\frac{1}{2}\cdot\frac{3}{4}\cdots\frac{2n-1}{2n}<\frac{1}{\sqrt{2n+1}}$$

is true.

Now multiply both sides of this inequality by $\frac{2n+1}{2n+2}$. This yields

$$\frac{1}{2}\cdot\frac{3}{4}\cdots\frac{2n-1}{2n}\cdot\frac{2n+1}{2n+2}<\frac{1}{\sqrt{2n+1}}\frac{2n+1}{2n+2}=\frac{\sqrt{2n+1}}{2n+2}$$

The theorem will be proved if this last expression is less than $\frac{1}{\sqrt{2n+3}}$. This happens if and only if

$$\left(\frac{1}{\sqrt{2n+3}}\right)^2=\frac{1}{2n+3}>\frac{2n+1}{(2n+2)^2}$$

which occurs if and only if $(2n+2)^2>(2n+3)(2n+1)$ and this is clearly true which may be seen from expanding both sides. This proves the inequality.

Let's review the process just used. If S is the set of integers at least as large as 1 for which the formula holds, the first step was to show $1 \in S$ and then that whenever $n \in S$, it follows $n+1 \in S$. Therefore, by the principle of mathematical induction, S contains $[1, \infty) \cap \mathbb{Z}$, all positive integers. In doing an inductive proof of this sort, the set S is normally not mentioned. One just verifies the steps above.

This page titled 10.2: Well Ordering and Induction is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) .

- **10.2: Well Ordering and Induction** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

CHAPTER OVERVIEW

11: Introduction

11.1: Notes About These Notes
11.2: Introduction to Differential Equations
11.3: Classification of Differential Equations
11.4: First order ODEs
11.E: Introduction (Exercises)

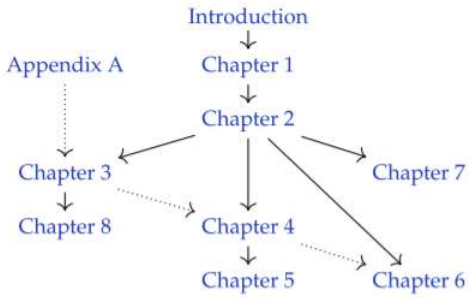
11: Introduction is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

11.1: Notes About These Notes

This book originated from my class notes for Math 286 at the University of Illinois at Urbana-Champaign (UIUC) in Fall 2008 and Spring 2009. It is a first course on differential equations for engineers. Using this book, I also taught Math 285 at UIUC, Math 20D at University of California, San Diego (UCSD), and Math 4233 at Oklahoma State University (OSU). Normally these courses are taught with Edwards and Penney, *Differential Equations and Boundary Value Problems: Computing and Modeling* [EP], or Boyce and DiPrima's *Elementary Differential Equations and Boundary Value Problems* [BD], and this book aims to be more or less a drop-in replacement. Other books I used as sources of information and inspiration are E.L. Ince's classic (and inexpensive) *Ordinary Differential Equations* [I], Stanley Farlow's *Differential Equations and Their Applications* [F], now available from Dover, Berg and McGregor's *Elementary Partial Differential Equations* [BM], and William Trench's free book *Elementary Differential Equations with Boundary Value Problems* [T]. See the Further Reading chapter at the end of the book.

11.1.1: Organization

The organization of this book to some degree requires chapters be done in order. Later chapters can be dropped. The dependence of the material covered is roughly:



There are a few references in chapters 4 and 5 to chapter 3 (some linear algebra), but these references are not essential and can be skimmed over, so chapter 3 can safely be dropped, while still covering chapters 4 and 5. Chapter 6 does not depend on chapter 4 except that the PDE section 6.5 makes a few references to chapter 4, although it could, in theory, be covered separately. The more in-depth appendix A on linear algebra can replace the short review Section 3.2 for a course that combines linear algebra and ODE.

11.1.2: Typical Types of Courses

Several typical types of courses can be run with the book. There are the two original courses at UIUC, both cover ODE as well some PDE. Either, there is the 4 hours-a-week for a semester (Math 286 at UIUC):

Introduction (0.2), chapter 1 (1.1-1.7), chapter 2, chapter 3, chapter 4 (4.1-4.9), chapter 5 (or 6 or 7 or 8).

Or, the second course at UIUC is at 3 hours-a-week (Math 285 at UIUC):

Introduction (0.2), chapter 1 (1.1-1.7), chapter 2, chapter 4 (4.1-4.9), (and maybe chapter 5, 6 or 7).

A semester-long course at 3 hours a week that doesn't cover either systems or PDE will cover, beyond the introduction, chapter 1, chapter 2, chapter 6, and chapter 7, (with sections skipped as above). On the other hand, a typical course that covers systems will probably need to skip Laplace and power series and cover chapter 1, chapter 2, chapter 3, and chapter 8.

If sections need to be skipped in the beginning, a good core of the sections on single ODE is: 0.2, 1.1-1.4, 1.6, 2.1, 2.2, 2.4-2.6.

The complete book can be covered at a reasonably fast pace at approximately 76 lectures (without appendix A) or 86 lectures (with appendix A replacing Section 3.2). This is not accounting for exams, review, or time spent in a computer lab. A two-quarter or a two-semester course can be easily run with the material. For example (with some sections perhaps strategically skipped):

Semester 1: Introduction, chapter 1, chapter 2, chapter 6, chapter 7.
Semester 2: Chapter 3, chapter 8, chapter 4, chapter 5.

A combined course on ODE with linear algebra can run as:

Introduction, chapter 1 (1.1-1.7), chapter 2, appendix A, chapter 3 (w/o Section 3.2), (possibly chapter 8).

The chapter on the Laplace transform (chapter 6), the chapter on Sturm–Liouville (chapter 5), the chapter on power series (chapter 7), and the chapter on nonlinear systems (chapter 8), are more or less interchangeable and can be treated as "topics". If chapter 8 is covered, it may be best to place it right after chapter 3, and chapter 5 is best covered right after chapter 4. If time is short, the first two sections of chapter 7 make a reasonable self-contained unit.

11.1.3: Computer Resources

The book's website https://www.jirka.org/diffyqs/ contains the following resources:

1. Interactive SAGE demos.
2. Online WeBWorK homeworks (using either your own WeBWorK installation or Edfinity) for most sections, customized for this book.
3. The PDFs of the figures used in this book.

I taught the UIUC courses using IODE (https://faculty.math.illinois.edu/iode/). IODE is a free software package that works with Matlab (proprietary) or Octave (free software). The graphs in the book were made with the Genius software (see https://www.jirka.org/genius.html). I use Genius in class to show these (and other) graphs.

The LaTeX source of the book is also available for possible modification and customization at github (https://github.com/jirilebl/diffyqs).

11.1.4: Acknowledgments

Firstly, I would like to acknowledge Rick Laugesen. I used his handwritten class notes the first time I taught Math 286. My organization of this book through chapter 5, and the choice of material covered, is heavily influenced by his notes. Many examples and computations are taken from his notes. I am also heavily indebted to Rick for all the advice he has given me, not just on teaching Math 286. For spotting errors and other suggestions, I would also like to acknowledge (in no particular order): John P. D'Angelo, Sean Raleigh, Jessica Robinson, Michael Angelini, Leonardo Gomes, Jeff Winegar, Ian Simon, Thomas Wicklund, Eliot Brenner, Sean Robinson, Jannett Susberry, Dana Al-Quadi, Cesar Alvarez, Cem Bagdatlioglu, Nathan Wong, Alison Shive, Shawn White, Wing Yip Ho, Joanne Shin, Gladys Cruz, Jonathan Gomez, Janelle Louie, Navid Froutan, Grace Victorine, Paul Pearson, Jared Teague, Ziad Adwan, Martin Weilandt, Sönmez Şahutoğlu, Pete Peterson, Thomas Gresham, Prentiss Hyde, Jai Welch, Simon Tse, Andrew Browning, James Choi, Dusty Grundmeier, John Marriott, Jim Kruidenier, Barry Conrad, Wesley Snider, Colton Koop, Sarah Morse, Erik Boczko, Asif Shakeel, Chris Peterson, Nicholas Hu, Paul Seeburger, Jonathan McCormick, David Leep, William Meisel, Shishir Agrawal, Tom Wan, Andres Valloud, and probably others I have forgotten. Finally, I would like to acknowledge NSF grants DMS-0900885 and DMS-1362337.

11.1: Notes About These Notes is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **0.1: Notes About These Notes** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

11.2: Introduction to Differential Equations

11.2.1: Differential Equations

The laws of physics are generally written down as differential equations. Therefore, all of science and engineering use differential equations to some degree. Understanding differential equations is essential to understanding almost anything you will study in your science and engineering classes. You can think of mathematics as the language of science, and differential equations are one of the most important parts of this language as far as science and engineering are concerned. As an analogy, suppose all your classes from now on were given in Swahili. It would be important to first learn Swahili, or you would have a very tough time getting a good grade in your classes.

You saw many differential equations already without perhaps knowing about it. And you even solved simple differential equations when you took calculus. Let us see an example you may not have seen:

$$\frac{dx}{dt} + x = 2\cos t. \tag{11.2.1}$$

Here x is the *dependent variable* and t is the *independent variable*. Equation (11.2.1) is a basic example of a *differential equation*. It is an example of a *first order differential equation*, since it involves only the first derivative of the dependent variable. This equation arises from Newton's law of cooling where the ambient temperature oscillates with time.

11.2.2: Solutions of Differential Equations

Solving the differential equation means finding x in terms of t. That is, we want to find a function of t, which we call x, such that when we plug x, t, and $\frac{dx}{dt}$ into (11.2.1), the equation holds; that is, the left hand side equals the right hand side. It is the same idea as it would be for a normal (algebraic) equation of just x and t. We claim that

$$x = x(t) = \cos t + \sin t$$

is a *solution*. How do we check? We simply plug x into equation (11.2.1)! First we need to compute $\frac{dx}{dt}$. We find that $\frac{dx}{dt} = -\sin t + \cos t$. Now let us compute the left-hand side of (11.2.1).

$$\frac{dx}{dt} + x = \underbrace{(-\sin t + \cos t)}_{\frac{dx}{dt}} + \underbrace{(\cos t + \sin t)}_{x} = 2\cos t.$$

Yay! We got precisely the right-hand side. But there is more! We claim $x = \cos t + \sin t + e^{-t}$ is also a solution. Let us try,

$$\frac{dx}{dt} = -\sin t + \cos t - e^{-t}.$$

We plug into the left-hand side of (11.2.1)

$$\frac{dx}{dt} + x = \underbrace{(-\sin t + \cos t - e^{-t})}_{\frac{dx}{dt}} + \underbrace{(\cos t + \sin t + e^{-t})}_{x} = 2\cos t.$$

And it works yet again!

So there can be many different solutions. For this equation all solutions can be written in the form

$$x = \cos t + \sin t + Ce^{-t},$$

for some constant C. Different constants C will give different solutions, so there are really infinitely many possible solutions. See Figure 11.2.1 for the graph of a few of these solutions. We will see how we find these solutions a few lectures from now.

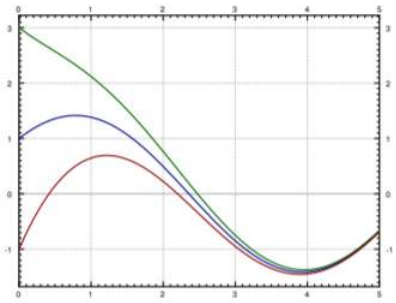


Figure 11.2.1: Few solutions of $\frac{dx}{dt} + x = 2\cos t$.

Solving differential equations can be quite hard. There is no general method that solves every differential equation. We will generally focus on how to get exact formulas for solutions of certain differential equations, but we will also spend a little bit of time on getting approximate solutions. And we will spend some time on understanding the equations without solving them.

Most of this book is dedicated to *ordinary differential equations* or ODEs, that is, equations with only one independent variable, where derivatives are only with respect to this one variable. If there are several independent variables, we get *partial differential equations* or PDEs.

Even for ODEs, which are very well understood, it is not a simple question of turning a crank to get answers. When you can find exact solutions, they are usually preferable to approximate solutions. It is important to understand how such solutions are found. Although in real applications you will leave much of the actual calculations to computers, you need to understand what they are doing. It is often necessary to simplify or transform your equations into something that a computer can understand and solve. You may even need to make certain assumptions and changes in your model to achieve this.

To be a successful engineer or scientist, you will be required to solve problems in your job that you never saw before. It is important to learn problem solving techniques, so that you may apply those techniques to new problems. A common mistake is to expect to learn some prescription for solving all the problems you will encounter in your later career. This course is no exception.

Below is a video on verifying a solution to a differential equation.



11.2.3: Differential Equations in Practice

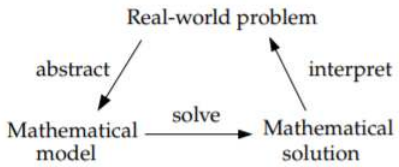


Figure 11.2.2

So how do we use differential equations in science and engineering? First, we have some *real-world problem* we wish to understand. We make some simplifying assumptions and create a *mathematical model*. That is, we translate the real-world situation into a set of differential equations. Then we apply mathematics to get some sort of a *mathematical solution*. There is still something

left to do. We have to interpret the results. We have to figure out what the mathematical solution says about the real-world problem we started with.

Learning how to formulate the mathematical model and how to interpret the results is what your physics and engineering classes do. In this course, we will focus mostly on the mathematical analysis. Sometimes we will work with simple real-world examples so that we have some intuition and motivation about what we are doing.

Let us look at an example of this process. One of the most basic differential equations is the standard *exponential growth model*. Let P denote the population of some bacteria on a Petri dish. We assume that there is enough food and enough space. Then the rate of growth of bacteria is proportional to the population—a large population grows quicker. Let t denote time (say in seconds) and P the population. Our model is

$$\frac{dP}{dt} = kP,$$

for some positive constant $k > 0$.

Example 11.2.1

Suppose there are 100 bacteria at time 0 and 200 bacteria 10 seconds later. How many bacteria will there be 1 minute from time 0 (in 60 seconds)?

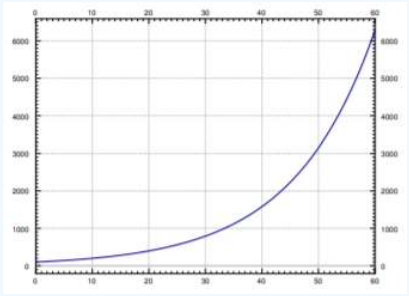


Figure 11.2.2: Bacteria growth in the first 60 seconds.

Solution

First we need to solve the equation. We claim that a solution is given by

$$P(t) = Ce^{kt},$$

where C is a constant. Let us try:

$$\frac{dP}{dt} = Cke^{kt} = kP.$$

And it really is a solution.

OK, now what? We do not know C, and we do not know k. But we know something. We know $P(0) = 100$, and we know $P(10) = 200$. Let us plug these conditions in and see what happens.

$$\begin{aligned} 100 &= P(0) = Ce^{k0} = C, \\ 200 &= P(10) = 100\, e^{k10}. \end{aligned} \tag{11.2.2}$$

Therefore, $2 = e^{10k}$ or $\frac{\ln 2}{10} = k \approx 0.069$. So

$$P(t) = 100\, e^{(\ln 2)t/10} \approx 100\, e^{0.069t}.$$

At one minute, $t = 60$, the population is $P(60) = 6400$. See Figure 11.2.2

Let us talk about the interpretation of the results. Does our solution mean that there must be exactly 6400 bacteria on the plate at 60s? No! We made assumptions that might not be true exactly, just approximately. If our assumptions are reasonable, then there will be approximately 6400 bacteria. Also, in real life P is a discrete quantity, not a real number. However, our model has no problem saying that for example at 61 seconds, $P(61) \approx 6859.35$

Normally, the k in $P' = kP$ is known, and we want to solve the equation for different *initial conditions*. What does that mean? Take $k = 1$ for simplicity. Suppose we want to solve the equation $\frac{dP}{dt} = P$ subject to $P(0) = 1000$ (the initial condition). Then the solution turns out to be (exercise)

$$P(t) = 1000\, e^t.$$

We call $P(t) = Ce^t$ *the general solution*, as every solution of the equation can be written in this form for some constant C. We need an initial condition to find out what C is, in order to find the *particular solution* we are looking for. Generally, when we say "particular solution," we just mean some solution.

Below is a video on verifying a solution to a differential equation and finding a particular solution.



11.2.4: Fundamental Equations

A few equations appear often and it is useful to just memorize what their solutions are. Let us call them the four fundamental equations. Their solutions are reasonably easy to guess by recalling properties of exponentials, sines, and cosines. They are also simple to check, which is something that you should always do. No need to wonder if you remembered the solution correctly.

First such equation is

$$\frac{dy}{dx} = ky,$$

for some constant $k > 0$. Here y is the dependent and x the independent variable. The general solution for this equation is

$$y(x) = Ce^{kx}.$$

We saw above that this function is a solution, although we used different variable names.

Next,

$$\frac{dy}{dx} = -ky,$$

for some constant $k > 0$. The general solution for this equation is

$$y(x) = Ce^{-kx}.$$

? Exercise 11.2.1

Check that the y given is really a solution to the equation.

Next, take the *second order differential equation*

$$\frac{d^2y}{dx^2} = -k^2y,$$

for some constant $k > 0$. The general solution for this equation is

$$y(x) = C_1 \cos(kx) + C_2 \sin(kx).$$

Since the equation is a second order differential equation, we have two constants in our general solution.

Exercise 11.2.2

Check that the y given is really a solution to the equation.

Finally, consider the second order differential equation

$$\frac{d^2y}{dx^2} = k^2 y,$$

for some constant $k > 0$. The general solution for this equation is

$$y(x) = C_1 e^{kx} + C_2 e^{-kx},$$

or

$$y(x) = D_1 \cosh(kx) + D_2 \sinh(kx).$$

For those that do not know, cosh and sinh are defined by

$$\cosh x = \frac{e^x + e^{-x}}{2}, \qquad \sinh x = \frac{e^x - e^{-x}}{2}.$$

They are called the *hyperbolic cosine* and *hyperbolic sine*. These functions are sometimes easier to work with than exponentials. They have some nice familiar properties such as $\cosh 0 = 1$, $\sinh 0 = 0$, and $\frac{d}{dx}\cosh x = \sinh x$ (no that is not a typo) and $\frac{d}{dx}\sinh x = \cosh x$.

Exercise 11.2.3

Check that both forms of the y given are really solutions to the equation.

Example 11.2.2

In equations of higher order, you get more constants you must solve for to get a particular solution. The equation $\frac{d^2y}{dx^2} = 0$ has the general solution $y = C_1 x + C_2$; simply integrate twice and don't forget about the constant of integration. Consider the initial conditions $y(0) = 2$ and $y'(0) = 3$. We plug in our general solution and solve for the constants:

$$2 = y(0) = C_1 \cdot 0 + C_2 = C_2, \qquad 3 = y'(0) = C_1.$$

In other words, $y = 3x + 2$ is the particular solution we seek.

An interesting note about cosh: The graph of cosh is the exact shape of a hanging chain. This shape is called a *catenary*. Contrary to popular belief this is not a parabola. If you invert the graph of cosh, it is also the ideal arch for supporting its weight. For example, the gateway arch in Saint Louis is an inverted graph of cosh—if it were just a parabola it might fall. The formula used in the design is inscribed inside the arch:

$$y = -127.7 \text{ ft} \cdot \cosh(x/127.7 \text{ ft}) + 757.7 \text{ ft}.$$

11.2: Introduction to Differential Equations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

11.3: Classification of Differential Equations

There are many types of differential equations, and we classify them into different categories based on their properties. Let us quickly go over the most basic classification. We already saw the distinction between ordinary and partial differential equations:

- *Ordinary differential equations* or (ODE) are equations where the derivatives are taken with respect to only one variable. That is, there is only one independent variable.
- *Partial differential equations* or (PDE) are equations that depend on partial derivatives of several variables. That is, there are several independent variables.

Let us see some examples of ordinary differential equations:

$$\begin{aligned} &\frac{dy}{dt} = ky, && \text{(Exponential growth)} \\ &\frac{dy}{dt} = k(A-y), && \text{(Newton's law of cooling)} \\ &m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx = f(t). && \text{(Mechanical vibrations)} \end{aligned} \tag{11.3.1}$$

And of partial differential equations:

$$\begin{aligned} &\frac{\partial y}{\partial t} + c\frac{\partial y}{\partial x} = 0, && \text{(Transport equation)} \\ &\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, && \text{(Heat equation)} \\ &\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}. && \text{(Wave equation in 2 dimensions)} \end{aligned} \tag{11.3.2}$$

If there are several equations working together, we have a so-called *system of differential equations*. For example,

$$y' = x, \qquad x' = y$$

is a simple system of ordinary differential equations. Maxwell's equations for electromagnetics,

$$\begin{aligned} &\nabla \cdot \vec{D} = \rho, && \nabla \cdot \vec{B} = 0, \\ &\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}, && \nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}, \end{aligned} \tag{11.3.3}$$

are a system of partial differential equations. The divergence operator $\nabla\cdot$ and the curl operator $\nabla\times$ can be written out in partial derivatives of the functions involved in the x, y, and z variables.

The next bit of information is the *order* of the equation (or system). The order is simply the order of the largest derivative that appears. If the highest derivative that appears is the first derivative, the equation is of first order. If the highest derivative that appears is the second derivative, then the equation is of second order. For example, Newton's law of cooling above is a first order equation, while the mechanical vibrations equation is a second order equation. The equation governing transversal vibrations in a beam,

$$a^4\frac{\partial^4 y}{\partial x^4} + \frac{\partial^2 y}{\partial t^2} = 0,$$

is a fourth order partial differential equation. It is fourth order as at least one derivative is the fourth derivative. It does not matter that the derivative in t is only of second order.

In the first chapter, we will start attacking first order ordinary differential equations, that is, equations of the form $\frac{dy}{dx} = f(x,y)$. In general, lower order equations are easier to work with and have simpler behavior, which is why we start with them.

We also distinguish how the dependent variables appear in the equation (or system). In particular, we say an equation is *linear* if the dependent variable (or variables) and their derivatives appear linearly, that is only as first powers, they are not multiplied together, and no other functions of the dependent variables appear. In other words, the equation is a sum of terms, where each term is some function of the independent variables or some function of the independent variables multiplied by a dependent variable or its

derivative. Otherwise, the equation is called *nonlinear*. For example, an ordinary differential equation is linear if it can be put into the form

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = b(x). \tag{11.3.4}$$

The functions a_0, a_1, …, a_n are called the *coefficients*. The equation is allowed to depend arbitrarily on the independent variable. So

$$e^x \frac{d^2 y}{dx^2} + \sin(x)\frac{dy}{dx} + x^2 y = \frac{1}{x} \tag{11.3.5}$$

is still a linear equation as y and its derivatives only appear linearly.

All the equations and systems above as examples are linear. It may not be immediately obvious for Maxwell's equations unless you write out the divergence and curl in terms of partial derivatives. Let us see some nonlinear equations. For example ,

$$\frac{\partial y}{\partial t} + y\frac{\partial y}{\partial x} = \nu \frac{\partial^2 y}{\partial x^2},$$

is a nonlinear second order partial differential equation. It is nonlinear because y and $\frac{\partial y}{\partial x}$ are multiplied together. The equation

$$\frac{dx}{dt} = x^2 \tag{11.3.6}$$

is a nonlinear first order differential equation as there is a second power of the dependent variable x.

A linear equation may further be called *homogenous* if all terms depend on the dependent variable. That is, if no term is a function of the independent variables alone. Otherwise, the equation is called *homogeneous* or *inhomogeneous*. For example, the exponential growth equation, the wave equation, or the transport equation above are homogeneous. The mechanical vibrations equation above is nonhomogeneous as long as $f(t)$ is not the zero function. Similarly, if the ambient temperature A is nonzero, Newton's law of cooling is nonhomogeneous. A homogeneous linear ODE can be put into the form

$$a_n(x)\frac{d^n y}{dx^n} + a_{n-1}(x)\frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1(x)\frac{dy}{dx} + a_0(x)y = 0.$$

Compare to (11.3.4) and notice there is no function $b(x)$.

If the coefficients of a linear equation are actually constant functions, then the equation is said to have *constant coefficients*. The coefficients are the functions multiplying the dependent variable(s) or one of its derivatives, not the function $b(x)$ standing alone. A constant coefficient nonhomogeneous ODE is an equation of the form

$$a_n\frac{d^n y}{dx^n} + a_{n-1}\frac{d^{n-1} y}{dx^{n-1}} + \cdots + a_1\frac{dy}{dx} + a_0 y = b(x),$$

where $a_0, a_1, \ldots, a_n$ are all constants, but b may depend on the independent variable x. The mechanical vibrations equation above is a constant coefficient nonhomogeneous second order ODE. The same nomenclature applies to PDEs, so the transport equation, heat equation and wave equation are all examples of constant coefficient linear PDEs.

Finally, an equation (or system) is called *autonomous* if the equation does not depend on the independent variable. For autonomous ordinary differential equations, the independent variable is then thought of as time. Autonomous equation means an equation that does not change with time. For example, Newton's law of cooling is autonomous, so is equation (11.3.6). On the other hand, mechanical vibrations or (11.3.5) are not autonomous.

Below is a video on defining and classifying differential equations.



11.3: Classification of Differential Equations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

11.4: First order ODEs

This page titled 11.4: First order ODEs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

11.E: Introduction (Exercises)

11.E.1: title="0.2: Introduction to Differential Equations" href="/Bookshelves/Differential_Equations/Book:_Differential_Equations_for_Engineers_(Lebl)/0:_Introduction/0.2:_Introduction_to_Differential_Equations">Introduction to Differential Equations

Exercise 11.E.0.2.1

Show that $x = e^{4t}$ is a solution to $x''' - 12x'' + 48x' - 64x = 0$.

Exercise 11.E.0.2.2

Show that $x = e^{t}$ is not a solution to $x''' - 12x'' + 48x' - 64x = 0$.

Exercise 11.E.0.2.3

Is $y = \sin t$ a solution to $\left(\frac{dy}{dt}\right)^2 = 1 - y^2$? Justify.

Exercise 11.E.0.2.4

Let $y'' + 2y' - 8y = 0$. Now try a solution of the form $y = e^{rx}$ for some (unknown) constant r. Is this a solution for some r? If so, find all such r.

Exercise 11.E.0.2.5

Verify that $x = Ce^{-2t}$ is a solution to $x' = -2x$. Find C to solve for the initial condition $x(0) = 100$.

Exercise 11.E.0.2.6

Verify that $x = C_1e^{-t} + C_2e^{2t}$ is a solution to $x'' - x' - 2x = 0$. Find C_1 and C_2 to solve for the initial conditions $x(0) = 10$ and $x'(0) = 0$.

Exercise 11.E.0.2.7

Find a solution to $(x')^2 + x^2 = 4$ using your knowledge of derivatives of functions that you know from basic calculus.

Exercise 11.E.0.2.8

Solve:

a. $\dfrac{dA}{dt} = -10A, \quad A(0) = 5$

b. $\dfrac{dH}{dx} = 3H, \quad H(0) = 1$

c. $\dfrac{d^2y}{dx^2} = 4y, \quad y(0) = 0, \quad y'(0) = 1$

d. $\dfrac{d^2x}{dy^2} = -9x, \quad x(0) = 1, \quad x'(0) = 0$

Exercise 11.E.0.2.9

Is there a solution to $y' = y$, such that $y(0) = y(1)$?

Exercise 11.E.0.2.10

The population of city X was 100 thousand 20 years ago, and the population of city X was 120 thousand 10 years ago. Assuming constant growth, you can use the exponential population model (like for the bacteria). What do you estimate the population is now?

Exercise 11.E.0.2.11

Suppose that a football coach gets a salary of one million dollars now, and a raise of 10% every year (so exponential model, like population of bacteria). Let s be the salary in millions of dollars, and t is time in years.

a. What is $s(0)$ and $s(1)$.
b. Approximately how many years will it take for the salary to be 10 million.
c. Approximately how many years will it take for the salary to be 20 million.
d. Approximately how many years will it take for the salary to be 30 million.

Exercise 11.E.0.2.12

Show that $x = e^{-2t}$ is a solution to $x'' + 4x' + 4x = 0$.

Answer

Compute $x' = -2e^{-2t}$ and $x'' = 4e^{-2t}$. Then $(4e^{-2t}) + 4(-2e^{-2t}) + 4(e^{-2t}) = 0$.

Exercise 11.E.0.2.13

Is $y = x^2$ a solution to $x^2y'' - 2y = 0$? Justify.

Answer

Yes.

Exercise 11.E.0.2.14

Let $xy'' - y' = 0$. Try a solution of the form $y = x^r$. Is this a solution for some r? If so, find all such r.

Answer

$y = x^r$ is a solution for $r = 0$ and $r = 2$.

Exercise 11.E.0.2.15

Verify that $x = C_1e^t + C_2$ is a solution to $x'' - x' = 0$. Find C_1 and C_2 so that x satisfies $x(0) = 10$ and $x'(0) = 100$.

Answer

$C_1 = 100$, $C_2 = -90$

Exercise 11.E.0.2.16

Solve $\frac{d\varphi}{ds} = 8\varphi$ and $\varphi(0) = -9$.

Answer

$\varphi = -9e^{8s}$

Exercise 11.E.0.2.17

Solve:

a. $\frac{dx}{dt} = -4x, \quad x(0) = 9$
b. $\frac{d^2x}{dt^2} = -4x, \quad x(0) = 1, \quad x'(0) = 2$
c. $\frac{dp}{dq} = 3p, \quad p(0) = 4$
d. $\frac{d^2T}{dx^2} = 4T, \quad T(0) = 0, \quad T'(0) = 6$

Answer

a. $x = 9e^{-4t}$
b. $x = \cos(2t) + \sin(2t)$
c. $p = 4e^{3q}$
d. $T = 3\sinh(2x)$

11.E.2: title="0.3: Classification of Differential Equations" href="/Bookshelves/Differential_Equations/Book:_Differential_Equations_for_Engineers_(Lebl)/0:_Introduction/0.3:_Classification_of_Differential_Equations">Classification of Differential Equations

Exercise 11.E.0.3.1

Classify the following equations. Are they ODE or PDE? Is it an equation or a system? What is the order? Is it linear or nonlinear, and if it is linear, is it homogeneous, constant coefficient? If it is an ODE, is it autonomous?

a. $\sin(t)\frac{d^2x}{dt^2} + \cos(t)x = t^2$
b. $\frac{\partial u}{\partial x} + 3\frac{\partial u}{\partial y} = xy$
c. $y'' + 3y + 5x = 0, \quad x'' + x - y = 0$
d. $\frac{\partial^2 u}{\partial t^2} + u\frac{\partial^2 u}{\partial s^2} = 0$
e. $x'' + tx^2 = t$
f. $\frac{d^4x}{dt^4} = 0$

Exercise 11.E.0.3.2

If $\vec{u} = (u_1, u_2, u_3)$ is a vector, we have the divergence $\nabla \cdot \vec{u} = \frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y} + \frac{\partial u_3}{\partial z}$ and curl $\nabla \times \vec{u} = \left(\frac{\partial u_3}{\partial y} - \frac{\partial u_2}{\partial z}, \frac{\partial u_1}{\partial z} - \frac{\partial u_3}{\partial x}, \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y} \right)$. Notice that curl of a vector is still a vector. Write out Maxwell's equations in terms of partial derivatives and classify the system.

Exercise 11.E.0.3.3

Suppose F is a linear function, that is, $F(x, y) = ax + by$ for constants a and b. What is the classification of equations of the form $F(y', y) = 0$.

Exercise 11.E.0.3.4

Write down an explicit example of a third order, linear, nonconstant coefficient, nonautonomous, nonhomogeneous system of two ODE such that every derivative that could appear, does appear.

Exercise 11.E.0.3.5

Classify the following equations. Are they ODE or PDE? Is it an equation or a system? What is the order? Is it linear or nonlinear, and if it is linear, is it homogeneous, constant coefficient? If it is an ODE, is it autonomous?

a. $\dfrac{\partial^2 v}{\partial x^2} + 3\dfrac{\partial^2 v}{\partial y^2} = \sin(x)$
b. $\dfrac{dx}{dt} + \cos(t)x = t^2 + t + 1$
c. $\dfrac{d^7 F}{dx^7} = 3F(x)$
d. $y'' + 8y' = 1$
e. $x'' + tyx' = 0, \quad y'' + txy = 0$
f. $\dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial s^2} + u^2$

Answer

a. PDE, equation, second order, linear, nonhomogeneous, constant coefficient.
b. ODE, equation, first order, linear, nonhomogeneous, not constant coefficient, not autonomous.
c. ODE, equation, seventh order, linear, homogeneous, constant coefficient, autonomous.
d. ODE, equation, second order, linear, nonhomogeneous, constant coefficient, autonomous.
e. ODE, system, second order, nonlinear.
f. PDE, equation, second order, nonlinear.

Exercise 11.E.0.3.6

Write down the general *zero*th order linear ordinary differential equation. Write down the general solution.

Answer

equation: $a(x)y = b(x)$, solution: $y = \frac{b(x)}{a(x)}$.

Exercise 11.E.0.3.7

For which k is $\frac{dx}{dt} + x^k = t^{k+2}$ linear. Hint: there are two answers.

Answer

$k = 0$ or $k = 1$

11.E: Introduction (Exercises) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **0.E: Introduction (Exercises)** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

CHAPTER OVERVIEW

12: First order ODEs

12.1: Integrals as solutions
12.2: Slope fields
12.3: Separable Equations
12.4: Linear equations and the integrating factor
12.5: Existence and Uniqueness of Solutions of Nonlinear Equations
12.5E: Existence and Uniqueness of Solutions of Nonlinear Equations (Exercises)
12.6: Substitution
12.7: Autonomous equations
12.8: Numerical methods- Euler’s method
12.9: Exact Equations
12.10: First Order Linear PDE
12.E: First order ODEs (Exercises)

This page titled 12: First order ODEs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

12.1: Integrals as solutions

A first order ODE is an equation of the form

$$\frac{dy}{dx} = f(x, y)$$

or just

$$y' = f(x, y)$$

In general, there is no simple formula or procedure one can follow to find solutions. In the next few lectures we will look at special cases where solutions are not difficult to obtain. In this section, let us assume that f is a function of x alone, that is, the equation is

$$y' = f(x) \tag{12.1.1}$$

We could just integrate (antidifferentiate) both sides with respect to x.

$$\int y'(x)dx = \int f(x)dx + C$$

that is

$$y(x) = \int f(x)dx + C$$

This $y(x)$ is actually the general solution. So to solve Equation 12.1.1, we find some antiderivative of $f(x)$ and then we add an arbitrary constant to get the general solution.

Now is a good time to discuss a point about calculus notation and terminology. Calculus textbooks muddy the waters by talking about the integral as primarily the so-called **indefinite integral**. The indefinite integral is really the **antiderivative** (in fact the whole one-parameter family of antiderivatives). There really exists only one integral and that is the definite integral. The only reason for the indefinite integral notation is that we can always write an antiderivative as a (definite) integral. That is, by the fundamental theorem of calculus we can always write $\int f(x)dx + C$ as

$$\int_{x_0}^{x} f(t)dt + C$$

Hence the terminology to integrate when we may really mean to antidifferentiate. Integration is just one way to compute the antiderivative (and it is a way that always works, see the following examples). Integration is defined as the area under the graph, it only happens to also compute antiderivatives. For sake of consistency, we will keep using the indefinite integral notation when we want an antiderivative, and you should always think of the definite integral.

Example 12.1.1

Find the general solution of $y' = 3x^2$.

Solution

Elementary calculus tells us that the general solution must be $y = x^3 + C$. Let us check: $y' = 3x^2$. We have gotten precisely our equation back.

Normally, we also have an initial condition such as $y(x_0) = y_0$ for some two numbersx_0 and y_0 x_0 is usually 0, but not always). We can then write the solution as a definite integral in a nice way. Suppose our problem is $y' = f(x)$, $y(x_0) = y_0$. Then the solution is

$$y(x) = \int_{x_0}^{x} f(s)ds + y_0 \tag{12.1.2}$$

Let us check! We compute $y' = f(x)$, via the fundamental theorem of calculus, and by Jupiter, y is a solution. Is it the one satisfying the initial condition? Well, $y(x_0) = \int_{x_0}^{x_0} f(x)dx + y_0 = y_0$. It is!

Do note that the definite integral and the indefinite integral (antidifferentiation) are completely different beasts. The definite integral always evaluates to a number. Therefore, Equation 12.1.2 is a formula we can plug into the calculator or a computer, and it will be happy to calculate specific values for us. We will easily be able to plot the solution and work with it just like with any other function. It is not so crucial to always find a closed form for the antiderivative.

Below is a video on using integration to solve a differential equation.



Example 12.1.2

Solve

$$y' = e^{-x^2}, \;\; y(0) = 1.$$

By the preceding discussion, the solution must be

$$y(x) = \int_0^x e^{-s^2}\, ds + 1.$$

Solution

Here is a good way to make fun of your friends taking second semester calculus. Tell them to find the closed form solution. Ha ha ha (bad math joke). It is not possible (in closed form). There is absolutely nothing wrong with writing the solution as a definite integral. This particular integral is in fact very important in statistics.

Using this method, we can also solve equations of the form

$$y' = f(y)$$

Let us write the equation in Leibniz notation.

$$\frac{dy}{dx} = f(y)$$

Now we use the inverse function theorem from calculus to switch the roles of x and y to obtain

$$\frac{dy}{dx} = \frac{1}{f(y)}$$

What we are doing seems like algebra with dx and dy. It is tempting to just do algebra with dx and dy as if they were numbers. And in this case it does work. Be careful, however, as this sort of hand-waving calculation can lead to trouble, especially when more than one independent variable is involved. At this point we can simply integrate,

$$x(y) = \int \frac{1}{f(y)} dy + C$$

Finally, we try to solve for y.

Example 12.1.3

Previously, we guessed $y' = ky$ (for some $k > 0$) has the solution $y = Ce^{kx}$. We can now find the solution without guessing. First we note that $y = 0$ is a solution. Henceforth, we assume $y \neq 0$. We write

$$\frac{dx}{dy} = \frac{1}{ky}$$

We integrate to obtain

$$x(y) = x = \frac{1}{k}\ln|y| + D$$

where D is an arbitrary constant. Now we solve for y (actually for $|y|$).

$$|y| = e^{kx-kD} = e^{-kD}e^{kx}$$

If we replace e^{-kD} with an arbitrary constant C we can get rid of the absolute value bars (which we can do as D was arbitrary). In this way, we also incorporate the solution $y = 0$. We get the same general solution as we guessed before, $y = Ce^{kx}$.

Example 12.1.4

Find the general solution of $y' = y^2$.

Solution

First we note that $y = 0$ is a solution. We can now assume that $y \neq 0$. Write

$$\frac{dx}{dy} = \frac{1}{y^2}$$

We integrate to get

$$x = \frac{-1}{y} + C$$

We solve for $y = \dfrac{1}{C - x}$. So the general solution is

$$y = \frac{1}{C-x} \text{ or } y = 0$$

Note the singularities of the solution. If for example $C = 1$, then the solution as we approach $x = 1$. See Figure 12.1.1. Generally, it is hard to tell from just looking at the equation itself how the solution is going to behave. The equation $y' = y^2$ is very nice and defined everywhere, but the solution is only defined on some interval $(-\infty, C)$ or (C, ∞). Usually when this happens we only consider one of these the solution. For example if we impose a condition $y(0) = 1$, then the solution is $y = \frac{1}{1-x}$, and we would consider this solution only for x on the interval $(-\infty, 1)$. In the figure, it is the left side of the graph.

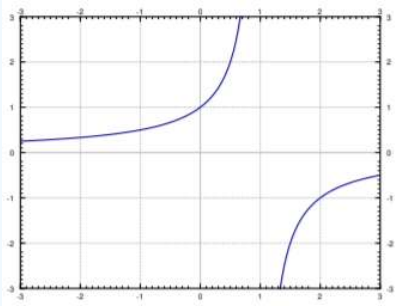


Figure 12.1.1: Plot of $y = \frac{1}{1-x}$.

Below is a video on using integration to solve an inital value problem.



Classical problems leading to differential equations solvable by integration are problems dealing with velocity, acceleration and distance. You have surely seen these problems before in your calculus class.

Example 12.1.5

Suppose a car drives at a speed $e^{t/2}$ meters per second, where t is time in seconds. How far did the car get in 2 seconds (starting at $t=0$)? How far in 10 seconds?

Solution

Let x denote the distance the car traveled. The equation is

$$x' = e^{t/2}$$

We can just integrate this equation to get that

$$x(t) = 2e^{t/2} + C$$

We still need to figure out C. We know that when $t=0$, then $x=0$. That is, $x(0)=0$. So

$$0 = x(0) = 2e^{0/2} + C = 2 + C$$

Thus $C=-2$ and

$$x(t) = 2e^{t/2} - 2$$

Now we just plug in to get where the car is at 2 and at 10 seconds. We obtain

$$x(2) = 2e^{2/2} - 2 \approx 3.44\text{~meters}, \quad x(10) = 2e^{10/2} - 2 \approx 294\text{~meters}$$

Example 12.1.6

Suppose that the car accelerates at a rate of $t^2 \frac{m}{s^2}$. At time $t=0$ the car is at the 1 meter mark and is traveling at 10 $^{\text{m}}/_{\text{s}}$. Where is the car at time $t=10$.

Solution

Well this is actually a second order problem. If x is the distance traveled, then x' is the velocity, and x'' is the acceleration. The equation with initial conditions is

$$x'' = t^2, \quad x(0) = 1, \quad x'(0) = 10$$

What if we say $x' = v$. Then we have the problem

$$v' = t^2, \quad v(0) = 10$$

Once we solve for v, we can integrate and find x.

This page titled 12.1: Integrals as solutions is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **1.1: Integrals as solutions** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

12.2: Slope fields

The general first order equation we are studying looks like

$$y' = f(x, y)$$

In general, we cannot simply solve these kinds of equations explicitly. It would be nice if we could at least figure out the shape and behavior of the solutions, or if we could find approximate solutions. At this point it may be good to first try the Lab I and/or Project I from the IODE website.

12.2.1: Slope fields

The equation $y' = f(x, y)$ gives you a slope at each point in the (x, y)-plane. And this is the slope a solution $y(x)$ would have at x if its value was y. In other words, $f(x, y)$ is the slope of a solution whose graph runs through the point (x, y). At a point (x, y), we plot a short line with the slope $f(x, y)$. For example, if $f(x, y) = xy$, then at point $(2, 1.5)$ we draw a short line of slope $xy = 2 \times 1.5 = 3$. So, if $y(x)$ is a solution and $y(2) = 1.5$, then the equation mandates that $y'(2) = 3$. See Figure 12.2.1.

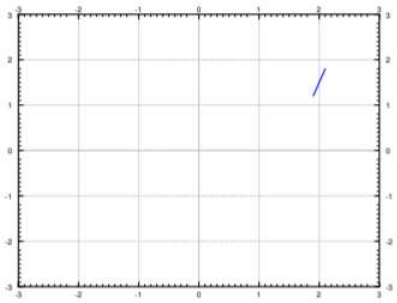


Figure 12.2.1: The slope $y' = xy$ at $(2, 1.5)$.

To get an idea of how solutions behave, we draw such lines at lots of points in the plane, not just the point $(2, 1.5)$. We would ideally want to see the slope at every point, but that is just not possible. Usually we pick a grid of points fine enough so that it shows the behavior, but not too fine so that we can still recognize the individual lines. We call this picture the of the equation. See Figure 12.2.2 for the slope field of the equation $y' = xy$. Usually in practice, one does not do this by hand, but has a computer do the drawing.

Below is a video on slope fields.



Suppose we are given a specific initial condition $y(x_0) = y_0$. A solution, that is, the graph of the solution, would be a curve that follows the slopes we drew. For a few sample solutions, see Figure 12.2.3. It is easy to roughly sketch (or at least imagine) possible solutions in the slope field, just from looking at the slope field itself. You simply sketch a line that roughly fits the little line segments and goes through your initial condition.

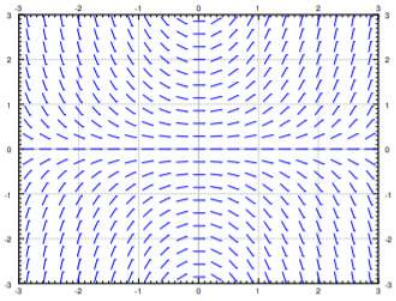


Figure 12.2.3: Slope field of $y' = xy$.

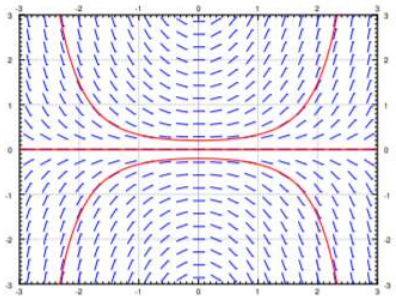


Figure 12.2.3: Slope field of $y' = xy$ with a graph of solutions satisfying $y(0) = 0.2$, $y(0) = 0$, and $y(0) = -0.2$.

By looking at the slope field we get a lot of information about the behavior of solutions without having to solve the equation. For example, in Figure 12.2.3 we see what the solutions do when the initial conditions are $y(0) > 0$, $y(0) = 0$ and $y(0) < 0$. A small change in the initial condition causes quite different behavior. We see this behavior just from the slope field and imagining what solutions ought to do.

We see a different behavior for the equation $y' = -y$. The slope field and a few solutions is in see Figure 12.2.4 If we think of moving from left to right (perhaps x is time and time is usually increasing), then we see that no matter what $y(0)$ is, all solutions tend to zero as x tends to infinity. Again that behavior is clear from simply looking at the slope field itself.

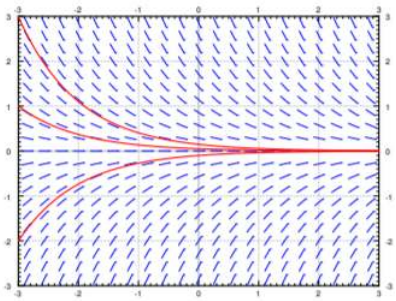


Figure 12.2.4: Slope field of $y' = -y$ with a graph of a few solutions.

Below is a video on choosing which differential equation corresponds to the given slope field.



12.2.2: Existence and uniqueness

We wish to ask two fundamental questions about the problem

$$y' = f(x, y), \quad y(x_0) = y_0$$

i. Does a solution exist?
ii. Is the solution unique (if it exists)?

What do you think is the answer? The answer seems to be yes to both does it not? Well, pretty much. But there are cases when the answer to either question can be no.

Since generally the equations we encounter in applications come from real life situations, it seems logical that a solution always exists. It also has to be unique if we believe our universe is deterministic. If the solution does not exist, or if it is not unique, we have probably not devised the correct model. Hence, it is good to know when things go wrong and why.

Example 12.2.1

Attempt to solve

$$y' = \frac{1}{x}, \quad y(0) = 0.$$

Solution

Integrate to find the general solution $y = \ln|x| + C$. Note that the solution does not exist at $x = 0$. See Figure 12.2.5 on the following page. The equation may have been written as the seemingly harmless $xy' = 1$.

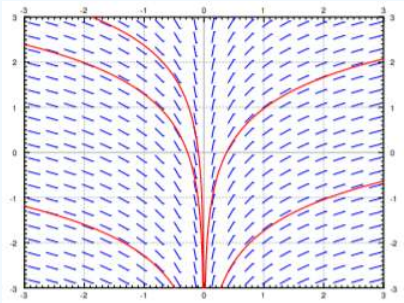


Figure 12.2.5: Slope field of $y' = \frac{1}{x}$.

Example 12.2.2

Solve:

$$y' = 2\sqrt{|y|}, \quad y(0) = 0.$$

Solution

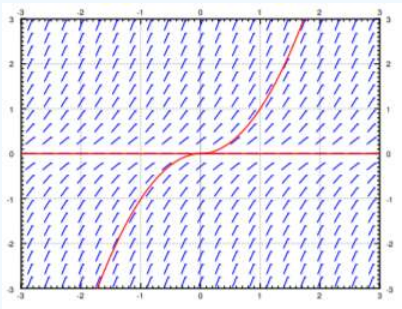


Figure 12.2.6: Slope field of $y' = 2\sqrt{|y|}$ with two solutions satisfying $y(0) = 0$.

Note that $y = 0$ is a solution. But another solution is the function

$$y(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ -x^2 & \text{if } x < 0 \end{cases}$$

It is hard to tell by staring at the slope field that the solution is not unique. Is there any hope? Of course there is. We have the following theorem, known as **Picard's theorem**.[1]

Theorem 12.2.1

Picard's theorem on existence and uniqueness

If $f(x, y)$ is *continuous* (*as a function of two variables*) *and* $\dfrac{\partial f}{\partial y}$ *exists and is continuous near some* (x_0, y_0), then a solution to

$$y' = f(x, y), \quad y(x_0) = y_0$$

exists (*at least for* x *in some small interval*) *and is unique*.

Note that the problems $y' = \dfrac{1}{x}$, $y(0) = 0$ and $y' = 2\sqrt{|y|}$, $y(0) = 0$ do not satisfy the hypothesis of the theorem. Even if we can use the theorem, we ought to be careful about this existence business. It is quite possible that the solution only exists for a short while.

Example 12.2.3

For some constant A, solve:

$$y' = y^2 \quad y(0) = A$$

.

Solution

We know how to solve this equation. First assume that $A \neq 0$, so y is not equal to zero at least for some x near 0. So $x' = \dfrac{1}{y^2}$, so $x = -\dfrac{1}{y} + C$, so $y = \dfrac{1}{C - x}$. If $y(0) = A$, then $C = \dfrac{1}{A}$ so

$$y = \frac{1}{\frac{1}{A} - x}$$

> If $A = 0$, then $y = 0$ is a solution.

For example, when $A = 1$ the solution "blows up" at $x = 1$. Hence, the solution does not exist for all x even if the equation is nice everywhere. The equation $y' = y^2$ certainly looks nice.

For most of this course we will be interested in equations where existence and uniqueness holds, and in fact holds "globally" unlike for the equation $y' = y^2$.

12.2.3: Footnotes

[1] Named after the French mathematician Charles Émile Picard (1856 – 1941)

This page titled 12.2: Slope fields is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **1.2: Slope fields** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

12.3: Separable Equations

When a differential equation is of the form $y' = f(x)$, we can just integrate: $y = \int f(x)dx + C$. Unfortunately this method no longer works for the general form of the equation $y' = f(x, y)$. Integrating both sides yields

$$y = \int f(x, y)dx + C$$

Notice the dependence on y in the integral.

12.3.0.1: Separable equations

Let us suppose that the equation is *separable*. That is, let us consider

$$y' = f(x)g(y),$$

for some functions $f(x)$ and $g(y)$. Let us write the equation in the Leibniz notation

$$\frac{dy}{dx} = f(x)g(y)$$

Then we rewrite the equation as

$$\frac{dy}{g(y)} = f(x)dx$$

Now both sides look like something we can integrate. We obtain

$$\int \frac{dy}{g(y)} = \int f(x)dx + C$$

If we can find closed form expressions for these two integrals, we can, perhaps, solve for y.

Example 12.3.1

Take the equation

$$y' = xy$$

First note that $y = 0$ is a solution, so assume $y \neq 0$ from now on, so that we can divide by y. Write the equation as $\frac{dy}{dx} = xy$, then

$$\int \frac{dy}{y} = \int x dx + C.$$

We compute the antiderivatives to get

$$\ln|y| = \frac{x^2}{2} + C$$

Or

$$|y| = e^{\frac{x^2}{2}+C} = e^{\frac{x^2}{2}}e^C = De^{\frac{x^2}{2}}$$

where $D > 0$ is some constant. Because $y = 0$ is a solution and because of the absolute value we actually can write:

$$y = De^{\frac{x^2}{2}}$$

for any number D (including zero or negative).

We check:

$$y' = Dxe^{\frac{x^2}{2}} = x\left(De^{\frac{x^2}{2}}\right) = xy$$

Yay!

We should be a little bit more careful with this method. You may be worried that we were integrating in two different variables. We seemingly did a different operation to each side. Let us work through this method more rigorously. Take

$$\frac{dy}{dx} = f(x)g(y)$$

We rewrite the equation as follows. Note that $y = y(x)$ is a function of x and so is $\frac{dy}{dx}$!

$$\frac{1}{g(y)}\frac{dy}{dx} = f(x)$$

We integrate both sides with respect to x.

$$\int \frac{1}{g(y)}\frac{dy}{dx}dx = \int f(x)dx + C$$

We use the change of variables formula (substitution) on the left hand side:

$$\int \frac{1}{g(y)}dy = \int f(x)dx + C$$

And we are done.

12.3.1: Implicit solutions

It is clear that we might sometimes get stuck even if we can do the integration. For example, take the separable equation

$$y' = \frac{xy}{y^2+1}$$

We separate variables,

$$\frac{y^2+1}{y}dy = \left(y + \frac{1}{y}\right)dy = xdx$$

We integrate to get

$$\frac{y^2}{2} + ln\,|y| = \frac{x^2}{2} + C$$

or perhaps the easier looking expression (where $D = 2C$)

$$y^2 + 2ln\,|y| = x^2 + D$$

It is not easy to find the solution explicitly as it is hard to solve for y. We, therefore, leave the solution in this form and call it an *implicit solution*. It is still easy to check that an implicit solution satisfies the differential equation. In this case, we differentiate with respect to x, and remember that y is a function of x, to get

$$y'\left(2y + \frac{2}{y}\right) = 2x$$

Multiply both sides by y and divide by $2(y^2+1)$ and you will get exactly the differential equation. We leave this computation to the reader.

If you have an implicit solution, and you want to compute values for y, you might have to be tricky. You might get multiple solutions y for each x, so you have to pick one. Sometimes you can graph x as a function of y, and then flip your paper. Sometimes you have to do more.

Computers are also good at some of these tricks. More advanced mathematical software usually has some way of plotting solutions to implicit equations. For example, for $C = 0$ if you plot all the points (x, y) that are solutions to $y^2 + 2\ln|y| = x^2$, you find the two curves in Figure 12.3.1. This is not quite a graph of a function. For each x there are two choices of y. To find a function you would have to pick one of these two curves. You pick the one that satisfies your initial condition if you have one. For example, the top curve satisfies the condition $y(1) = 1$. So for each C we really got two solutions. As you can see, computing values from an implicit solution can be somewhat tricky. But sometimes, an implicit solution is the best we can do.

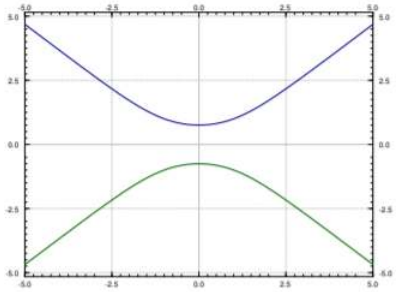


Figure 12.3.1: The implicit solution $y^2 + 2\ln|y| = x^2$ to $y' = \frac{xy}{y^2+1}$.

The equation above also has the solution $y = 0$. So the general solution is

$$y^2 + 2\ln|y| = x^2 + C, \qquad \text{and} \qquad y = 0.$$

These outlying solutions such as $y = 0$ are sometimes called *singular solutions*.

Below is a video on solving a separable differential equation.



Example 12.3.2

Solve $x^2y' = 1 - x^2 + y^2 - x^2y^2$, $y(1) = 0$.

Solution

First factor the right hand side to obtain

$$x^2y' = \left(1 - x^2\right)\left(1 + y^2\right)$$

Separate variables, integrate, and solve for y

$$\begin{aligned} \frac{y'}{1+y^2} &= \frac{1-x^2}{x^2}, \\ \frac{y'}{1+y^2} &= \frac{1}{x^2}-1, \\ \arctan(y) &= -\frac{1}{x^2}-x+C, \\ y &= \tan\left(-\frac{1}{x}-x+C\right) \end{aligned} \tag{12.3.1}$$

Solve for the initial condition, $0=\tan(-2+C)$ to get $C=2$ (or $C=2+\pi$, or $C=2+2\pi$, etc.). The particular solution we seek is, therefore,

$$y=\tan\left(\frac{-1}{x}-x+2\right).$$

Example 12.3.3

Juan made a cup of coffee, and Juan likes to drink coffee only once reaches 60 degrees Celsius and will not burn him. Initially at time $t=0$ minutes, Juan measured the temperature and the coffee was 89 degrees Celsius. One minute later, Juan measured the coffee again and it had 85 degrees. The temperature of the room (the ambient temperature) is 22 degrees. When should Juan start drinking?

Solution

Let T be the temperature of the coffee in degrees Celsius, and let A be the ambient (room) temperature, also in degrees Celsius. states that the rate at which the temperature of the coffee is changing is proportional to the difference between the ambient temperature and the temperature of the coffee. That is,

$$\frac{dT}{dt}=k(A-T),$$

for some constant k. For our setup $A=22$, $T(0)=89$, $T(1)=85$. We separate variables and integrate (let C and D denote arbitrary constants)

$$\begin{aligned} \frac{1}{T-A}\frac{dT}{dt} &= -k, \\ \ln(T-A) &= -kt+C, \quad (\text{note that} T-A>0) \\ T-A &= De^{-kt}, \\ T &= A+De^{-kt} \end{aligned} \tag{12.3.2}$$

That is, $T=22+D\,e^{-kt}$. We plug in the first condition: $89=T(0)=22+D$, and hence $D=67$. So $T=22+67\,e^{-kt}$. The second condition says $85=T(1)=22+67\,e^{-k}$. Solving for k we get $k=-\ln\frac{85-22}{67}\approx 0.0616$. Now we solve for the time t that gives us a temperature of 60 degrees. Namely, we solve

$$60=22+67e^{-0.0616t}$$

to get $t=-\frac{\ln\frac{60-22}{67}}{0.0616}\approx 9.21$ minutes. So Juan can begin to drink the coffee at just over 9 minutes from the time Juan made it. That is probably about the amount of time it took us to calculate how long it would take. See Figure 12.3.2

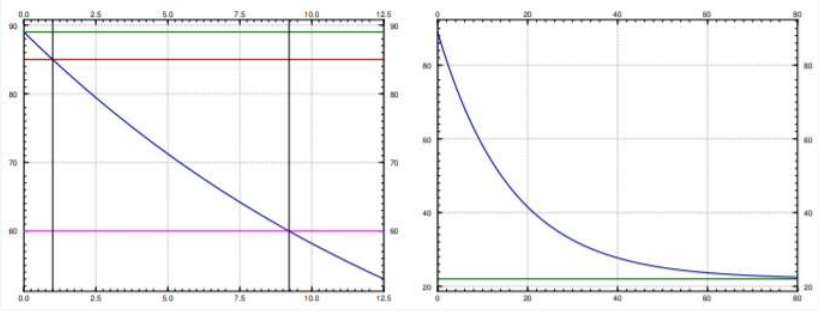


Figure 12.3.2: Graphs of the coffee temperature function $T(t)$. On the left, horizontal lines are drawn at temperatures 60, 85, and 89. Vertical lines are drawn at $t = 1$ and $t = 9.21$. Notice that the temperature of the coffee hits 85 at $t = 1$, and 60 at $t \approx 9.21$. On the right, the graph is over a longer period of time, with a horizontal line at the ambient temperature 22.

Example 12.3.4

Find the general solution to $y' = \frac{-xy^2}{3}$ (including singular solutions).

Solution

First note that $y = 0$ is a solution (a singular solution). Now assume that $y \neq 0$.

$$\begin{aligned} -\frac{3}{y^2}y' &= x, \\ \frac{3}{y} &= \frac{x^2}{2} + C, \\ y &= \frac{3}{\frac{x^2}{2} + C} = \frac{6}{x^2 + 2C}. \end{aligned} \tag{12.3.3}$$

So the general solution is,

$$y = \frac{6}{x^2 + 2C}, \quad \text{and} \quad y = 0.$$

This page titled 12.3: Separable Equations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **1.3: Separable Equations** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

12.4: Linear equations and the integrating factor

One of the most important types of equations we will learn how to solve are the so-called linear equations. In fact, the majority of the course is about linear equations. In this lecture we focus on the first order linear equation. A first order equation is linear if we can put it into the form:

$$y' + p(x)y = f(x). \tag{12.4.1}$$

Here the word "linear" means linear in y and y'; no higher powers nor functions of y or y' appear. The dependence on x can be more complicated.

Solutions of linear equations have nice properties. For example, the solution exists wherever $p(x)$ and $f(x)$ are defined, and has the same regularity (read: it is just as nice). But most importantly for us right now, there is a method for solving linear first order equations. The trick is to rewrite the left hand side of (12.4.1) as a derivative of a product of y with another function. To this end we find a function $r(x)$ such that

$$r(x)y' + r(x)p(x)y = \frac{d}{dx}[r(x)y]$$

This is the left hand side of (12.4.1) multiplied by $r(x)$. So if we multiply (12.4.1) by $r(x)$, we obtain

$$\frac{d}{dx}[r(x)y] = r(x)f(x)$$

Now we integrate both sides. The right hand side does not depend on y and the left hand side is written as a derivative of a function. Afterwards, we solve for y. The function $r(x)$ is called the integrating factor and the method is called the integrating factor method.

We are looking for a function $r(x)$, such that if we differentiate it, we get the same function back multiplied by $p(x)$. That seems like a job for the exponential function! Let

$$r(x) = e^{\int p(x)dx}$$

We compute:

$$\begin{aligned} y' + p(x)y &= f(x), \\ e^{\int p(x)dx}y' + e^{\int p(x)dx}p(x)y &= e^{\int p(x)dx}f(x), \\ \frac{d}{dx}\left[e^{\int p(x)dx}y\right] &= e^{\int p(x)dx}f(x), \\ e^{\int p(x)dx}y &= \int e^{\int p(x)dx}f(x)dx + C, \\ y &= e^{-\int p(x)dx}\left(\int e^{\int p(x)dx}f(x)dx + C\right). \end{aligned} \tag{12.4.2}$$

Of course, to get a closed form formula for y, we need to be able to find a closed form formula for the integrals appearing above.

Below is a video on solving a differential equation using an integrating factor.



Example 12.4.1

Solve

$$y' + 2xy = e^{x-x^2}, \quad y(0) = -1$$

Solution

First note that $p(x) = 2x$ and $f(x) = e^{x-x^2}$. The integrating factor is $r(x) = e^{\int p(x)dx} = e^{x^2}$. We multiply both sides of the equation by $r(x)$ to get

$$\begin{aligned} e^{x^2}y' + 2xe^{x^2}y &= e^{x-x^2}e^{x^2}, \\ \frac{d}{dx}\left[e^{x^2}y\right] &= e^x. \end{aligned} \tag{12.4.3}$$

We integrate

$$\begin{aligned} e^{x^2}y &= e^x + C, \\ y &= e^{x-x^2} + Ce^{-x^2}. \end{aligned} \tag{12.4.4}$$

Next, we solve for the initial condition $-1 = y(0) = 1 + C$, so $C = -2$. The solution is

$$y = e^{x-x^2} - 2e^{-x^2}.$$

Note that we do not care which antiderivative we take when computing $e^{\int p(x)dx}$. You can always add a constant of integration, but those constants will not matter in the end.

Exercise 12.4.1

Try it! Add a constant of integration to the integral in the integrating factor and show that the solution you get in the end is the same as what we got above. An advice: Do not try to remember the formula itself, that is way too hard. It is easier to remember the process and repeat it.

Since we cannot always evaluate the integrals in closed form, it is useful to know how to write the solution in definite integral form. A definite integral is something that you can plug into a computer or a calculator. Suppose we are given

$$y' + p(x)y = f(x), \quad y(x_0) = y_0$$

. Look at the solution and write the integrals as definite integrals.

$$y(x) = e^{\int_{x_0}^{x} -p(s)ds}\left(\int_{x_0}^{x} e^{\int_{x_0}^{t} p(s)ds} f(t)dt + y_0\right) \tag{12.4.5}$$

You should be careful to properly use dummy variables here. If you now plug such a formula into a computer or a calculator, it will be happy to give you numerical answers.

Exercise 12.4.2

Check that $y(x_0) = y_0$ in formula (12.4.5).

Exercise 12.4.3

Write the solution of the following problem as a definite integral, but try to simplify as far as you can. You will not be able to find the solution in closed form.

$$y' + y = e^{x^2 - x}, \quad y(0) = 10$$

Note

Before we move on, we should note some interesting properties of linear equations. First, for the linear initial value problem $y' + p(x)y = f(x)$, $y(x_0) = y_0$, there is always an explicit formula (12.4.5) for the solution. Second, it follows from the formula (12.4.5) that if $p(x)$ and $f(x)$ are continuous on some interval (a, b), then the solution $y(x)$ exists and is differentiable on (a, b). Compare with the simple nonlinear example we have seen previously, $y' = y^2$, and compare to Theorem 1.2.1.

Example 12.4.2

Let us discuss a common simple application of linear equations. This type of problem is used often in real life. For example, linear equations are used in figuring out the concentration of chemicals in bodies of water (rivers and lakes).



Figure 12.4.1

A 100 liter tank contains 10 kilograms of salt dissolved in 60 liters of water. Solution of water and salt (brine) with concentration of 0.1 kilograms per liter is flowing in at the rate of 5 liters a minute. The solution in the tank is well stirred and flows out at a rate of 3 liters a minute. How much salt is in the tank when the tank is full?

Solution

Let us come up with the equation. Let x denote the kilograms of salt in the tank, let t denote the time in minutes. For a small change Δt in time, the change in x (denoted Δx) is approximately

$$\Delta x \approx (\text{rate in x concentration in})\Delta t - (\text{rate out x concentration out})\Delta t.$$

Dividing through by Δt and taking the limit $\Delta t \to 0$ we see that

$$\frac{dx}{dt} = (\text{rate in x concentration in}) - (\text{rate out x concentration out})$$

In our example, we have

$$\begin{aligned} \text{rate in} &= 5, \\ \text{concentration in} &= 0.1, \\ \text{rate out} &= 3, \\ \text{concentration out} &= \frac{x}{\text{volume}} = \frac{x}{60+(5-3)t}. \end{aligned} \tag{12.4.6}$$

Our equation is, therefore,

$$\frac{dx}{dt} = (5 \text{ x } 0.1) - \left(3\frac{x}{60+2t}\right)$$

Or in the form (12.4.1)

$$\frac{dx}{dt} + \frac{3}{60+2t}x = 0.5$$

Let us solve. The integrating factor is

$$r(t) = \exp\left(\int \frac{3}{60+2t}dt\right) = \exp\left(\frac{3}{2}\ln(60+2t)\right) = (60+2t)^{3/2}$$

We multiply both sides of the equation to get

$$\begin{aligned} (60+2t)^{3/2}\frac{dx}{dt} + (60+2t)^{3/2}\frac{3}{60+2t}x &= 0.5(60+2t)^{3/2}, \\ \frac{d}{dt}\left[(60+2t)^{3/2}x\right] &= 0.5(60+2t)^{3/2}, \\ (60+2t)^{3/2}x &= \int 0.5(60+2t)^{3/2}dt + C, \\ x &= (60+2t)^{-3/2}\int \frac{(60+2t)^{3/2}}{2}dt + C(60+2t)^{-3/2}, \\ x &= (60+2t)^{-3/2}\frac{1}{10}(60+2t)^{5/2} + C(60+2t)^{-3/2}, \\ x &= \frac{(60+2t)}{10} + C(60+2t)^{-3/2}. \end{aligned} \tag{12.4.7}$$

We need to find C. We know that at $t = 0$, $x = 10$. So

$$10 = x(0) = \frac{60}{10} + C(60)^{-3/2} = 6 + C(60)^{-3/2}$$

or

$$C = 4(60^{3/2}) \approx 1859.03.$$

We are interested in x when the tank is full. So we note that the tank is full when $60+2t = 100$, or when $t = 20$. So

$$\begin{aligned} x(20) &= \frac{60+40}{10} + C(60+40)^{-3/2} \\ &\approx 10 + 1859.03(100)^{-3/2} \approx 11.86. \end{aligned} \tag{12.4.8}$$

See Figure 12.4.2 for the graph of x over t.

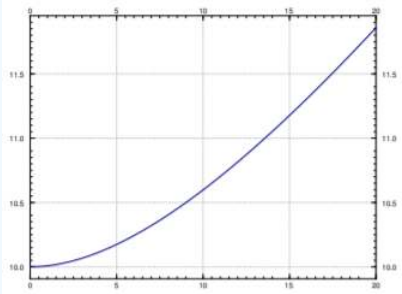


Figure 12.4.2: Graph of the solution x kilograms of salt in the tank at time t.

The concentration at the end is approximately $0.1186\,{}^{\text{kg}}/_{\text{liter}}$ and we started with $\frac{1}{6}$ or $0.167\,{}^{\text{kg}}/_{\text{liter}}$.

Playlist of Videos on Integrating Factor

This page titled 12.4: Linear equations and the integrating factor is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **1.4: Linear equations and the integrating factor** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

12.5: Existence and Uniqueness of Solutions of Nonlinear Equations

Although there are methods for solving some nonlinear equations, it is impossible to find useful formulas for the solutions of most. Whether we are looking for exact solutions or numerical approximations, it is useful to know conditions that imply the existence and uniqueness of solutions of initial value problems for nonlinear equations. In this section we state such a condition and illustrate it with examples.

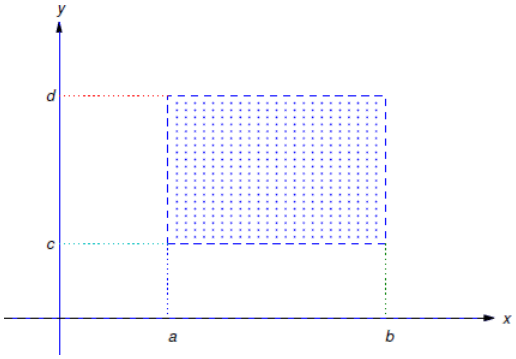


Figure 2.5.1 : An open rectangle

Some terminology: an *open rectangle* R is a set of points (x, y) such that

$$a < x < b \quad \text{and} \quad c < y < d$$

(Figure 2.5.1). We'll denote this set by $R : \{a < x < b, c < y < d\}$. "Open" means that the boundary rectangle (indicated by the dashed lines in Figure 2.5.1) is not included in R.

The next theorem gives sufficient conditions for existence and uniqueness of solutions of initial value problems for first order nonlinear differential equations. We omit the proof, which is beyond the scope of this book.

Theorem 2.5.1 : existence and uniqueness

a. If f is continuous on an open rectangle

$$R : \{a < x < b, c < y < d\}$$

that contains (x_0, y_0) then the initial value problem

$$y' = f(x, y), \quad y(x_0) = y_0 \tag{12.5.1}$$

has at least one solution on some open subinterval of (a, b) that contains x_0.

b. If both f and f_y are continuous on R then Equation 12.5.1 has a unique solution on some open subinterval of (a, b) that contains x_0

It's important to understand exactly what Theorem 2.5.1 says.

- (a) is an *existence theorem*. It guarantees that a solution exists on some open interval that contains x_0, but provides no information on how to find the solution, or to determine the open interval on which it exists. Moreover, (a) provides no information on the number of solutions that Equation 12.5.1 may have. It leaves open the possibility that Equation 12.5.1 may have two or more solutions that differ for values of x arbitrarily close to x_0. We will see in Example 2.5.6 that this can happen.
- (b) is a *uniqueness theorem*. It guarantees that Equation 12.5.1 has a unique solution on some open interval (a,b) that contains x_0. However, if $(a, b) \neq (-\infty, \infty)$, Equation 12.5.1 may have more than one solution on a larger interval that contains (a, b). For example, it may happen that $b < \infty$ and all solutions have the same values on (a, b), but two solutions y_1 and y_2 are defined on some interval (a, b_1) with $b_1 > b$, and have different values for $b < x < b_1$; thus, the graphs of the y_1 and y_2 "branch off" in different directions at $x = b$. (See Example 2.5.7 and Figure 2.5.3). In this case, continuity implies that $y_1(b) = y_2(b)$ (call their common value y), and y_1 and y_2 are both solutions of the initial value problem

$$y = f(x, y), \quad y(b) = \overline{y} \tag{12.5.2}$$

that differ on every open interval that contains b. Therefore f or f_y must have a discontinuity at some point in each open rectangle that contains (b, y), since if this were not so, 12.5.2 would have a unique solution on some open interval that contains b. We leave it to you to give a similar analysis of the case where $a > -\infty$.

Example 2.5.1

Consider the initial value problem

$$y' = \frac{x^2 - y^2}{1 + x^2 + y^2}, \quad y(x_0) = y_0. \tag{12.5.3}$$

Since

$$f(x, y) = \frac{x^2 - y^2}{1 + x^2 + y^2} \quad \text{and} \quad f_y(x, y) = -\frac{2y(1 + 2x^2)}{(1 + x^2 + y^2)^2}$$

are continuous for all (x, y), Theorem 2.5.1 implies that if (x_0, y_0) is arbitrary, then Equation 12.5.3 has a unique solution on some open interval that contains x_0.

Example 2.5.2

Consider the initial value problem

$$y' = \frac{x^2 - y^2}{x^2 + y^2}, \quad y(x_0) = y_0. \tag{12.5.4}$$

Here

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2} \quad \text{and} \quad f_y(x, y) = -\frac{4x^2 y}{(x^2 + y^2)^2}$$

are continuous everywhere except at $(0, 0)$. If $(x_0, y_0) \neq (0, 0)$, there's an open rectangle R that contains (x_0, y_0) that does not contain $(0, 0)$. Since f and f_y are continuous on R, Theorem 2.5.1 implies that if $(x_0, y_0) \neq (0, 0)$ then Equation 12.5.4 has a unique solution on some open interval that contains x_0.

Example 2.5.3

Consider the initial value problem

$$y' = \frac{x + y}{x - y}, \quad y(x_0) = y_0. \tag{12.5.5}$$

Here

$$f(x, y) = \frac{x + y}{x - y} \quad \text{and} \quad f_y(x, y) = \frac{2x}{(x - y)^2}$$

are continuous everywhere except on the line $y = x$. If $y_0 \neq x_0$, there's an open rectangle R that contains (x_0, y_0) that does not intersect the line $y = x$. Since f and f_y are continuous on R, Theorem 2.5.1 implies that if $y_0 \neq x_0$, Equation 12.5.5 has a unique solution on some open interval that contains x_0.

Example 2.5.4

In Example 2.2.4, we saw that the solutions of

$$y' = 2xy^2 \tag{12.5.6}$$

are

$$y \equiv 0 \quad \text{and} \quad y = -\frac{1}{x^2 + c},$$

where c is an arbitrary constant. In particular, this implies that no solution of Equation 12.5.6 other than $y \equiv 0$ can equal zero for any value of x. Show that Theorem 12.5.1b implies this.

We'll obtain a contradiction by assuming that Equation 12.5.6 has a solution y_1 that equals zero for some value of x, but is not identically zero. If y_1 has this property, there's a point x_0 such that $y_1(x_0) = 0$, but $y_1(x) \neq 0$ for some value of x in every open interval that contains x_0. This means that the initial value problem

$$y' = 2xy^2, \quad y(x_0) = 0 \tag{12.5.7}$$

has two solutions $y \equiv 0$ and $y = y_1$ that differ for some value of x on every open interval that contains x_0. This contradicts Theorem 2.5.1 (b), since in Equation 12.5.6 the functions

$$f(x, y) = 2xy^2 \quad \text{and} \quad f_y(x, y) = 4xy.$$

are both continuous for all (x, y), which implies that Equation 12.5.7 has a unique solution on some open interval that contains x_0.

Below is a video on finding values where there is no guarantee of existence and uniqueness of a solution to a differential equation.



Example 2.5.5

Consider the initial value problem

$$y' = \frac{10}{3}xy^{2/5}, \quad y(x_0) = y_0. \tag{12.5.8}$$

a. For what points (x_0, y_0) does Theorem 12.5.1a imply that Equation 12.5.8 has a solution?
b. For what points (x_0, y_0) does Theorem 12.5.1b imply that Equation 12.5.8 has a unique solution on some open interval that contains x_0?

Solution a

Since

$$f(x, y) = \frac{10}{3}xy^{2/5}$$

is continuous for all (x, y), Theorem 2.5.1 implies that Equation 12.5.8 has a solution for every (x_0, y_0).

Solution b

Here

$$f_y(x, y) = \frac{4}{3}xy^{-3/5}$$

is continuous for all (x, y) with $y \neq 0$. Therefore, if $y_0 \neq 0$ there's an open rectangle on which both f and f_y are continuous, and Theorem 2.5.1 implies that Equation 12.5.8 has a unique solution on some open interval that contains x_0.

If $y = 0$ then $f_y(x, y)$ is undefined, and therefore discontinuous; hence, Theorem 2.5.1 does not apply to Equation 12.5.8 if $y_0 = 0$.

Example 2.5.6

Example 2.5.5 leaves open the possibility that the initial value problem

$$y' = \frac{10}{3} x y^{2/5}, \quad y(0) = 0 \tag{12.5.9}$$

has more than one solution on every open interval that contains $x_0 = 0$. Show that this is true.

Solution

By inspection, $y \equiv 0$ is a solution of the differential equation

$$y' = \frac{10}{3} x y^{2/5}. \tag{12.5.10}$$

Since $y \equiv 0$ satisfies the initial condition $y(0) = 0$, it is a solution of Equation 12.5.9.

Now suppose y is a solution of Equation 12.5.10 that is not identically zero. Separating variables in Equation 12.5.10 yields

$$y^{-2/5} y' = \frac{10}{3} x$$

on any open interval where y has no zeros. Integrating this and rewriting the arbitrary constant as $5c/3$ yields

$$\frac{5}{3} y^{3/5} = \frac{5}{3} (x^2 + c).$$

Therefore

$$y = (x^2 + c)^{5/3}. \tag{12.5.11}$$

Since we divided by y to separate variables in Equation 12.5.10, our derivation of Equation 12.5.11 is legitimate only on open intervals where y has no zeros. However, Equation 12.5.11 actually defines y for all x, and differentiating Equation 12.5.11 shows that

$$y' = \frac{10}{3} x (x^2 + c)^{2/3} = \frac{10}{3} x y^{2/5}, \ -\infty < x < \infty$$

Therefore Equation 12.5.11 satisfies Equation 12.5.10 on $(-\infty, \infty)$ even if $c \leq 0$, so that $y(\sqrt{|c|}) = y(-\sqrt{|c|}) = 0$. In particular, taking $c = 0$ in Equation 12.5.11 yields

$$y = x^{10/3}$$

as a second solution of Equation 12.5.9. Both solutions are defined on $(-\infty, \infty)$, and they differ on every open interval that contains $x_0 = 0$ (Figure 2.5.2). In fact, there are *four* distinct solutions of Equation 12.5.9 defined on $(-\infty, \infty)$ that differ from each other on every open interval that contains $x_0 = 0$. Can you identify the other two?

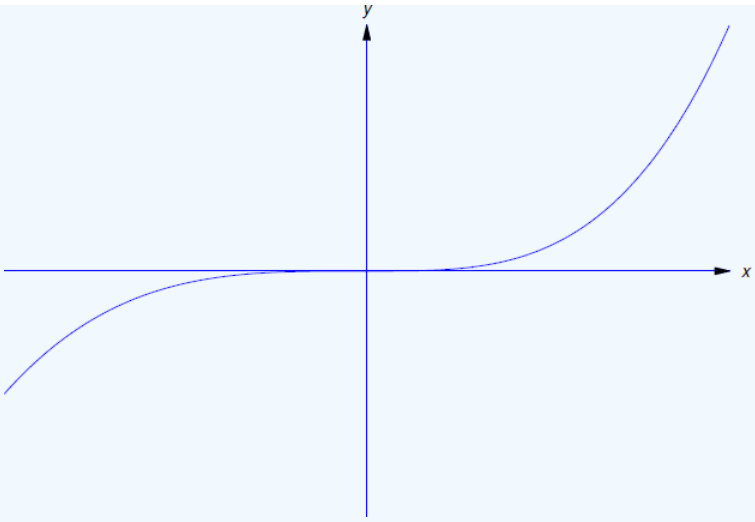


Figure 2.5.2 : Two solutions ($y = 0$ and $y = x^{1/2}$) of Equation 12.5.9 that differ on every interval containing $x_0 = 0$

Example 2.5.7

From Example 2.5.5 , the initial value problem

$$y' = \frac{10}{3}xy^{2/5}, \quad y(0) = -1 \tag{12.5.12}$$

has a unique solution on some open interval that contains $x_0 = 0$. Find a solution and determine the largest open interval (a, b) on which it is unique.

Solution

Let y be any solution of Equation 12.5.12 Because of the initial condition $y(0) = -1$ and the continuity of y, there's an open interval I that contains $x_0 = 0$ on which y has no zeros, and is consequently of the form Equation 12.5.11. Setting $x = 0$ and $y = -1$ in Equation 12.5.11yields $c = -1$, so

$$y = (x^2 - 1)^{5/3} \tag{12.5.13}$$

for x in I. Therefore every solution of Equation 12.5.12differs from zero and is given by Equation 12.5.13on $(-1, 1)$; that is, Equation 12.5.13is the unique solution of Equation 12.5.12on $(-1, 1)$. This is the largest open interval on which Equation 12.5.12has a unique solution. To see this, note that Equation 12.5.13is a solution of Equation 12.5.12on $(-\infty, \infty)$. From *Exercise 2.2.15*, there are infinitely many other solutions of Equation 12.5.12that differ from Equation 12.5.13on every open interval larger than $(-1, 1)$. One such solution is

$$y = \begin{cases} (x^2 - 1)^{5/3}, & -1 \le x \le 1, \\ 0, & |x| > 1. \end{cases}$$

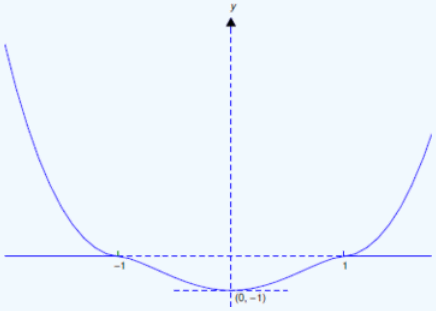


Figure 2.5.3 : Two solutions of Equation 12.5.12 on (−1,1) that coincide on (−1, 1), but on no larger open interval. (right)

Example 2.5.8

From Example 2.5.5), the initial value problem

$$y' = \frac{10}{3}xy^{2/5}, \quad y(0) = 1 \tag{12.5.14}$$

has a unique solution on some open interval that contains $x_0 = 0$. Find the solution and determine the largest open interval on which it is unique.

Solution

Let y be any solution of Equation 12.5.14. Because of the initial condition $y(0) = 1$ and the continuity of y, there's an open interval I that contains $x_0 = 0$ on which y has no zeros, and is consequently of the form Equation 12.5.11. Setting $x = 0$ and $y = 1$ in Equation 12.5.11yields $c = 1$, so

$$y = (x^2 + 1)^{5/3} \tag{12.5.15}$$

for x in I. Therefore every solution of Equation 12.5.14differs from zero and is given by Equation 12.5.15on $(-\infty, \infty)$; that is, Equation 12.5.15is the unique solution of Equation 12.5.14on $(-\infty, \infty)$. Figure 2.5.4) shows the graph of this solution.

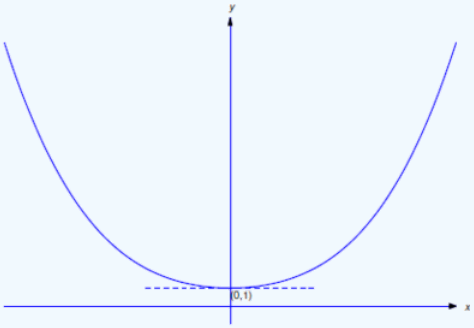


Figure 2.5.4 : The unique solution of Equation 12.5.14.

This page titled 12.5: Existence and Uniqueness of Solutions of Nonlinear Equations is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **2.3: Existence and Uniqueness of Solutions of Nonlinear Equations** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

12.5E: Existence and Uniqueness of Solutions of Nonlinear Equations (Exercises)

Q2.3.1

In **Exercises 2.3.1-2.3.13**, find all (x_0, y_0) for which Theorem 2.3.1 implies that the initial value problem $y' = f(x, y),\ y(x_0) = y_0$ has (a) a solution and (b) a unique solution on some open interval that contains x_0.

1. $y' = \frac{x^2+y^2}{\sin x}$
2. $y' = \frac{e^x+y}{x^2+y^2}$
3. $y' = \tan xy$
4. $y' = \frac{x^2+y^2}{\ln xy}$
5. $y' = (x^2+y^2)y^{1/3}$
6. $y' = 2xy$
7. $y' = \ln(1+x^2+y^2)$
8. $y' = \frac{2x+3y}{x-4y}$
9. $y' = (x^2+y^2)^{1/2}$
10. $y' = x(y^2-1)^{2/3}$
11. $y' = (x^2+y^2)^2$
12. $y' = (x+y)^{1/2}$
13. $y' = \frac{\tan y}{x-1}$

Q2.3.2

14. Apply Theorem 2.3.1 to the initial value problem

$$y' + p(x)y = q(x), \quad y(x_0) = y_0$$

for a linear equation, and compare the conclusions that can be drawn from it to those that follow from Theorem 2.1.2.

15.

a. Verify that the function

$$y = \begin{cases} (x^2-1)^{5/3}, & -1 < x < 1, \\ 0, & |x| \geq 1, \end{cases}$$

is a solution of the initial value problem

$$y' = \frac{10}{3}xy^{2/5}, \quad y(0) = -1$$

on $(-\infty, \infty)$. HINT: You'll need the definition

$$y'(\overline{x}) = \lim_{x \to \overline{x}} \frac{y(x) - y(\overline{x})}{x - \overline{x}}$$

to verify that y satisfies the differential equation at $\overline{x} = \pm 1$.

b. Verify that if $\epsilon_i = 0$ or 1 for $i = 1, 2$ and $a, b > 1$, then the function

$$y=\begin{cases}\epsilon_1(x^2-a^2)^{5/3}, & -\infty<x<-a,\\ 0, & -a\le x\le -1,\\ (x^2-1)^{5/3}, & -1<x<1,\\ 0, & 1\le x\le b,\\ \epsilon_2(x^2-b^2)^{5/3}, & b<x<\infty,\end{cases}$$

is a solution of the initial value problem of a on $(-\infty,\infty)$.

16. Use the ideas developed in **Exercise 2.3.15** to find infinitely many solutions of the initial value problem

$$y'=y^{2/5},\quad y(0)=1$$

on $(-\infty,\infty)$.

17. Consider the initial value problem

$$y'=3x(y-1)^{1/3},\quad y(x_0)=y_0. \tag{A}$$

a. For what points (x_0,y_0) does Theorem 2.3.1 imply that (A) has a solution?
b. For what points (x_0,y_0) does Theorem 2.3.1 imply that (A) has a unique solution on some open interval that contains x_0?

18. Find nine solutions of the initial value problem

$$y'=3x(y-1)^{1/3},\quad y(0)=1$$

that are all defined on $(-\infty,\infty)$ and differ from each other for values of x in every open interval that contains $x_0=0$.

19. From Theorem 2.3.1, the initial value problem

$$y'=3x(y-1)^{1/3},\quad y(0)=9$$

has a unique solution on an open interval that contains $x_0=0$. Find the solution and determine the largest open interval on which it is unique.

20.

a. From Theorem 2.3.1, the initial value problem

$$y'=3x(y-1)^{1/3},\quad y(3)=-7 \tag{A}$$

has a unique solution on some open interval that contains $x_0=3$. Determine the largest such open interval, and find the solution on this interval.
b. Find infinitely many solutions of (A), all defined on $(-\infty,\infty)$.

21. Prove:

a. If

$$f(x,y_0)=0,\quad a<x<b, \tag{A}$$

andx_0 is in (a,b), then $y\equiv y_0$ is a solution of

$$y'=f(x,y),\quad y(x_0)=y_0$$

on (a,b).
b. If f and f_y are continuous on an open rectangle that contains (x_0,y_0) and (A) holds, no solution of $y'=f(x,y)$ other than $y\equiv y_0$ can equal y_0 at any point in (a,b).

This page titled 12.5E: Existence and Uniqueness of Solutions of Nonlinear Equations (Exercises) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **2.3E: Existence and Uniqueness of Solutions of Nonlinear Equations (Exercises)** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

12.6: Substitution

Just as when solving integrals, one method to try is to change variables to end up with a simpler equation to solve.

12.6.1: Substitution

The equation

$$y' = (x - y + 1)^2 \tag{12.6.1}$$

is neither separable nor linear. What can we do? How about trying to change variables, so that in the new variables the equation is simpler. We use another variable v, which we treat as a function of x. Let us try

$$v = x - y + 1. \tag{12.6.2}$$

We need to figure out y' in terms of v', v and x. We differentiate (in x) to obtain $v' = 1 - y'$. So $y' = 1 - v'$. We plug this into the equation to get

$$1 - v' = v^2 \tag{12.6.3}$$

In other words, $v' = 1 - v^2$. Such an equation we know how to solve by separating variables:

$$\frac{1}{1 - v^2}\, dv = dx \tag{12.6.4}$$

So

$$\frac{1}{2}\ln\left|\frac{v+1}{v-1}\right| = x + C, \quad \text{or} \quad \left|\frac{v+1}{v-1}\right| = e^{2x+2C}, \quad \text{or} \quad \frac{v+1}{v-1} = De^{2x},$$

for some constant D. Note that $v = 1$ and $v = -1$ are also solutions.

Now we need to "unsubstitute" to obtain

$$\frac{x - y + 2}{x - y} = De^{2x} \tag{12.6.5}$$

and also the two solutions $x - y + 1 = 1$ or $y = x$, and $x - y + 1 = -1$ or $y = x + 2$. We solve the first equation for y.

$$\begin{aligned} x - y + 2 &= (x - y)\, De^{2x}, \\ x - y + 2 &= Dxe^{2x} - yDe^{2x}, \\ -y + yDe^{2x} &= Dxe^{2x} - x - 2, \\ y(-1 + De^{2x}) &= Dxe^{2x} - x - 2, \\ y &= \frac{Dxe^{2x} - x - 2}{De^{2x} - 1}. \end{aligned} \tag{12.6.6}$$

Note that $D = 0$ gives $y = x + 2$, but no value of D gives the solution $y = x$.

Substitution in differential equations is applied in much the same way that it is applied in calculus. You guess. Several different substitutions might work. There are some general things to look for. We summarize a few of these in a table.

When you see	Try substituting
yy'	$v = y^2$
y^2y'	$v = y^3$
$(\cos y)y'$	$v = \sin y$
$(\sin y)y'$	$v = \cos y$
$y'e^y$	$v = e^y$

Usually you try to substitute in the "most complicated" part of the equation with the hopes of simplifying it. The above table is just a rule of thumb. You might have to modify your guesses. If a substitution does not work (it does not make the equation any simpler), try a different one.

12.6.2: Bernoulli Equations

There are some forms of equations where there is a general rule for substitution that always works. One such example is the so-called Bernoulli equation.[1]

$$y' + p(x)\,y = q(x)\,y^n \tag{12.6.7}$$

This equation looks a lot like a linear equation except for the y^n. If $n = 0$ or $n = 1$, then the equation is linear and we can solve it. Otherwise, the substitution $v = y^{1-n}$ transforms the Bernoulli equation into a linear equation. Note that n need not be an integer.

Example 1.5.1: Bernoulli Equation

Solve

$$xy' + y(x+1) + xy^5 = 0, \;\; y(1) = 1$$

Solution

First, the equation is Bernoulli $p(x) = \dfrac{x+1}{x}$ (and $q(x) = -1$). We substitute

$$v = y^{1-5} = y^{-4}, \quad v' = -4y^{-5}y'$$

In other words, $\left(\dfrac{-1}{4}\right) y^5 v' = y'$. So

$$\begin{aligned} xy' + y(x+1) + xy^5 &= 0, \\ \frac{-xy^5}{4}v' + y(x+1) + xy^5 &= 0, \\ \frac{-x}{4}v' + y^{-4}(x+1) + x &= 0, \\ \frac{-x}{4}v' + v(x+1) + x &= 0, \end{aligned} \tag{12.6.8}$$

and finally

$$v' - \frac{4(x+1)}{x}v = 4$$

Now the equation is linear. We can use the integrating factor method. In particular, we use formula (1.4.17). Let us assume that $x > 0$ so $|x| = x$. This assumption is OK, as our initial condition is $x = 1$. Let us compute the integrating factor. Here $p(s)$ from formula (1.4.17) is $\dfrac{-4(s+1)}{s}$.

$$\begin{aligned} e^{\int_1^x p(s)ds} &= \exp\left(\int_1^x \frac{-4(s+1)}{s}ds\right) = e^{-4x-4ln(x)+4} = e^{-4x+4}x^{-4} = \frac{e^{-4x+4}}{x^4}, \\ e^{-\int_1^x p(s)ds} &= e^{4x+4ln(x)-4} = e^{4x-4}x^4 \end{aligned} \tag{12.6.9}$$

We now plug in to (1.4.17)

$$\begin{aligned} v(x) &= e^{-\int_1^x p(s)ds}\left(\int_1^x e^{\int_1^t p(s)ds}4dt + 1\right), \\ &= e^{4x-4}x^4\left(\int_1^x 4\frac{e^{-4t+4}}{t^4}dt + 1\right) \end{aligned} \tag{12.6.10}$$

Note that the integral in this expression is not possible to find in closed form. As we said before, it is perfectly fine to have a definite integral in our solution. Now "unsubstitute"

$$
\begin{aligned}
y^{-4} &= e^{4x-4}x^4\left(4\int_1^x \frac{e^{-4t+4}}{t^4}dt+1\right),\\
y &= \frac{e^{-x+1}}{x\left(4\int_1^x \frac{e^{-4t+4}}{t^4}dt+1\right)^{1/4}}
\end{aligned}
\tag{12.6.11}
$$

12.6.3: Homogeneous Equations

Another type of equations we can solve by substitution are the so-called homogeneous equations. Suppose that we can write the differential equation as

$$y' = F\left(\frac{y}{x}\right) \tag{12.6.12}$$

Here we try the substitutions

$$v = \frac{y}{x} \quad \text{and therefore} \quad y' = v + xv' \tag{12.6.13}$$

We note that the equation is transformed into

$$v + xv' = F(v) \quad \text{or} \quad xv' = F(v) - v \quad \text{or} \quad \frac{v'}{F(v)-v} = \frac{1}{x} \tag{12.6.14}$$

Hence an implicit solution is

$$\int \frac{1}{F(v)-v}dv = \ln|x| + C \tag{12.6.15}$$

Example 1.5.2

Solve

$$x^2y' = y^2 + xy, \quad y(1) = 1$$

Solution

We put the equation into the form $y' = \left(\frac{y}{x}\right)^2 + \frac{y}{x}$. We substitute $v = \frac{y}{x}$ to get the separable equation

$$xv' = v^2 + v - v = v^2$$

$$v^{-2}dv = \frac{1}{x}dx$$

which has a solution

$$
\begin{aligned}
\int \frac{1}{v^2}dv &= ln\,|x| + C,\\
\frac{-1}{v} &= ln\,|x| + C,\\
v &= \frac{-1}{ln\,|x| + C}.
\end{aligned}
\tag{12.6.16}
$$

We unsubstitute

$$
\begin{aligned}
\frac{y}{x} &= \frac{-1}{ln\,|x| + C},\\
y &= \frac{-x}{ln\,|x| + C}
\end{aligned}
\tag{12.6.17}
$$

We want $y(1) = 1$, so

$$1 = y(1) = \frac{-1}{ln\,|1| + C} = \frac{-1}{C}$$

Thus $C = -1$ and the solution we are looking for is

$$y = \frac{-x}{ln\,|x| - 1}$$

12.6.4: Footnotes

[1] There are several things called Bernoulli equations, this is just one of them. The Bernoullis were a prominent Swiss family of mathematicians. These particular equations are named for Jacob Bernoulli (1654–1705).

- ◦ Jiří Lebl (Oklahoma State University).These pages were supported by NSF grants DMS-0900885 and DMS-1362337.

This page titled 12.6: Substitution is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **1.5: Substitution** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

12.7: Autonomous equations

Let us consider general differential equation problems of the form

$$\frac{dx}{dt} = f(x)$$

where the derivative of solutions depends only on x (the dependent variable). Such equations are called **autonomous** equations. If we think of t as time, the naming comes from the fact that the equation is independent of time.

Let us come back to the cooling coffee problem (see Example 1.3.3). Newton's law of cooling says that

$$\frac{dx}{dt} = -k(x - A)$$

where x is the temperature, t is time, k is some constant, and A is the ambient temperature. See Figure 12.7.1 for an example with $k = 0.3$ and $A = 5$.

Note the solution $x = A$ (in the figure $x = 5$). We call these constant solutions the equilibrium solutions. The points on the x axis where $f(X) = 0$ are called critical points. The point $x = A$ is a critical point. In fact, each critical point corresponds to an equilibrium solution. Note also, by looking at the graph, that the solution $x = A$ is "stable" in that small perturbations in x do not lead to substantially different solutions as t grows. If we change the initial condition a little bit, then as $t \to \infty$ we get $x \to A$. We call such critical points stable. In this simple example it turns out that all solutions in fact go to A as $t \to \infty$. If a critical point is not stable we would say it is unstable.

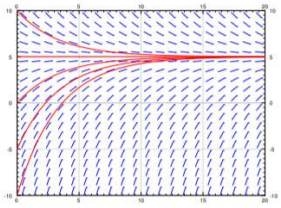


Figure 12.7.1: The slope field and some solutions of $x' = 0.3(5 - x)$.

Let us consider the logistic equation

$$\frac{dx}{dt} = kx(M - x)$$

for some positive k and M. This equation is commonly used to model population if we know the limiting population M, that is the maximum sustainable population. The logistic equation leads to less catastrophic predictions on world population than $x' = kx$. In the real world there is no such thing as negative population, but we will still consider negative x for the purposes of the math (see Figure 12.7.2 for an example).

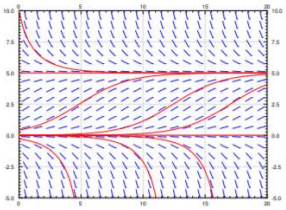


Figure 12.7.2: The slope field and some solutions of $x' = 0.1x(5 - x)$.

Note two critical points, $x = 0$ and $x = 5$. The critical point at $x = 5$. is stable. On the other hand the critical point at $x = 0$. is unstable.

It is not really necessary to find the exact solutions to talk about the long term behavior of the solutions. For example, from the above slope field plot, we can easily see that

$$\lim_{t \to \infty} x(t) = \begin{cases} 5 & \text{if } x(0) > 0, \\ 0 & \text{if } x(0) = 0, \\ \text{DNE or } -\infty & \text{if } x(0) < 0. \end{cases}$$

Where DNE means "does not exist." From just looking at the slope field we cannot quite decide what happens if $x(0) < 0$. It could be that the solution does not exist for t all the way to ∞. Think of the equation $x' = x^2$, we have seen that it only exists for some finite period of time. Same can happen here. In our example equation above it will actually turn out that the solution does not exist for all time, but to see that we would have to solve the equation. In any case, the solution does go to $-\infty$, but it may get there rather quickly.

If we are interested only in the long term behavior of the solution, we would be doing unnecessary work if we solved the equation exactly. We could draw the slope field, but it is easier to just look at the or , which is a simple way to visualize the behavior of autonomous equations. In this case there is one dependent variable x. We draw the x-axis, we mark all the critical points, and then we draw arrows in between. Since x is the dependent variable we draw the axis vertically, as it appears in the slope field diagrams above. If $f(x) > 0$, we draw an up arrow. If $f(x) < 0$, we draw a down arrow. To figure this out, we could just plug in some x between the critical points, $f(x)$ will have the same sign at all x between two critical points as long $f(x)$ is continuous. For example, $f(6) = -0.6 < 0$, so $f(x) < 0$ for $x > 5$, and the arrow above $x = 5$ is a down arrow. Next, $f(1) = 0.4 > 0$, so $f(x) > 0$ whenever $0 < x < 5$, and the arrow points up. Finally, $f(-1) = -0.6 < 0$ so $f(x) < 0$ when $x < 0$, and the arrow points down.

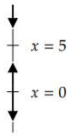
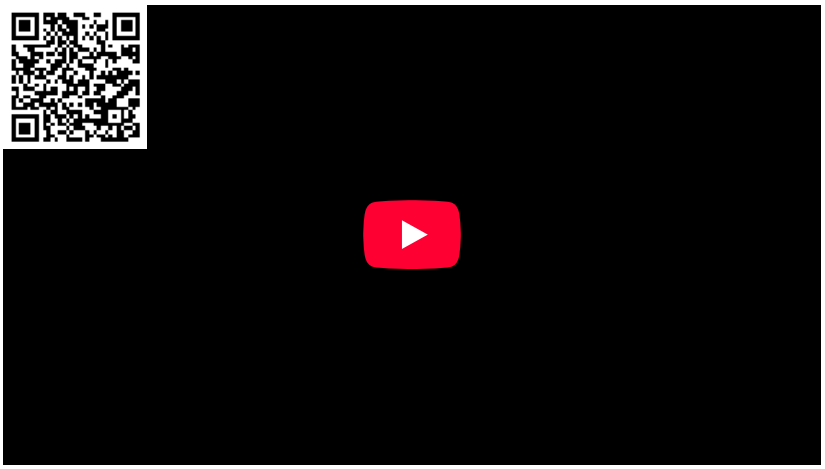


Figure 12.7.3

Armed with the phase diagram, it is easy to sketch the solutions approximately: As time t moves from left to right, the graph of a solution goes up if the arrow is up, and it goes down if the arrow is down.

Below is a video on solving an autonomous differential equation that describes logistic growth.



Exercise 12.7.1

Try sketching a few solutions simply from looking at the phase diagram. Check with the preceding graphs if you are getting the type of curves.

Once we draw the phase diagram, we can easily classify critical points as stable or unstable.[1]



Figure 12.7.4

Since any mathematical model we cook up will only be an approximation to the real world, unstable points are generally bad news.

Let us think about the logistic equation with harvesting. Suppose an alien race really likes to eat humans. They keep a planet with humans on it and harvest the humans at a rate of h million humans per year. Suppose$x$ is the number of humans in millions on the planet and t is time in years. Let M be the limiting population when no harvesting is done and $k > 0$ is some constant depending on how fast humans multiply. Our equation becomes

$$\frac{dx}{dt} = kx(M - x) - h$$

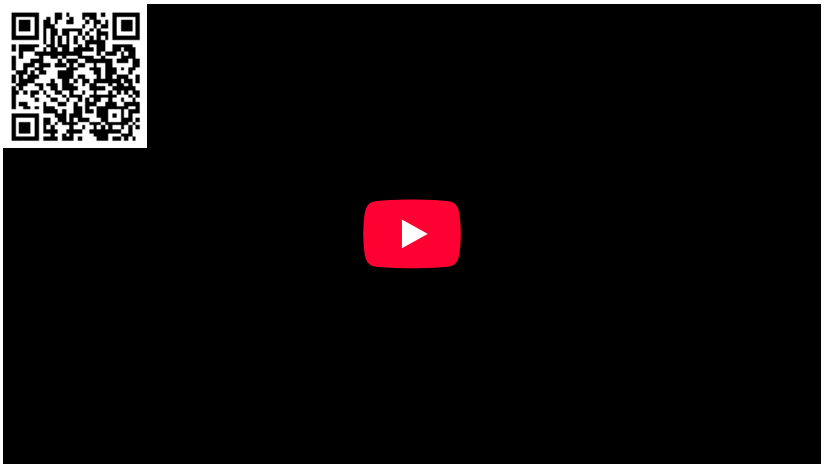
We expand the right hand side and solve for critical points

$$\frac{dx}{dt} = -kx^2 + kMx - h$$

Solving for the critical points A and B from the quadratic equations:

$$A = \frac{kM + \sqrt{(kM)^2 - 4hk}}{2k}, \quad B = \frac{kM - \sqrt{(kM)^2 - 4hk}}{2k}$$

Below is a video on solving an autonomous initial value problem.



Exercise 12.7.2

Sketch a phase diagram for different possibilities. Note that these possibilities are $A > B$, or $A = B$, or A and B both complex (i.e. no real solutions). Hint: Fix some simple k and M and then vary h.

For example, let $M = 8$ and $k = 0.1$. When $h = 1$, then A and B are distinct and positive. The slope field we get is in Figure 12.7.5. As long as the population starts above B, which is approximately 1.55 million, then the population will not die out. It will in fact tend towards $A \approx 6.45$ million. If ever some catastrophe happens and the population drops below B, humans will die out, and the fast food restaurant serving them will go out of business.

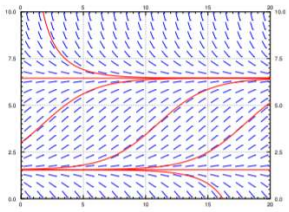


Figure 12.7.5: Slope field and some solutions of $x' = 0.1x(8-x) - 1$.

When $h = 1.6$, then $A = B = 4$ and there is only **one** critical point and it is **unstable**. When the population starts above 4 million it will tend towards 4 million. If it ever drops below 4 million, humans will die out on the planet. This scenario is not one that we (as the human fast food proprietor) want to be in. A small perturbation of the equilibrium state and we are out of business; there is no room for error (see Figure 12.7.6).

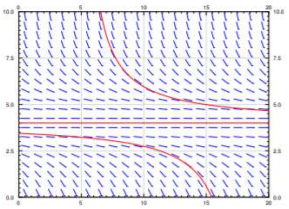


Figure 12.7.6: The slope field and some solutions of $x' = 0.1x(8-x) - 1.6$.

Finally if we are harvesting at 2 million humans per year, there are no critical points. The population will always plummet towards zero, no matter how well stocked the planet starts (see Figure 12.7.7).

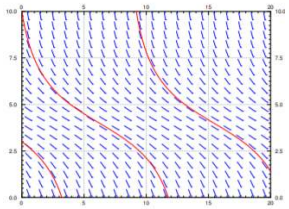


Figure 12.7.7: Slope field and some solutions of $x' = 0.1x(8-x) - 2$.

12.7.1: Footnotes

[1] Unstable points with one of the arrows pointing towards the critical point are sometimes called semistable.

12.7.2: References

1. Paul W. Berg and James L. McGregor, Elementary Partial Differential Equations, Holden-Day, San Francisco, CA, 1966.
2. William E. Boyce, Richard C. DiPrima, Elementary Differential Equations and Boundary Value Problems, 9th edition, John Wiley & Sons Inc., New York, NY, 2008.
3. C.H. Edwards and D.E. Penney, Differential Equations and Boundary Value Problems: Computing and Modeling, 4th edition, Prentice Hall, 2008.
4. Stanley J. Farlow, An Introduction to Differential Equations and Their Applications, McGraw-Hill, Inc., Princeton, NJ, 1994. (Published also by Dover Publications, 2006.)
5. E.L. Ince, Ordinary Differential Equations, Dover Publications, Inc., New York, NY, 1956.

This page titled 12.7: Autonomous equations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **1.6: Autonomous equations** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

12.8: Numerical methods- Euler's method

At this point it may be good to first try the Lab II and/or Project II from the IODE website: www.math.uiuc.edu/iode/. As we said before, unless $f(x,y)$ is of a special form, it is generally very hard if not impossible to get a nice formula for the solution of the problem

$$y' = f(x,y), \quad y(x_0) = y_0$$

If the equation can be solved in closed form, we should do that. But what if we have an equation that cannot be solved in closed form? What if we want to find the value of the solution at some particular x? Or perhaps we want to produce a graph of the solution to inspect the behavior. In this section we will learn about the basics of numerical approximation of solutions.

The simplest method for approximating a solution is Euler's Method.[1] It works as follows: Take x_0 and compute the slope $k = f(x_0, y_0)$. The slope is the change in y per unit change in x. Follow the line for an interval of length h on the x-axis. Hence if $y = y_0$ at x_0, then we say that y_1 (the approximate value of y at $x_1 = x_0 + h$) is $y_1 = y_0 + hk$. Rinse, repeat! Let $k = f(x_1, y_1)$, and then compute $x_2 = x_1 + h$, and $y_2 = y_1 + hk$. Now compute x_3 and y_3 using x_2 and y_2, etc. Consider the equation $y' = \frac{y^2}{3}$, $y(0) = 1$, and $h = 1$. Then $x_0 = 0$ and $y_0 = 1$. We compute

$$\begin{aligned} x_1 &= x_0 + h = 0 + 1 = 1, \qquad & y_1 &= y_0 + h\, f(x_0, y_0) = 1 + 1 \cdot \frac{1}{3} = \frac{4}{3} \approx 1.333, \\ x_2 &= x_1 + h = 1 + 1 = 2, \qquad & y_2 &= y_1 + h\, f(x_1, y_1) = \frac{4}{3} + 1 \cdot \frac{\left(\frac{4}{3}\right)^2}{3} = \frac{52}{27} \approx 1.926. \end{aligned} \tag{12.8.1}$$

We then draw an approximate graph of the solution by connecting the points (x_0, y_0), (x_1, y_1), (x_2, y_2),.... For the first two steps of the method see Figure 12.8.1.

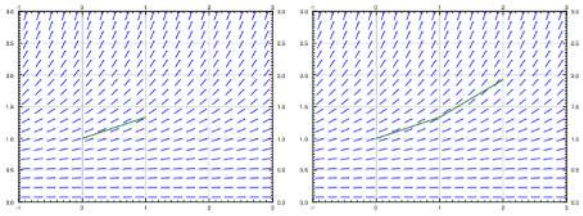


Figure 12.8.1: First two steps of Euler's method with $h = 1$ for the equation $y' = \frac{y^2}{3}$ with initial conditions $y(0) = 1$.

More abstractly, for any $i = 0, 1, 2, 3, \ldots$, we compute

$$x_{i+1} = x_i + h, \qquad y_{i+1} = y_i + h\, f(x_i, y_i).$$

The line segments we get are an approximate graph of the solution. Generally it is not exactly the solution. See Figure 12.8.2 for the plot of the real solution and the approximation.

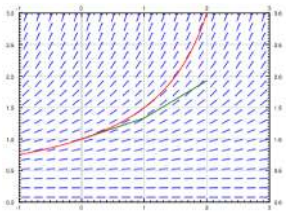


Figure 12.8.2: Two steps of Euler's method (step size 1) and the exact solution for the equation $y' = \frac{y^2}{3}$ with initial conditions $y(0) = 1$.

We continue with the equation $y' = \frac{y^2}{3}$, $y(0) = 1$. Let us try to approximate $y(2)$ using Euler's method. In Figures 12.8.1 and 12.8.2 we have graphically approximated $y(2)$ with step size 1. With step size 1, we have $y(2) \approx 1.926$. The real answer is 3. We are approximately 1.074 off. Let us halve the step size. Computing y_4 with $h = 0.5$, we find that $y(2) \approx 2.209$, so an error of about 0.791. Table 12.8.1 gives the values computed for various parameters.

? Exercise 12.8.1

Solve this equation exactly and show that $y(2) = 3$.

The difference between the actual solution and the approximate solution we will call the error. We will usually talk about just the size of the error and we do not care much about its sign. The main point is, that we usually do not know the real solution, so we only have a vague understanding of the error. If we knew the error exactly …what is the point of doing the approximation?

h	Approximate $y(2)$	Error	$\frac{\text{Error}}{\text{Previous error}}$
1	1.92593	1.07407	
0.5	2.20861	0.79139	0.73681
0.25	2.47250	0.52751	0.66656
0.125	2.68034	0.31966	0.60599
0.0625	2.82040	0.17960	0.56184
0.03125	2.90412	0.09588	0.53385
0.015625	2.95035	0.04965	0.51779
0.0078125	2.97472	0.02528	0.50913

Table 12.8.1: Euler's method approximation of $y(2)$ where of $y' = \frac{y^2}{3}$, $y(0) = 1$.

We notice that except for the first few times, every time we halved the interval the error approximately halved. This halving of the error is a general feature of Euler's method as it is a first order method. In the IODE Project II you are asked to implement a second order method. A second order method reduces the error to approximately one quarter every time we

halve the interval (second order as $\frac{1}{4} = \frac{1}{2}\text{x}\frac{1}{2}$).

To get the error to be within 0.1 of the answer we had to already do 64 steps. To get it to within 0.01 we would have to halve another three or four times, meaning doing 512 to 1024 steps. That is quite a bit to do by hand. The improved Euler method from IODE Project II should quarter the error every time we halve the interval, so we would have to approximately do half as many "halvings" to get the same error. This reduction can be a big deal. With 10 halvings (starting at $h = 1$) we have 1024 steps, whereas with 5 halvings we only have to do 32 steps, assuming that the error was comparable to start with. A computer may not care about this difference for a problem this simple, but suppose each step would take a second to compute (the function may be substantially more difficult to compute than $\frac{y^2}{3}$). Then the difference is 32 seconds versus about 17 minutes. Note: We are not being altogether fair, a second order method would probably double the time to do each step. Even so, it is 1 minute versus 17 minutes. Next, suppose that we have to repeat such a calculation for different parameters a thousand times. You get the idea.

Below is a video on using Euler's Method to approximate the solution to a differential equation.



Note that in practice we do not know how large the error is! How do we know what is the right step size? Well, essentially we keep halving the interval, and if we are lucky, we can estimate the error from a few of these calculations and the assumption that the error goes down by a factor of one half each time (if we are using standard Euler).

? Exercise 12.8.2

In the table above, suppose you do not know the error. Take the approximate values of the function in the last two lines, assume that the error goes down by a factor of 2. Can you estimate the error in the last time from this? Does it (approximately) agree with the table? Now do it for the first two rows. Does this agree with the table?

Let us talk a little bit more about the example $y' = \frac{y^2}{3}$, $y(0) = 1$. Suppose that instead of the value $y(2)$ we wish to find $y(3)$. The results of this effort are listed in Table 12.8.2 for successive halvings of h. What is going on here? Well, you should solve the equation exactly and you will notice that the solution does not exist at $x = 3$. In fact, the solution goes to infinity when you approach $x = 3$.

h	Approximate $y(3)$
1	3.16232
0.5	4.54329
0.25	6.86079
0.125	10.80321
0.0625	17.59893
0.03125	29.46004
0.015625	50.40121
0.0078125	87.75769

Table 12.8.2 Attempts to use Euler's to approximate $y(3)$ where of $y' = \frac{y^2}{3}$, $y(0) = 1$.

Another case when things can go bad is if the solution oscillates wildly near some point. Such an example is given in IODE Project II. The solution may exist at all points, but even a much better numerical method than Euler would need an insanely small step size to approximate the solution with reasonable precision. And computers might not be able to easily handle such a small step size.

In real applications we would not use a simple method such as Euler's. The simplest method that would probably be used in a real application is the standard Runge-Kutta method (see exercises). That is a fourth order method, meaning that if we halve the interval, the error generally goes down by a factor of 16 (it is fourth order as $\frac{1}{16} = \frac{1}{2}\text{x}\frac{1}{2}\text{x}\frac{1}{2}\text{x}\frac{1}{2}$).

Choosing the right method to use and the right step size can be very tricky. There are several competing factors to consider.

- Computational time: Each step takes computer time. Even if the function f is simple to compute, we do it many times over. Large step size means faster computation, but perhaps not the right precision.
- Roundoff errors: Computers only compute with a certain number of significant digits. Errors introduced by rounding numbers off during our computations become noticeable when the step size becomes too small relative to the quantities we are working with. So reducing step size may in fact make errors worse.
- Stability: Certain equations may be numerically unstable. What may happen is that the numbers never seem to stabilize no matter how many times we halve the interval. We may need a ridiculously small interval size, which may not be practical due to roundoff errors or computational time considerations. Such problems are sometimes called stiff. In the worst case, the numerical computations might be giving us bogus numbers that look like a correct answer. Just because the numbers have stabilized after successive halving, does not mean that we must have the right answer.

Below is a video on using Euler's Method to appoximation a solution to a differential equation.



We have seen just the beginnings of the challenges that appear in real applications. Numerical approximation of solutions to differential equations is an active research area for engineers and mathematicians. For example, the general purpose method used for the ODE solver in Matlab and Octave (as of this writing) is a method that appeared in the literature only in the 1980s.

12.8.1: Footnotes

[1] Named after the Swiss mathematician Leonhard Paul Euler (1707–1783). The correct pronunciation of the name sounds more like “oiler.”

This page titled 12.8: Numerical methods- Euler’s method is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **1.7: Numerical methods: Euler’s method** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

12.9: Exact Equations

Another type of equation that comes up quite often in physics and engineering is an *exact equation*. Suppose $F(x,y)$ is a function of two variables, which we call the *potential function*. The naming should suggest potential energy, or electric potential. Exact equations and potential functions appear when there is a conservation law at play, such as conservation of energy. Let us make up a simple example. Let

$$F(x,y) = x^2 + y^2.$$

We are interested in the lines of constant energy, that is lines where the energy is conserved; we want curves where $F(x,y) = C$, for some constant C. In our example, the curves $x^2 + y^2 = C$ are circles. See Figure 12.9.1.

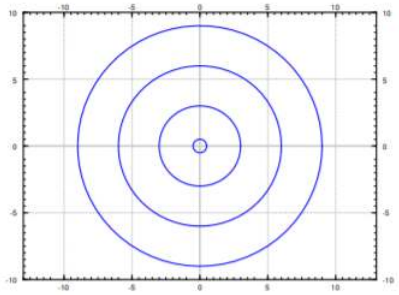


Figure 12.9.1: Solutions to $F(x,y) = x^2 + y^2 = C$ for various C.

We take the *total derivative* of F:

$$dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy.$$

For convenience, we will make use of the notation of $F_x = \frac{\partial F}{\partial x}$ and $F_y = \frac{\partial F}{\partial y}$. In our example,

$$dF = 2x\,dx + 2y\,dy.$$

We apply the total derivative to $F(x,y) = C$, to find the differential equation $dF = 0$. The differential equation we obtain in such a way has the form

$$M\,dx + N\,dy = 0, \qquad \text{or} \qquad M + N\,\frac{dy}{dx} = 0.$$

An equation of this form is called *exact* if it was obtained as $dF = 0$ for some potential function F. In our simple example, we obtain the equation

$$2x\,dx + 2y\,dy = 0, \qquad \text{or} \qquad 2x + 2y\,\frac{dy}{dx} = 0.$$

Since we obtained this equation by differentiating $x^2 + y^2 = C$, the equation is exact. We often wish to solve for y in terms of x. In our example,

$$y = \pm\sqrt{C^2 - x^2}.$$

An interpretation of the setup is that at each point $\vec{v} = (M,N)$ is a vector in the plane, that is, a direction and a magnitude. As M and N are functions of (x,y), we have a *vector field*. The particular vector field $\vec{v}$ that comes from an exact equation is a so-called *conservative vector field*, that is, a vector field that comes with a potential function $F(x,y)$, such that

$$\vec{v} = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}\right).$$

Let γ be a path in the plane starting at (x_1,y_1) and ending at (x_2,y_2). If we think of $\vec{v}$ as force, then the work required to move along γ is

$$\int_\gamma \vec{v}(\vec{r}) \cdot d\vec{r} = \int_\gamma M\,dx + N\,dy = F(x_2,y_2) - F(x_1,y_1).$$

That is, the work done only depends on endpoints, that is where we start and where we end. For example, suppose F is gravitational potential. The derivative of F given by $\vec{v}$ is the gravitational force. What we are saying is that the work required to move a heavy box from the ground floor to the roof, only depends on the change in potential energy. That is, the work done is the same no matter what path we took; if we took the stairs or the elevator. Although if we took the elevator, the elevator is doing the work for us. The curves $F(x,y)=C$ are those where no work need be done, such as the heavy box sliding along without accelerating or breaking on a perfectly flat roof, on a cart with incredibly well oiled wheels.

An exact equation is a conservative vector field, and the implicit solution of this equation is the potential function.

12.9.1: Solving exact equations

Now you, the reader, should ask: Where did we solve a differential equation? Well, in applications we generally know M and N, but we do not know F. That is, we may have just started with $2x+2y\frac{dy}{dx}=0$, or perhaps even

$$x+y\frac{dy}{dx}=0.$$

It is up to us to find some potential F that works. Many different F will work; adding a constant to F does not change the equation. Once we have a potential function F, the equation $F\big(x,y(x)\big)=C$ gives an implicit solution of the ODE.

Below is a video on solving an exact first order differential equation.



Below is another video on solving an exact first order differential equation.



Example 12.9.1

Let us find the general solution to $2x+2y\frac{dy}{dx}=0$. Forget we knew what F was.

Solution

If we know that this is an exact equation, we start looking for a potential function F. We have $M = 2x$ and $N = 2y$. If F exists, it must be such that $F_x(x, y) = 2x$. Integrate in the x variable to find

$$F(x, y) = x^2 + A(y), \tag{12.9.1}$$

for some function $A(y)$. The function A is the , though it is only constant as far as x is concerned, and may still depend on y. Now differentiate (12.9.1) in y and set it equal to N, which is what F_y is supposed to be:

$$2y = F_y(x, y) = A'(y).$$

Integrating, we find $A(y) = y^2$. We could add a constant of integration if we wanted to, but there is no need. We found $F(x, y) = x^2 + y^2$. Next for a constant C, we solve

$$F\big(x, y(x)\big) = C.$$

for y in terms of x. In this case, we obtain $y = \pm\sqrt{C^2 - x^2}$ as we did before.

? Exercise 12.9.1

Why did we not need to add a constant of integration when integrating $A'(y) = 2y$? Add a constant of integration, say 3, and see what F you get. What is the difference from what we got above, and why does it not matter?

The procedure, once we know that the equation is exact, is:

i. Integrate $F_x = M$ in x resulting in $F(x, y) = \text{something} + A(y)$.
ii. Differentiate this F in y, and set that equal to N, so that we may find $A(y)$ by integration.

The procedure can also be done by first integrating in y and then differentiating in x. Pretty easy huh? Let's try this again.

✓ Example 12.9.2

Consider now $2x + y + xy\frac{dy}{dx} = 0$.

OK, so $M = 2x + y$ and $N = xy$. We try to proceed as before. Suppose F exists. Then $F_x(x, y) = 2x + y$. We integrate:

$$F(x, y) = x^2 + xy + A(y)$$

for some function $A(y)$. Differentiate in y and set equal to N:

$$N = xy = F_y(x, y) = x + A'(y).$$

But there is no way to satisfy this requirement! The function xy cannot be written as x plus a function of y. The equation is not exact; no potential function F exists.

But there is no way to satisfy this requirement! The function xy cannot be written as x plus a function of y. The equation is not exact; no potential function F exists

Below is a video on solving an exact first order differential equation.



Below is another video on solving an exact first order differential equation.



Below is another video on solving an exact first order differential equation.



Below is another video on solving an exact first order differential equation.



Is there an easier way to check for the existence of F, other than failing in trying to find it? Turns out there is. Suppose $M = F_x$ and $N = F_y$. Then as long as the second derivatives are continuous,

$$\frac{\partial M}{\partial y} = \frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 F}{\partial x \partial y} = \frac{\partial N}{\partial x}.$$

Let us state it as a theorem. Usually this is called the Poincaré Lemma.[1]

Theorem 12.9.1

Pointcaré

If M and N are continuously differentiable functions of (x, y), and $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$, then near any point there is a function $F(x, y)$ such that $M = \frac{\partial F}{\partial x}$ and $N = \frac{\partial F}{\partial y}$.

The theorem doesn't give us a global F defined everywhere. In general, we can only find the potential locally, near some initial point. By this time, we have come to expect this from differential equations.

Let us return to Example 12.9.2where $M = 2x + y$ and $N = xy$. Notice $M_y = 1$ and $N_x = y$, which are clearly not equal. The equation is not exact.

Example 12.9.3

Solve

$$\frac{dy}{dx} = \frac{-2x - y}{x - 1}, \qquad y(0) = 1.$$

Solution

We write the equation as

$$(2x + y) + (x - 1)\frac{dy}{dx} = 0,$$

so $M = 2x + y$ and $N = x - 1$. Then

$$M_y = 1 = N_x.$$

The equation is exact. Integrating M in x, we find

$$F(x, y) = x^2 + xy + A(y).$$

Differentiating in y and setting to N, we find

$$x - 1 = x + A'(y).$$

So $A'(y) = -1$, and $A(y) = -y$ will work. Take $F(x,y) = x^2 + xy - y$. We wish to solve $x^2 + xy - y = C$. First let us find C. As $y(0) = 1$ then $F(0,1) = C$. Therefore $0^2 + 0 \times 1 - 1 = C$, so $C = -1$. Now we solve $x^2 + xy - y = -1$ for y to get

$$y = \frac{-x^2 - 1}{x - 1}.$$

Example 12.9.4

Solve

$$-\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy = 0, \qquad y(1) = 2.$$

Solution

We leave to the reader to check that $M_y = N_x$.

This vector field (M, N) is not conservative if considered as a vector field of the entire plane minus the origin. The problem is that if the curve γ is a circle around the origin, say starting at $(1,0)$ and ending at $(1,0)$ going counterclockwise, then if F existed we would expect

$$0 = F(1,0) - F(1,0) = \int_\gamma F_x\,dx + F_y\,dy = \int_\gamma \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy = 2\pi.$$

That is nonsense! We leave the computation of the path integral to the interested reader, or you can consult your multivariable calculus textbook. So there is no potential function F defined everywhere outside the origin $(0,0)$.

If we think back to the theorem, it does not guarantee such a function anyway. It only guarantees a potential function locally, that is only in some region near the initial point. As $y(1) = 2$ we start at the point $(1,2)$. Considering $x > 0$ and integrating M in x or N in y, we find

$$F(x,y) = \arctan\Big(\frac{y}{x}\Big).$$

The implicit solution is $\arctan\big(\frac{y}{x}\big) = C$. Solving, $y = \tan(C)x$. That is, the solution is a straight line. Solving $y(1) = 2$ gives us that $\tan(C) = 2$, and so $y = 2x$ is the desired solution. See Figure 12.9.1, and note that the solution only exists for $x > 0$.

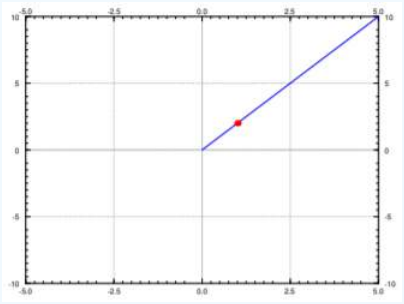


Figure 12.9.1: Solution to $-\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy = 0$, $y(1) = 2$, with initial point marked.

Example 12.9.5

Solve

$$x^2 + y^2 + 2y(x+1)\frac{dy}{dx} = 0.$$

Solution

The reader should check that this equation is exact. Let $M = x^2 + y^2$ and $N = 2y(x+1)$. We follow the procedure for exact equations

$$F(x,y) = \frac{1}{3}x^3 + xy^2 + A(y),$$

and

$$2y(x+1) = 2xy + A'(y).$$

Therefore $A'(y) = 2y$ or $A(y) = y^2$ and $F(x,y) = \frac{1}{3}x^3 + xy^2 + y^2$. We try to solve $F(x,y) = C$. We easily solve for y^2 and then just take the square root:

$$y^2 = \frac{C - (\frac{1}{3})x^3}{x+1}, \qquad \text{so} \qquad y = \pm\sqrt{\frac{C - (\frac{1}{3})x^3}{x+1}}.$$

When $x = -1$, the term in front of $\frac{dy}{dx}$ vanishes. You can also see that our solution is not valid in that case. However, one could in that case try to solve for x in terms of y starting from the implicit solution $\frac{1}{3}x^3 + xy^2 + y^2 = C$. The solution is somewhat messy and we leave it as implicit.

12.9.2: Integrating factors

Sometimes an equation $M\,dx + N\,dy = 0$ is not exact, but it can be made exact by multiplying with a function $u(x,y)$. That is, perhaps for some nonzero function $u(x,y)$,

$$u(x,y)M(x,y)\,dx + u(x,y)N(x,y)\,dy = 0$$

is exact. Any solution to this new equation is also a solution to $M\,dx + N\,dy = 0$.

In fact, a linear equation

$$\frac{dy}{dx} + p(x)y = f(x), \qquad \text{or} \qquad \big(p(x)y - f(x)\big)\,dx + dy = 0$$

is always such an equation. Let $r(x) = e^{\int p(x)\,dx}$ be the integrating factor for a linear equation. Multiply the equation by $r(x)$ and write it in the form of $M + N\frac{dy}{dx} = 0$.

$$r(x)p(x)y - r(x)f(x) + r(x)\frac{dy}{dx} = 0.$$

Then $M = r(x)p(x)y - r(x)f(x)$, so $M_y = r(x)p(x)$, while $N = r(x)$, so $N_x = r'(x) = r(x)p(x)$. In other words, we have an exact equation. Integrating factors for linear functions are just a special case of integrating factors for exact equations.

But how do we find the integrating factor u? Well, given an equation

$$M\,dx + N\,dy = 0,$$

u should be a function such that

$$\frac{\partial}{\partial y}\big[uM\big] = u_y M + uM_y = \frac{\partial}{\partial x}\big[uN\big] = u_x N + uN_x.$$

Therefore,

$$(M_y - N_x)u = u_x N - u_y M.$$

At first it may seem we replaced one differential equation by another. True, but all hope is not lost.

A strategy that often works is to look for a u that is a function of x alone, or a function of y alone. If u is a function of x alone, that is $u(x)$, then we write $u'(x)$ instead of u_x, and u_y is just zero. Then

$$\frac{M_y - N_x}{N}u = u'.$$

In particular, $\frac{M_y - N_x}{N}$ ought to be a function of x alone (not depend on y). If so, then we have a linear equation

$$u' - \frac{M_y - N_x}{N} u = 0.$$

Letting $P(x) = \frac{M_y - N_x}{N}$, we solve using the standard integrating factor method, to find $u(x) = Ce^{\int P(x)\,dx}$. The constant in the solution is not relevant, we need any nonzero solution, so we take $C = 1$. Then $u(x) = e^{\int P(x)\,dx}$ is the integrating factor.

Similarly we could try a function of the form $u(y)$. Then

$$\frac{M_y - N_x}{M} u = -u'.$$

In particular, $\frac{M_y - N_x}{M}$ ought to be a function of y alone. If so, then we have a linear equation

$$u' + \frac{M_y - N_x}{M} u = 0.$$

Letting $Q(y) = \frac{M_y - N_x}{M}$, we find $u(y) = Ce^{-\int Q(y)\,dy}$. We take $C = 1$. So $u(y) = e^{-\int Q(y)\,dy}$ is the integrating factor.

Example 12.9.6

Solve

$$\frac{x^2 + y^2}{x+1} + 2y\frac{dy}{dx} = 0.$$

Solution

Let $M = \frac{x^2+y^2}{x+1}$ and $N = 2y$. Compute

$$M_y - N_x = \frac{2y}{x+1} - 0 = \frac{2y}{x+1}.$$

As this is not zero, the equation is not exact. We notice

$$P(x) = \frac{M_y - N_x}{N} = \frac{2y}{x+1}\frac{1}{2y} = \frac{1}{x+1}$$

is a function of x alone. We compute the integrating factor

$$e^{\int P(x)\,dx} = e^{\ln(x+1)} = x+1.$$

We multiply our given equation by $(x+1)$ to obtain

$$x^2 + y^2 + 2y(x+1)\frac{dy}{dx} = 0,$$

which is an exact equation that we solved in Example 12.9.5. The solution was

$$y = \pm\sqrt{\frac{C - (\frac{1}{3})x^3}{x+1}}.$$

Example 12.9.7

Solve

$$y^2 + (xy+1)\frac{dy}{dx} = 0.$$

Solution

First compute

$$M_y - N_x = 2y - y = y.$$

As this is not zero, the equation is not exact. We observe

$$Q(y) = \frac{M_y - N_x}{M} = \frac{y}{y^2} = \frac{1}{y}$$

is a function of y alone. We compute the integrating factor

$$e^{-\int Q(y)\,dy} = e^{-\ln y} = \frac{1}{y}.$$

Therefore we look at the exact equation

$$y + \frac{xy+1}{y}\frac{dy}{dx} = 0.$$

The reader should double check that this equation is exact. We follow the procedure for exact equations

$$F(x,y) = xy + A(y),$$

and

$$\frac{xy+1}{y} = x + \frac{1}{y} = x + A'(y).$$

Consequently $A'(y) = \frac{1}{y}$ or $A(y) = \ln y$. Thus $F(x,y) = xy + \ln y$. It is not possible to solve $F(x,y) = C$ for y in terms of elementary functions, so let us be content with the implicit solution:

$$xy + \ln y = C.$$

We are looking for the general solution and we divided by y above. We should check what happens when $y = 0$, as the equation itself makes perfect sense in that case. We plug in $y = 0$ to find the equation is satisfied. So $y = 0$ is also a solution.

12.9.3: Footnotes

[1] Named for the French polymath Jules Henri Poincaré (1854–1912).

12.9: Exact Equations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.8: Exact Equations** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

12.10: First Order Linear PDE

We only considered ODE so far, so let us solve a linear first order PDE. Consider the equation

$$a(x,t)\,u_x + b(x,t)\,u_t + c(x,t)\,u = g(x,t), \qquad u(x,0) = f(x), \qquad -\infty < x < \infty, \quad t > 0,$$

where $u(x,t)$ is a function of x and t. The *initial condition* $u(x,0) = f(x)$ is now a function of x rather than just a number. In these problems, it is useful to think of x as position and t as time. The equation describes the evolution of a function of x as time goes on. Below, the coefficients a, b, c, and the function g are mostly going to be constant or zero. The method we describe works with nonconstant coefficients, although the computations may get difficult quickly.

This method we use is the . The idea is that we find lines along which the equation is an ODE that we solve. We will see this technique again for second order PDE when we encounter the wave equation in Section 4.8.

Example 12.10.1

Consider the equation

$$u_t + \alpha u_x = 0, \qquad u(x,0) = f(x).$$

This particular equation, $u_t + \alpha u_x = 0$, is called the *transport equation*.

The data will propagate along curves called characteristics. The idea is to change to the so-called *characteristic coordinates*. If we change to these coordinates, the equation simplifies. The change of variables for this equation is

$$\xi = x - \alpha t, \qquad s = t.$$

Let's see what the equation becomes. Remember the chain rule in several variables.

$$\begin{aligned} u_t &= u_\xi \xi_t + u_s s_t = -\alpha u_\xi + u_s, \\ u_x &= u_\xi \xi_x + u_s s_x = u_\xi. \end{aligned} \tag{12.10.1}$$

The equation in the coordinates ξ and s becomes

$$\underbrace{(-\alpha u_\xi + u_s)}_{u_t} + \alpha \underbrace{(u_\xi)}_{u_x} = 0,$$

or in other words

$$u_s = 0.$$

That is trivial to solve. Treating ξ as simply a parameter, we have obtained the ODE $\frac{du}{ds} = 0$.

The solution is a function that does not depend on s (but it does depend on ξ). That is, there is some function A such that

$$u = A(\xi) = A(x - \alpha t).$$

The initial condition says that:

$$f(x) = u(x,0) = A(x - \alpha 0) = A(x),$$

so $A = f$. In other words,

$$u(x,t) = f(x - \alpha t).$$

Everything is simply moving right at speed α as t increases. The curve given by the equation

$$\xi = \text{constant}$$

is called the characteristic. See Figure 12.10.1 In this case, the solution does not change along the characteristic.

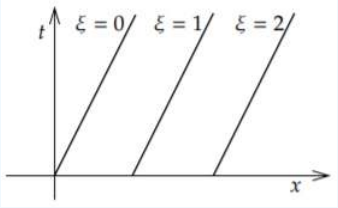


Figure 12.10.1: Characteristic curves.

In the (x,t) coordinates, the characteristic curves satisfy $t = \frac{1}{\alpha}(x-\xi)$, and are in fact lines. The slope of characteristic lines is $\frac{1}{\alpha}$, and for each different ξ we get a different characteristic line.

We see why $u_t + \alpha u_x = 0$ is called the transport equation: everything travels at some constant speed. Sometimes this is called . An example application is material being moved by a river where the material does not diffuse and is simply carried along. In this setup, x is the position along the river, t is the time, and $u(x,t)$ the concentration the material at position x and time t. See Figure 12.10.2for an example.

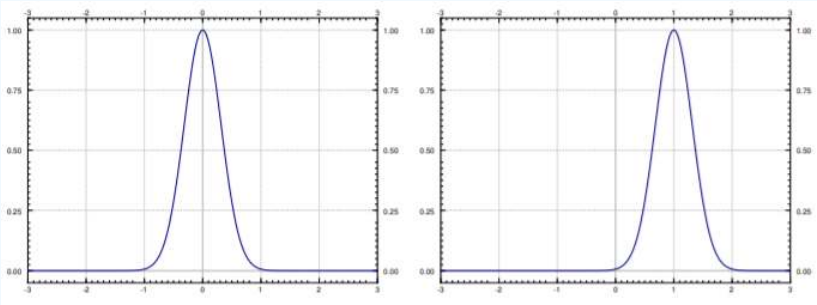


Figure 12.10.2: Example of "transport" in $u_t - u_x = 0$ (that is, $\alpha = 1$) where the initial condition $f(x)$ is a peak at the origin. On the left is a graph of the initial condition $u(x,0)$. On the right is a graph of the function $u(x,1)$, that is at time $t = 1$. Notice it is the same graph shifted one unit to the right.

We use similar idea in the more general case:

$$au_x + bu_t + cu = g, \qquad u(x,0) = f(x).$$

We change coordinates to the characteristic coordinates. Let us call these coordinates (ξ, s). These are coordinates where $au_x + bu_t$ becomes differentiation in the s variable.

Along the characteristic curves (where ξ is constant), we get a new ODE in the s variable. In the transport equation, we got the simple $\frac{du}{ds} = 0$. In general, we get the linear equation

$$\frac{du}{ds} + cu = g. \tag{12.10.2}$$

We think of everything as a function of ξ and s, although we are thinking of ξ as a parameter rather than an independent variable. So the equation is an ODE. It is a linear ODE that we can solve using the integrating factor.

To find the characteristics, think of a curve given parametrically $\big(x(s), t(s)\big)$. We try to have the curve satisfy

$$\frac{dx}{ds} = a, \qquad \frac{dt}{ds} = b.$$

Why? Because when we think of x and t as functions of s we find, using the chain rule,

$$\frac{du}{ds} + cu = \underbrace{\left(u_x \frac{dx}{ds} + u_t \frac{dt}{ds}\right)}_{\frac{du}{ds}} + cu = au_x + bu_t + cu = g.$$

So we get the ODE (12.10.2), which then describes the value of the solution u of the PDE along this characteristic curve. It is also convenient to make sure that $s = 0$ corresponds to $t = 0$, that is $t(0) = 0$. It will be convenient also for $x(0) = \xi$. See Figure 12.10.3

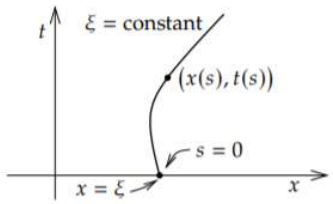


Figure 12.10.3: General characteristic curve.

Example 12.10.2

Consider

$$u_x + u_t + u = x, \qquad u(x,0) = e^{-x^2}.$$

We find the characteristics, that is, the curves given by

$$\frac{dx}{ds} = 1, \qquad \frac{dt}{ds} = 1.$$

So

$$x = s + c_1, \qquad t = s + c_2,$$

for some c_1 and c_2. At $s = 0$ we want $t = 0$, and x should be ξ. So we let $c_1 = \xi$ and $c_2 = 0$:

$$x = s + \xi, \qquad t = s.$$

The ODE is $\frac{du}{ds} + u = x$, and $x = s + \xi$. So, the ODE to solve along the characteristic is

$$\frac{du}{ds} + u = s + \xi.$$

The general solution of this equation, treating ξ as a parameter, is $u = Ce^{-s} + s + \xi - 1$, for some constant C. At $s = 0$, our initial condition is that u is $e^{-\xi^2}$, since at $s = 0$ we have $x = \xi$. Given this initial condition, we find $C = e^{-\xi^2} - \xi + 1$. So,

$$\begin{aligned} u &= \left(e^{-\xi^2} - \xi + 1\right)e^{-s} + s + \xi - 1 \\ &= e^{-\xi^2 - s} + (1 - \xi)e^{-s} + s + \xi - 1. \end{aligned} \tag{12.10.3}$$

Substitute $\xi = x - t$ and $s = t$ to find u in terms of x and t:

$$\begin{aligned} u &= e^{-\xi^2 - s} + (1 - \xi)e^{-s} + s + \xi - 1 \\ &= e^{-(x-t)^2 - t} + (1 - x + t)e^{-t} + x - 1. \end{aligned} \tag{12.10.4}$$

See Figure 12.10.4for a plot of $u(x,t)$ as a function of two variables.

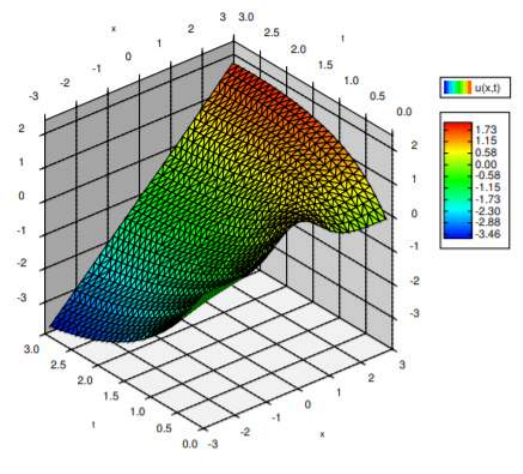


Figure 12.10.4: Plot of the solution $u(x,t)$ to $u_x + u_t + u = x$, $u(x,0) = e^{-x^2}$.

When the coefficients are not constants, the characteristic curves are not going to be straight lines anymore.

Example 12.10.3

Consider the following variable coefficient equation:

$$xu_x + u_t + 2u = 0, \qquad u(x,0) = \cos(x). \tag{12.10.5}$$

We find the characteristics, that is, the curves given by

$$\frac{dx}{ds} = x, \qquad \frac{dt}{ds} = 1.$$

So

$$x = c_1 e^s, \qquad t = s + c_2.$$

At $s = 0$, we wish to get the line $t = 0$, and x should be ξ. So

$$x = \xi e^s, \qquad t = s.$$

OK, the ODE we need to solve is

$$\frac{du}{ds} + 2u = 0.$$

This is for a fixed ξ. At $s = 0$, we should get that u is $\cos(\xi)$, so that is our initial condition. Consequently,

$$u = e^{-2s}\cos(\xi) = e^{-2t}\cos(xe^{-t}).$$

We make a few closing remarks. One thing to keep in mind is that we would get into trouble if the coefficient in front of u_t, that is the b, is ever zero. Let us consider a quick example of what can go wrong:

$$u_x + u = 0, \qquad u(x,0) = \sin(x).$$

This problem has no solution. If we had a solution, it would imply that $u_x(x,0) = \cos(x)$, but $u_x(x,0) + u(x,0) = \cos(x) + \sin(x) \neq 0$. The problem is that the characteristic curve is now the line $t = 0$, and the solution is already provided on that line!

As long as b is nonzero, it is convenient to ensure that b is positive by multiplying by -1 if necessary, so that positive s means positive t.

Another remark is that if a or b in the equation are variable, the computations can quickly get out of hand, as the expressions for the characteristic coordinates become messy and then solving the ODE becomes even messier. In the examples above, b was

always 1, meaning we got $s = t$ in the characteristic coordinates. If b is not constant, your expression for s will be more complicated.

Finding the characteristic coordinates is really a system of ODE in general if a depends on t or if b depends on x. In that case, we would need techniques of systems of ODE to solve, see Chapter 3 or Chapter 8. In general, if a and b are not linear functions or constants, finding closed form expressions for the characteristic coordinates may be impossible.

Finally, the method of characteristics applies to nonlinear first order PDE as well. In the nonlinear case, the characteristics depend not only on the differential equation, but also on the initial data. This leads to not only more difficult computations, but also the formation of singularities where the solution breaks down at a certain point in time. An example application where first order nonlinear PDE come up is traffic flow theory, and you have probably experienced the formation of singularities: traffic jams. But we digress.

12.10: First Order Linear PDE is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **1.9: First Order Linear PDE** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

12.E: First order ODEs (Exercises)

These are homework exercises to accompany Libl's "Differential Equations for Engineering" Textmap. This is a textbook targeted for a one semester first course on differential equations, aimed at engineering students. Prerequisite for the course is the basic calculus sequence.

12.E.1: 1.1: Integrals as solutions

Exercise 12.E.1.1.1

Solve for v, and then solve for x. Find $x(10)$ to answer the question.

Exercise 12.E.1.1.2

Solve $\dfrac{dy}{dx} = x^2 + x$ for $y(1) = 3$.

Exercise 12.E.1.1.3

Solve $\dfrac{dy}{dx} = \sin(5x)$ for $y(0) = 2$.

Exercise 12.E.1.1.4

Solve $\dfrac{dy}{dx} = \dfrac{1}{x^2 - 1}$ for $y(0) = 0$.

Exercise 12.E.1.1.5

Solve $y' = y^3$ for $y(0) = 1$.

Exercise 12.E.1.1.6: (little harder)

Solve $y' = (y-1)(y+1)$ for $y(0) = 3$.

Exercise 12.E.1.1.7

Solve $\dfrac{dy}{dx} = \dfrac{1}{y+1}$ for $y(0) = 0$.

Exercise 12.E.1.1.8: (harder)

Solve $y'' = \sin x$ for $y(0) = 0$, $y'(0) = 2$.

Exercise 12.E.1.1.9

A spaceship is traveling at the speed $2t^2 + 1$ $^{\text{km}}/_{\text{s}}$ (t is time in seconds). It is pointing directly away from earth and at time $t = 0$ it is 1000 kilometers from earth. How far from earth is it at one minute from time $t = 0$?

Exercise 12.E.1.1.10

Solve $\dfrac{dx}{dt} = \sin(t^2) + t$, $x(0) = 20$. It is OK to leave your answer as a definite integral.

Exercise 12.E.1.1.11

A dropped ball accelerates downwards at a constant rate 9.8 meters per second squared. Set up the differential equation for the height above ground h in meters. Then supposing $h(0) = 100$ meters, how long does it take for the ball to hit the ground.

Exercise 12.E.1.1.12

Find the general solution of $y' = e^x$, and then $y' = e^y$.

Exercise 12.E.1.1.13

Solve $\dfrac{dy}{dx} = e^x + x$ and $y(0) = 10$.

Answer

$y = e^x + \frac{x^2}{2} + 9$

Exercise 12.E.1.1.14

Solve $x' = \dfrac{1}{x^2}$, $x(1) = 1$.

Answer

$x = (3t - 2)^{1/3}$

Exercise 12.E.1.1.15

Solve $x' = \frac{1}{\cos(x)}$, $x(0) = \frac{\pi}{2}$.

Answer

A$x = \sin^{-1}(t + 1)$

Exercise 12.E.1.1.16

Sid is in a car traveling at speed $10t + 70$ miles per hour away from Las Vegas, where t is in hours. At $t = 0$ the Sid is 10 miles away from Vegas. How far from Vegas is Sid 2 hours later?

Answer

170

Exercise 12.E.1.1.17

Solve $y' = y''$, $y(0) = 1$, where n is a positive integer. Hint: You have to consider different cases.

Answer

If $n \neq 1$, then $y = ((1 - n)x + 1)^{1/(1-n)}$. If $n = 1$, then $y = e^x$.

Exercise 12.E.1.1.18

The rate of change of the volume of a snowball that is melting is proportional to the surface area of the snowball. Suppose the snowball is perfectly spherical. The volume (in centimeters cubed) of a ball of radius r centimeters is $\frac{4}{3})\pi r^3$. The surface area is $4\pi r^2$. Set up the differential equation for how the radius r is changing. Then, suppose that at time $t = 0$ minutes, the radius is 10 centimeters. After 5 minutes, the radius is 8 centimeters. At what time t will the snowball be completely melted?

Answer

The equation is $r' = -C$ for some constant C. The snowball will be completely melted in 25 minutes from time $t = 0$.

Exercise 12.E.1.1.19

Find the general solution to $y'''' = 0$. How many distinct constants do you need?

Answer

$y = Ax^3 + Bx^2 + Cx + D$, so 4 constants.

12.E.2: 1.2: Slope fields

Exercise 12.E.1.2.1

Sketch slope field for $y' = e^{x-y}$. How do the solutions behave as x grows? Can you guess a particular solution by looking at the slope field?

Exercise 12.E.1.2.2

Sketch slope field for $y' = x^2$.

Exercise 12.E.1.2.3

Sketch slope field for $y' = y^2$.

Exercise 12.E.1.2.4

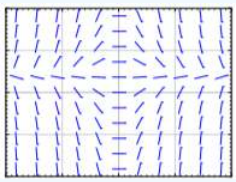
Is it possible to solve the equation $y' = \frac{xy}{\cos x}$ for $y(0) = 1$? Justify.

Exercise 12.E.1.2.5

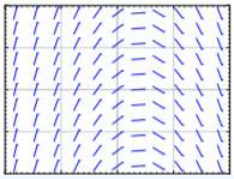
Is it possible to solve the equation $y' = y\sqrt{|x|}$ for $y(0) = 0$? Is the solution unique? Justify.

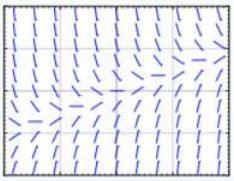
Exercise 12.E.1.2.6

Match equations $y' = 1 - x$, $y' = x - 2y$, $y' = x(1 - y)$ to slope fields. Justify.

a. 

https://math.libretexts.org/@go/page/128064

b. 

c. 

Exercise 12.E.1.2.7: (challenging)

Take $y' = f(x, y)$, $y(0) = 0$, where $f(x, y) > 1$ for all x and y. If the solution exists for all x, can you say what happens to $y(x)$ as x goes to positive infinity? Explain.

Exercise 12.E.1.2.8: (challenging)

Take $(y - x)y' = 0$, $x(0) = 0$.

a. Find two distinct solutions.
b. Explain why this does not violate Picard's theorem.

Exercise 12.E.1.2.9

Suppose $y' = f(x, y)$. What will the slope field look like, explain and sketch an example, if you know the following about $f(x, y)$:

a. f does not depend on y.
b. f does not depend on x.
c. $f, (t, t) = 0$ for any number t.
d. $f(x, 0) = 0$ and $f(x, 1) = 1$ for all x.

Exercise 12.E.1.2.10

Find a solution to $y' = |y|$, $y(0) = 0$. Does Picard's theorem apply?

Exercise 12.E.1.2.11

Take an equation $y' = (y - 2x)g(x, y) + 2$ for some function $g(x, y)$. Can you solve the problem for the initial condition $y(0) = 0$, and if so what is the solution?

Exercise 12.E.1.2.12: (challenging)

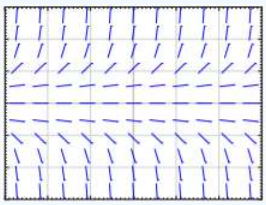
Suppose $y' = f(x, y)$ is such that $f(x, 1) = 0$ for every x, f is continuous and $\frac{\partial f}{\partial y}$ exists and is continuous for every x and y.

a. Guess a solution given the initial condition $y(0) = 1$.
b. Can graphs of two solutions of the equation for different initial conditions ever intersect?
c. Given $y(0) = 0$, what can you say about the solution. In particular, can $y(x) > 1$ for any x? Can $y(x) = 1$ for any x? Why or why not?

Exercise 12.E.1.2.13

Sketch the slope field of $y' = y^3$. Can you visually find the solution that satisfies $y(0) = 0$?

Answer



$y = 0$ is a solution such that $y(0) = 0$

Exercise 12.E.1.2.14

Is it possible to solve $y' = xy$ for $y(0) = 0$? Is the solution unique?

Answer

Yes a solution exists. The equation is $y' = f(x, y)$ where $f(x, y) = xy$. The function $f(x, y)$ is continuous and $\frac{\partial f}{\partial y} = x$, which is also continuous near $(0, 0)$. So a solution exists and is unique. (In fact, $y = 0$ is the solution.)

Exercise 12.E.1.2.15

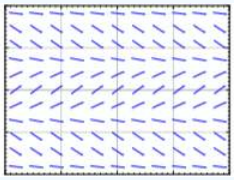
Is it possible to solve $y' = \frac{x}{x^2-1}$ for $y(1) = 0$?

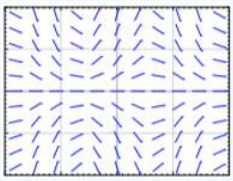
Answer

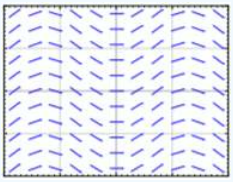
No, the equation is not defined at $(x, y) = (1, 0)$.

Exercise 12.E.1.2.16

Match equations $y' = \sin x$, $y' = \cos y$, $y' = y\cos(x)$ to slope fields. Justify.

a. 

b. 

c. 

Answer

a. $y' = \cos y$

https://math.libretexts.org/@go/page/128064

b. $y' = y\cos(x)$
c. $y' = \sin x$

Justification left to reader.

Exercise 12.E.1.2.17: (tricky)

Suppose

$$f(y) = \begin{cases} 0 & if\ y > 0, \\ 1 & if\ y \leq 0. \end{cases}$$

Does $y' = f(y)$, $y(0) = 0$ have a continuously differentiable solution? Does Picard apply? Why, or why not?

Answer

Picard does not apply as f is not continuous at $y = 0$. The equation does not have a continuously differentiable solution. Suppose it did. Notice that $y'(0) = 1$. By the first derivative test, $y(x) > 0$ for small positive x. But then for those x we would have $y'(x) = 0$, so clearly the derivative cannot be continuous.

Exercise 12.E.1.2.18

Consider an equation of the form $y' = f(x)$ for some continuous function f, and an initial condition $y(x_0) = y_0$. Does a solution exist for all x? Why or why not?

Answer

The solution is $y(x) = \int_{x_0}^{x} f(s)ds + y_0$, and this does indeed exist for every x.

12.E.3: 1.3: Separable Equations

Exercise 12.E.1.3.1

Solve $y' = \frac{x}{y}$.

Exercise 12.E.1.3.2

Solve $y' = x^2 y$.

Exercise 12.E.1.3.3

Solve $\frac{dx}{dt} = \left(x^2 - 1\right)$, for $x(0) = 0$.

Exercise 12.E.1.3.4

Solve $\frac{dx}{dt} = x\sin(t)$, for $x(0) = 1$.

Exercise 12.E.1.3.5

Solve $\frac{dy}{dx} = xy + x + y + 1$. Hint: Factor the right hand side.

Exercise 12.E.1.3.6

Solve $xy' = y + 2x^2y$, where $y(1) = 1$.

Exercise 12.E.1.3.7

Solve $\frac{dy}{dx} = \frac{y^2+1}{x^2+1}$, for $y(0) = 1$.

Exercise 12.E.1.3.8

Find an implicit solution for $\frac{dy}{dx} = \frac{x^2+1}{y^2+1}$, for $y(0) = 1$.

Exercise 12.E.1.3.9

Find an explicit solution for $y' = xe^{-y}$, $y(0) = 1$.

Exercise 12.E.1.3.10

Find an explicit solution for $xy' = e^{-y}$, for $y(1) = 1$.

Exercise 12.E.1.3.11

Find an explicit solution for $y' = ye^{-x^2}$, $y(0) = 1$. It is alright to leave a definite integral in your answer.

Exercise 12.E.1.3.12

Suppose a cup of coffee is at 100 degrees Celsius at time $t = 0$, it is at 70 degrees at $t = 10$ minutes, and it is at 50 degrees at $t = 20$ minutes. Compute the ambient temperature.

Exercise 12.E.1.3.13

Solve $y' = 2xy$.

Answer

$y = Ce^{x^2}$

Exercise 12.E.1.3.14

Solve $x' = 3xt^2 - 3t^2$, $x(0) = 2$.

Answer

$y = e^{t^3} + 1$

Exercise 12.E.1.3.15

Find an implicit solution for $x' = \frac{1}{3x^2+1}$ $x(0) = 1$.

Answer

$x^3 + x = t + 2$

Exercise 12.E.1.3.16

Find an explicit solution to $xy' = y^2$, $y(1) = 1$.

Answer

$y = \frac{1}{1-\ln x}$

Exercise 12.E.1.3.17

Find an implicit solution to $y' = \frac{\sin(x)}{\cos(y)}$.

Answer

$\sin(y) = -\cos(x) + C$

Exercise 12.E.1.3.18

Take Example 1.3.3 with the same numbers: 89 degrees at $t = 0$, 85 degrees at $t = 1$, and ambient temperature of 22 degrees. Suppose these temperatures were measured with precision of ± 0.5 degrees. Given this imprecision, the time it takes the coffee to cool to (exactly) 60 degrees is also only known in a certain range. Find this range. Hint: Think about what kind of error makes the cooling time longer and what shorter.

Answer

The range is approximately 7.45 to 12.15 minutes.

Exercise 12.E.1.3.19

A population x of rabbits on an island is modeled by $x' = x - (\frac{1}{1000})x^2$, where the independent variable is time in months. At time $t = 0$, there are 40 rabbits on the island.

a. Find the solution to the equation with the initial condition.
b. How many rabbits are on the island in 1 month, 5 months, 10 months, 15 months (round to the nearest integer)

Answer

a. $x = \frac{1000e^t}{e^t+24}$.
b. 102 rabbits after one month, 861 after 5 months, 999 after 10 months, 1000 after 15 months.

12.E.4: 1.4: Linear equations and the integrating factor

In the exercises, feel free to leave answer as a definite integral if a closed form solution cannot be found. If you can find a closed form solution, you should give that.

Exercise 12.E.1.4.1

Solve $y' + xy = x$.

Exercise 12.E.1.4.2

Solve $y' + 6y = e^x$.

Exercise 12.E.1.4.3

Solve $y' + 3x^2 y = \sin(x)\, e^{-x^3}$ with $y(0) = 1$.

Exercise 12.E.1.4.4

Solve $y' + \cos(x)\, y = \cos(x)$.

Exercise 12.E.1.4.5

Solve $\frac{1}{x^2+1} y' + xy = 3$ with $y(0) = 0$.

Exercise 12.E.1.4.6

Suppose there are two lakes located on a stream. Clean water flows into the first lake, then the water from the first lake flows into the second lake, and then water from the second lake flows further downstream. The in and out flow from each lake is 500 liters per hour. The first lake contains 100 thousand liters of water and the second lake contains 200 thousand liters of water. A truck with 500 kg of toxic substance crashes into the first lake. Assume that the water is being continually mixed perfectly by the stream.

a. Find the concentration of toxic substance as a function of time in both lakes.
b. When will the concentration in the first lake be below 0.001 kg per liter?
c. When will the concentration in the second lake be maximal?

Exercise 12.E.1.4.7

Newton's law of cooling states that $\frac{dx}{dt} = -k(x - A)$ where x is the temperature, t is time, A is the ambient temperature, and $k > 0$ is a constant. Suppose that $A = A_0 \cos(\omega t)$ for some constants A_0 and ω. That is, the ambient temperature oscillates (for example night and day temperatures).

a. Find the general solution.
b. In the long term, will the initial conditions make much of a difference? Why or why not?

Exercise 12.E.1.4.8

Initially 5 grams of salt are dissolved in 20 liters of water. Brine with concentration of salt 2 grams of salt per liter is added at a rate of 3 liters a minute. The tank is mixed well and is drained at 3 liters a minute. How long does the process have to continue until there are 20 grams of salt in the tank?

Exercise 12.E.1.4.9

Initially a tank contains 10 liters of pure water. Brine of unknown (but constant) concentration of salt is flowing in at 1 liter per minute. The water is mixed well and drained at 1 liter per minute. In 20 minutes there are 15 grams of salt in the tank. What is the concentration of salt in the incoming brine?

Exercise 12.E.1.4.10

Solve $y' + 3x^2 y + x^2$.

Answer

$y = Ce^{-x^3} + \frac{1}{3}$

Exercise 12.E.1.4.11

Solve $y' + 2\sin(2x)y = 2\sin(2x)$ with $y(\pi/2) = 3$.

Answer

$y = 2e^{\cos(2x)+1} + 1$

Exercise 12.E.1.4.12

Suppose a water tank is being pumped out at $3\ \frac{\text{L}}{\text{min}}$. The water tank starts at $10\ \text{L}$ of clean water. Water with toxic substance is flowing into the tank at $2\ \frac{\text{L}}{\text{min}}$, with concentration $20t\ \frac{\text{g}}{\text{L}}$ at time t. When the tank is half empty, how many grams of toxic substance are in the tank (assuming perfect mixing)?

Answer

$250\,\text{grams}$

Exercise 12.E.1.4.13

Suppose we have bacteria on a plate and suppose that we are slowly adding a toxic substance such that the rate of growth is slowing down. That is, suppose that $\frac{dP}{dt} = (2 - 0.1\,t)P$. If $P(0) = 1000$, find the population at $t = 5$.

Answer

$P(5) = 1000e^{2\times 5 - 0.05\times 5^2} = 1000e^{8.75} \approx 6.31 \times 10^6$

Exercise 12.E.1.4.14

A cylindrical water tank has water flowing in at I cubic meters per second. Let A be the area of the cross section of the tank in meters. Suppose water is flowing from the bottom of the tank at a rate proportional to the height of the water level. Set up the differential equation for h, the height of the water, introducing and naming constants that you need. You should also give the units for your constants.

Answer

$Ah' = I - kh$, where k is a constant with units $\frac{\text{m}^2}{\text{s}}$.

12.E.5: 1.5: Substitution

Hint: Answers need not always be in closed form.

Exercise 12.E.1.5.1

Solve $y' + y(x^2 - 1) + xy^6 = 0$, with $y(1) = 1$.

Exercise 12.E.1.5.2

Solve $2yy' + 1 = y^2 + x$, with $y(0) = 1$.

Exercise 12.E.1.5.3

Solve $y' + xy = y^4$, with $y(0) = 1$.

Exercise 12.E.1.5.4

Solve $yy' + x = \sqrt{x^2 + y^2}$.

Exercise 12.E.1.5.5

Solve $y' = (x + y - 1)^2$.

Exercise 12.E.1.5.6

Solve $y' = \dfrac{x^2 - y^2}{xy}$, with $y(1) = 2$.

Exercise 12.E.1.5.7

Solve $xy' + y + y^2 = 0$, $y(1) = 2$.

Answer

$y = \frac{2}{3x-2}$

Exercise 12.E.1.5.8

Solve $xy' + y + x = 0$, $y(1) = 1$.

Answer

$y = \frac{3-x^2}{2x}$

Exercise 12.E.1.5.9

Solve $y^2 y' = y^3 - 3x$, $y(0) = 2$.

Answer

$y = \left(7e^{3x} + 3x + 1\right)^{1/3}$

Exercise 12.E.1.5.10

Solve $2yy' = e^{y^2 - x^2} + 2x$.

Answer

$y = \sqrt{x^2 - \ln(C - x)}$

12.E.6: 1.6: Autonomous equations

Exercise 12.E.1.6.1

Consider $x' = x^2$.

a. Draw the phase diagram, find the critical points and mark them stable or unstable.
b. Sketch typical solutions of the equation.
c. Find $\lim_{t \to \infty} x(t)$ for the solution with the initial condition $x(0) = -1$.

Exercise 12.E.1.6.2

Let $x' = \sin x$.

a. Draw the phase diagram for $-4\pi \leq x \leq 4\pi$. On this interval mark the critical points stable or unstable.
b. Sketch typical solutions of the equation.
c. Find $\lim_{t\to\infty} x(t)$ for the solution with the initial condition $x(0) = 1$.

Exercise 12.E.1.6.3

Suppose $f(x)$ is positive for $0 < x < 1$, it is zero when $x = 0$ and $x = 1$, and it is negative for all other x.

a. Draw the phase diagram for $x' = f(x)$, find the critical points and mark them stable or unstable.
b. Sketch typical solutions of the equation.
c. Find $\lim_{t\to\infty} x(t)$ for the solution with the initial condition $x(0) = 0.5$.

Exercise 12.E.1.6.4

Start with the logistic equation $\frac{dx}{dt} = kx(M - x)$. Suppose that we modify our harvesting. That is we will only harvest an amount proportional to current population. In other words we harvest hx per unit of time for some $h > 0$ (Similar to earlier example with h replaced with hx).

a. Construct the differential equation.
b. Show that if $kM > h$, then the equation is still logistic.
c. What happens when $kM < h$?

Exercise 12.E.1.6.5

A disease is spreading through the country. Let x be the number of people infected. Let the constant S be the number of people susceptible to infection. The infection rate $\frac{dx}{dt}$ is proportional to the product of already infected people, x, and the number of susceptible but uninfected people, $S - x$.

a. Write down the differential equation.
b. Supposing $x(0) > 0$, that is, some people are infected at time $t = 0$, what is $\lim_{t\to\infty} x(t)$.
c. Does the solution to part b) agree with your intuition? Why or why not?

Exercise 12.E.1.6.6

Let $x' = (x-1)(x-2)x^2$.

a. Sketch the phase diagram and find critical points.
b. Classify the critical points.
c. If $x(0) = 0.5$ then find $\lim_{t\to\infty} x(t)$.

Answer

a. $0,\ 1,\ 2$ are critical points.
b. $x = 0$ is unstable (semistable), $x = 1$ is stable, and $x = 2$ is unstable.
c. 1

Exercise 12.E.1.6.7

Let $x' = e^{-x}$.

a. Find and classify all critical points.
b. Find $\lim_{t\to\infty} x(t)$ given any initial condition.

Answer

a. There are no critical points.
b. ∞

Exercise 12.E.1.6.8

Assume that a population of fish in a lake satisfies $\frac{dx}{dt} = kx(M-x)$. Now suppose that fish are continually added at A fish per unit of time.

a. Find the differential equation for x.
b. What is the new limiting population?

Answer

a. $\frac{dx}{dt} = kx(M-x) + A$
b. $\frac{kM + \sqrt{(kM)^2 + 4Ak}}{2k}$

Exercise 12.E.1.6.9

Suppose $\frac{dx}{dt} = (x-\alpha)(x-\beta)$ for two numbers $\alpha < \beta$.

a. Find the critical points, and classify them.

For b), c), d), find $\lim_{t\to\infty} x(t)$ based on the phase diagram.

b. $x(0) < \alpha$,
c. $\alpha < x(0) < \beta$,
d. $\beta < x(0)$.

Answer

1. α is a stable critical point, β is an unstable one.
2. α
3. α
4. ∞ or DNE.

12.E.7: 1.7: Numerical methods: Euler's method

Exercise 12.E.1.7.1

Consider $\frac{dx}{dt} = (2t-x)^2$, $x(0) = 2$. Use Euler's method with step size $h = 0.5$ to approximate $x(1)$.

Exercise 12.E.1.7.2

Consider $\frac{dx}{dt} = t - x$, $x(0) = 1$.

a. Use Euler's method with step sizes $h = 1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$ to approximate $x(1)$.
b. Solve the equation exactly.
c. Describe what happens to the errors for each h you used. That is, find the factor by which the error changed each time you halved the interval.

Exercise 12.E.1.7.3

Approximate the value of e by looking at the initial value problem $y' = y$ with $y(0) = 1$ and approximating $y(1)$ using Euler's method with a step size of 0.2.

Exercise 12.E.1.7.4

Example of numerical instability: Take $y' = -5y$, $y(0) = 1$. We know that the solution should decay to zero as x grows. Using Euler's method, start with $h = 1$ and compute y_1, y_2, y_3, y_4 to try to approximate $y(4)$. What happened? Now halve the interval. Keep halving the interval and approximating $y(4)$ until the numbers you are getting start to stabilize (that is, until they start going towards zero). Note: You might want to use a calculator.

The simplest method used in practice is the Runge-Kutta method. Consider $\frac{dy}{dx} = f(x,y)$, $y(x_0) = y_0$ and a step size h. Everything is the same as in Euler's method, except the computation of y_{i+1} and x_{i+1}.

$$\begin{aligned} k_1 &= f(x_i, y_i), \\ k_2 &= f(x_i + \frac{h}{2}, y_i + k_1\frac{h}{2}) \qquad & x_{i+1} &= x_i + h, \\ k_3 &= f(x_i + \frac{h}{2}, y_i + k_2\frac{h}{2}) \qquad & y_{i+1} &= y_i + \frac{k_1 + 2k_2 + 2k_3 + k_4}{6}h, \\ k_4 &= f(x_i + h, y_i + k_3 h). \end{aligned} \tag{12.E.1}$$

Exercise 12.E.1.7.5

Consider $\frac{dy}{dx} = yx^2$, $y(0) = 1$.

a. Use Runge-Kutta (see above) with step sizes $h = 1$ and $h = \frac{1}{2}$ to approximate $y(1)$.
b. Use Euler's method with $h = 1$ and $h = \frac{1}{2}$.
c. Solve exactly, find the exact value of $y(1)$, and compare.

Exercise 12.E.1.7.6

Let $x' = \sin(xt)$, and $x(0) = 1$. Approximate $x(1)$ using Euler's method with step sizes 1, 0.5, 0.25. Use a calculator and compute up to 4 decimal digits.

Answer

Approximately: $1.0000,\ 1.2397,\ 1.382$

Exercise 12.E.1.7.7

Let $x' = 2t$, and $x(0) = 0$.

a. Approximate $x(4)$ using Euler's method with step sizes 4, 2, and 1.
b. Solve exactly, and compute the errors.
c. Compute the factor by which the errors changed.

Answer

a. $0,\ 8,\ 12$
b. $x(4) = 16$, so errors are: $16,\ 8,\ 4$
c. Factors are $0.5,\ 0.5,\ 0.5$

Exercise 12.E.1.7.8

Let $x' = xe^{xt+1}$, and $x(0) = 0$.

a. Approximate $x(4)$ using Euler's method with step sizes 4, 2, and 1.
b. Guess an exact solution based on part a) and compute the errors.

Answer

a. $0,\ 0,\ 0$

b. $x=0$ is a solution so errors are: $0,\ 0,\ 0$.

There is a simple way to improve Euler's method to make it a second order method by doing just one extra step. Consider $\frac{dy}{dx}=f(x,y)$, $y(x_0)=y_0$, and a step size h. What we do is to pretend we compute the next step as in Euler, that is, we start with (x_i,y_i), we compute a slope $k_1=f(x_i,y_i)$, and then look at the point (x_i+h,y_i+k_1h). Instead of letting our new point be (x_i+h,y_i+k_1h), we compute the slope at that point, call it k_2, and then take the average of k_1 and k_2, hoping that the average is going to be closer to the actual slope on the interval from x_i to x_i+h. And we are correct, if we halve the step, the error should go down by a factor of $2^2=4$. To summarize, the setup is the same as for regular Euler, except the computation of y_{i+1} and x_{i+1}.

$$\begin{aligned} k_1&=f(x_i,y_i), & x_{i+1}&=x_i+h,\\ k_2&=f(x_i+h,y_i+k_1h), & y_{i+1}&=y_i+\frac{k_1+k_2}{2}\,h. \end{aligned} \tag{12.E.2}$$

? Exercise 12.E.1.7.9

Consider $\dfrac{dy}{dx}=x+y$, $y(0)=1$.

a. Use the improved Euler's method (see above) with step sizes $h=\frac{1}{4}$ and $h=\frac{1}{8}$ to approximate $y(1)$.
b. Use Euler's method with $h=\frac{1}{4}$ and $h=\frac{1}{8}$.
c. Solve exactly, find the exact value of $y(1)$.
d. Compute the errors, and the factors by which the errors changed.

Answer

a. Improved Euler: $y(1)\approx 3.3897$ for $h=1/4$, $y(1)\approx 3.4237$ for $h=1/8$,
b. Standard Euler: $y(1)\approx 2.8828$ for $h=1/4$, $y(1)\approx 3.1316$ for $h=1/8$,
c. $y=2e^x-x-1$, so $y(2)$ is approximately 3.4366
d. Approximate errors for improved Euler: 0.046852 for $h=1/4$, and 0.012881 for $h=1/8$. For standard Euler: 0.55375 for $h=1/4$, and 0.30499 for $h=1/8$. Factor is approximately 0.27 for improved Euler, and 0.55 for standard Euler.

12.E.8: 1.8 Exact Equations

? Exercise 12.E.1.8.1

Solve the following exact equations, implicit general solutions will suffice:

a. $(2xy+x^2)\,dx+(x^2+y^2+1)\,dy=0$
b. $x^5+y^5\frac{dy}{dx}=0$
c. $e^x+y^3+3xy^2\frac{dy}{dx}=0$
d. $(x+y)\cos(x)+\sin(x)+\sin(x)y'=0$

? Exercise 12.E.1.8.2

Find the integrating factor for the following equations making them into exact equations:

a. $e^{xy}\,dx+\frac{y}{x}e^{xy}\,dy=0$
b. $\frac{e^x+y^3}{y^2}\,dx+3x\,dy=0$
c. $4(y^2+x)\,dx+\frac{2x+2y^2}{y}\,dy=0$
d. $2\sin(y)\,dx+x\cos(y)\,dy=0$

https://math.libretexts.org/@go/page/128064

Exercise 12.E.1.8.3

Suppose you have an equation of the form: $f(x)+g(y)\frac{dy}{dx}=0$.

a. Show it is exact.
b. Find the form of the potential function in terms of f and g.

Exercise 12.E.1.8.4

Suppose that we have the equation $f(x)\,dx-dy=0$.

a. Is this equation exact?
b. Find the general solution using a definite integral.

Exercise 12.E.1.8.5

Find the potential function $F(x,y)$ of the exact equation $\frac{1+xy}{x}\,dx+\left(\frac{1}{y}+x\right)dy=0$ in two different ways.

a. Integrate M in terms of x and then differentiate in y and set to N.
b. Integrate N in terms of y and then differentiate in x and set to M.

Exercise 12.E.1.8.6

A function $u(x,y)$ is said to be a if $u_{xx}+u_{yy}=0$.

a. Show if u is harmonic, $-u_y\,dx+u_x\,dy=0$ is an exact equation. So there exists (at least locally) the so-called function $v(x,y)$ such that $v_x=-u_y$ and $v_y=u_x$.

Verify that the following u are harmonic and find the corresponding harmonic conjugates v:

a. $u=2xy$
b. $u=e^x\cos y$
c. $u=x^3-3xy^2$

Exercise 12.E.1.8.7

Solve the following exact equations, implicit general solutions will suffice:

a. $\cos(x)+ye^{xy}+xe^{xy}y'=0$
b. $(2x+y)\,dx+(x-4y)\,dy=0$
c. $e^x+e^y\frac{dy}{dx}=0$
d. $(3x^2+3y)\,dx+(3y^2+3x)\,dy=0$

Answer

a. $e^{xy}+\sin(x)=C$
b. $x^2+xy-2y^2=C$
c. $e^x+e^y=C$
d. $x^3+3xy+y^3=C$

Exercise 12.E.1.8.8

Find the integrating factor for the following equations making them into exact equations:

a. $\frac{1}{y}\,dx+3y\,dy=0$
b. $dx-e^{-x-y}\,dy=0$
c. $\left(\frac{\cos(x)}{y^2}+\frac{1}{y}\right)dx+\frac{x}{y^2}\,dy=0$

d. $\left(2y+\frac{y^2}{x}\right)dx+(2y+x)\,dy=0$

Answer

a. Integrating factor is y, equation becomes $dx+3y^2dy=0$.
b. Integrating factor is e^x, equation becomes $e^xdx-e^{-y}dy=0$.
c. Integrating factor is y^2, equation becomes $(\cos(x)+y)dx+xdy=0$.
d. Integrating factor is x, equation becomes $(2xy+y^2)dx+(x^2+2xy)dy=0$.

Exercise 12.E.1.8.9

a. Show that every separable equation $y'=f(x)g(y)$ can be written as an exact equation, and verify that it is indeed exact.
b. Using this rewrite $y'=xy$ as an exact equation, solve it and verify that the solution is the same as it was in Example 1.3.1.

Answer

a. The equation is $-f(x)dx+\frac{1}{g(y)}dy$, and this is exact because $M=-f(x)$, $N=\frac{1}{g(y)}$, so $M_y=0=N_x$.
b. $-xdx+\frac{1}{y}dy=0$, leads to potential function $F(x,y)=-\frac{x^2}{2}+\ln|y|$, solving $F(x,y)=C$ leads to the same solution as the example.

12.E.9: 1.9: First Order Linear PDE

Exercise 12.E.1.9.1

Solve

a. $u_t+9u_x=0$, $u(x,0)=\sin(x)$,
b. $u_t-8u_x=0$, $u(x,0)=\sin(x)$,
c. $u_t+\pi u_x=0$, $u(x,0)=\sin(x)$,
d. $u_t+\pi u_x+u=0$, $u(x,0)=\sin(x)$.

Exercise 12.E.1.9.2

Solve $u_t+3u_x=1$, $u(x,0)=x^2$.

Exercise 12.E.1.9.3

Solve $u_t+3u_x=x$, $u(x,0)=e^x$.

Exercise 12.E.1.9.4

Solve $u_x+u_t+xu=0$, $u(x,0)=\cos(x)$.

Exercise 12.E.1.9.5

a. Find the characteristic coordinates for the following equations:
$u_x+u_t+u=1$, $u(x,0)=\cos(x)$,) $2u_x+2u_t+2u=2$, $u(x,0)=\cos(x)$.
b. Solve the two equations using the coordinates.
c. Explain why you got the same solution, although the characteristic coordinates you found were different.

Exercise 12.E.1.9.6

Solve $(1+x^2)u_t+x^2u_x+e^xu=0$, $u(x,0)=0$. Hint: Think a little out of the box.

Exercise 12.E.1.9.7

Solve

a. $u_t - 5u_x = 0$, $u(x,0) = \frac{1}{1+x^2}$,
b. $u_t + 2u_x = 0$, $u(x,0) = \cos(x)$.

Answer

a. $u = \frac{1}{1+(x+5t)^2}$
b. $u = \cos(x-2t)$

Exercise 12.E.1.9.8

Solve $u_x + u_t + tu = 0$, $u(x,0) = \cos(x)$.

Answer

$u = \cos(x-t)e^{-t^2/2}$

Exercise 12.E.1.9.9

Solve $u_x + u_t = 5$, $u(x,0) = x$.

Answer

$u = x + 4t$

This page titled 12.E: First order ODEs (Exercises) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **1.E: First order ODEs (Exercises)** has no license indicated.

CHAPTER OVERVIEW

13: Applications of First Order Equations

In this chapter, we consider applications of first order differential equations.

13.1: Growth and Decay

13.1E: Growth and Decay (Exercises)

13.2: Cooling and Mixing

13.2E: Cooling and Mixing (Exercises)

13.3: Elementary Mechanics

13.3E: Elementary Mechanics (Exercises)

This page titled 13: Applications of First Order Equations is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

13.1: Growth and Decay

This section begins with a discussion of exponential growth and decay, which you have probably already seen in calculus. We consider applications to radioactive decay, carbon dating, and compound interest. We also consider more complicated problems where the rate of change of a quantity is in part proportional to the magnitude of the quantity, but is also influenced by other other factors for example, a radioactive substance is manufactured at a certain rate, but decays at a rate proportional to its mass, or a saver makes regular deposits in a savings account that draws compound interest.

Since the applications in this section deal with functions of time, we'll denote the independent variable by t. If Q is a function of t, Q' will denote the derivative of Q with respect to t; thus,

$$Q' = \frac{dQ}{dt}.$$

Exponential Growth and Decay

One of the most common mathematical models for a physical process is the *exponential model*, where it is assumed that the rate of change of a quantity Q is proportional to Q; thus

$$Q' = aQ, \tag{13.1.1}$$

where a is the constant of proportionality.

From Example 2.1.3, the general solution of Equation 13.1.1 is

$$Q = ce^{at}$$

and the solution of the initial value problem

$$Q' = aQ, \quad Q(t_0) = Q_0$$

is

$$Q = Q_0 e^{a(t-t_0)}. \tag{13.1.2}$$

Since the solutions of $Q' = aQ$ are exponential functions, we say that a quantity Q that satisfies this equation *grows exponentially* if $a > 0$, or *decays exponentially* if $a < 0$ (Figure 3.1.1).

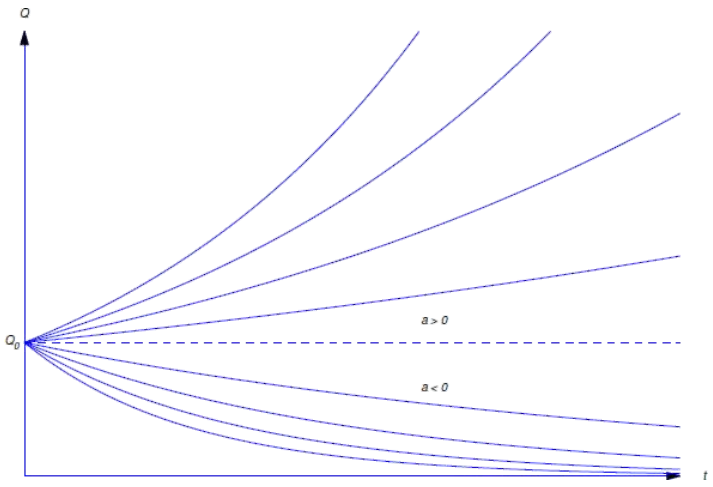


Figure 3.1.1 : Exponential growth and decay

Radioactive Decay

Experimental evidence shows that radioactive material decays at a rate proportional to the mass of the material present. According to this model the mass $Q(t)$ of a radioactive material present at time t satisfies Equation 13.1.1, where a is a negative constant whose value for any given material must be determined by experimental observation. For simplicity, we'll replace the negative constant a by $-k$, where k is a positive number that we'll call the *decay constant* of the material. Thus, Equation 13.1.1 becomes

$$Q' = -kQ.$$

If the mass of the material present at $t = t_0$ is Q_0, the mass present at time t is the solution of

$$Q' = -kQ, \quad Q(t_0) = Q_0.$$

From Equation 13.1.2 with $a = -k$, the solution of this initial value problem is

$$Q = Q_0 e^{-k(t-t_0)}. \tag{13.1.3}$$

The *half–life* τ of a radioactive material is defined to be the time required for half of its mass to decay; that is, if $Q(t_0) = Q_0$, then

$$Q(\tau + t_0) = \frac{Q_0}{2}. \tag{13.1.4}$$

From Equation 13.1.3 with $t = \tau + t_0$, Equation 13.1.4 is equivalent to

$$Q_0 e^{-k\tau} = \frac{Q_0}{2},$$

so

$$e^{-k\tau} = \frac{1}{2}.$$

Taking logarithms yields

$$-k\tau = \ln \frac{1}{2} = -\ln 2,$$

so the half-life is

$$\tau = \frac{1}{k} \ln 2. \tag{13.1.5}$$

(Figure 3.1.2). The half-life is independent of t_0 and Q_0, since it is determined by the properties of material, not by the amount of the material present at any particular time.

Example 3.1.1

A radioactive substance has a half-life of 1620 years.

a. If its mass is now 4 g (grams), how much will be left 810 years from now?
b. Find the time t_1 when 1.5 g of the substance remain.

Solution a

From Equation 13.1.3 with $t_0 = 0$ and $Q_0 = 4$,

$$Q = 4e^{-kt}, \tag{13.1.6}$$

where we determine k from Equation 13.1.5, with τ= 1620 years:

$$k = \frac{\ln 2}{\tau} = \frac{\ln 2}{1620}.$$

Substituting this in Equation 13.1.6 yields

$$Q = 4e^{-(t \ln 2)/1620}. \tag{13.1.7}$$

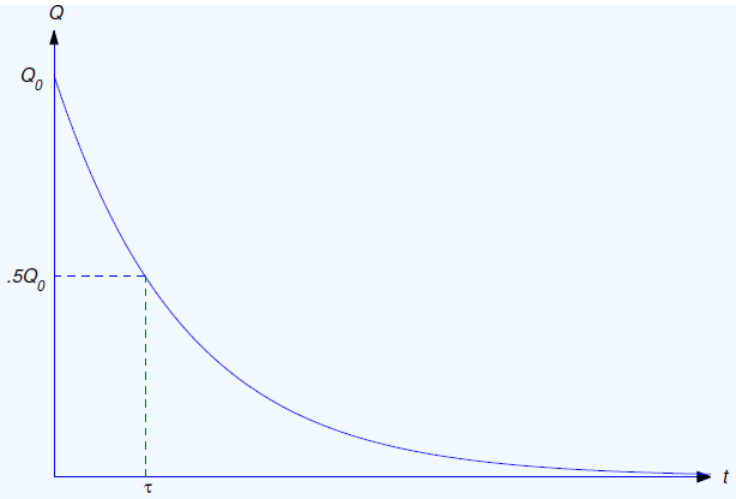


Figure 3.1.2 : Half-life of a radioactive substance

Therefore the mass left after 810 years will be

$$\begin{aligned} Q(810) &= 4e^{-(810\ln 2)/1620} = 4e^{-(\ln 2)/2} \\ &= 2\sqrt{2}\,\text{g}. \end{aligned}$$

Solution b

Setting $t = t_1$ in Equation 13.1.7 and requiring that $Q(t_1) = 1.5$ yields

$$\frac{3}{2} = 4e^{(-t_1 \ln 2)/1620}.$$

Dividing by 4 and taking logarithms yields

$$\ln\frac{3}{8} = -\frac{t_1 \ln 2}{1620}.$$

Since $\ln 3/8 = -\ln 8/3$,

$$t_1 = 1620\frac{\ln 8/3}{\ln 2} \approx 2292.4 \text{ years}.$$

Interest Compounded Continuously

Suppose we deposit an amount of money Q_0 in an interest-bearing account and make no further deposits or withdrawals for t years, during which the account bears interest at a constant annual rate r. To calculate the value of the account at the end of t years, we need one more piece of information: how the interest is added to the account, or—as the bankers say—how it is *compounded*. If the interest is compounded annually, the value of the account is multiplied by $1+r$ at the end of each year. This means that after t years the value of the account is

$$Q(t) = Q_0(1+r)^t.$$

If interest is compounded semiannually, the value of the account is multiplied by $(1+r/2)$ every 6 months. Since this occurs twice annually, the value of the account after t years is

$$Q(t) = Q_0\left(1+\frac{r}{2}\right)^{2t}.$$

In general, if interest is compounded n times per year, the value of the account is multiplied n times per year by $(1+r/n)$; therefore, the value of the account after t years is

$$Q(t) = Q_0\left(1+\frac{r}{n}\right)^{nt}. \tag{13.1.8}$$

Thus, increasing the frequency of compounding increases the value of the account after a fixed period of time. Table 3.1.1 shows the effect of increasing the number of compoundings over $t = 5$ years on an initial deposit of $Q_0 = 100$ (dollars), at an annual interest rate of 6%.

Table 3.1.1 : The effect of compound interest

n (number of compoundings per year)	$\$100\left(1+\frac{.06}{n}\right)^{5n}$ (value in dollars after 5 years)
1	$133.82
2	$134.39
4	$134.68
8	$134.83
364	$134.98

You can see from Table 3.1.1 that the value of the account after 5 years is an increasing function of n. Now suppose the maximum allowable rate of interest on savings accounts is restricted by law, but the time intervals between successive compoundings isn't; then competing banks can attract savers by compounding often. The ultimate step in this direction is to *compound continuously*, by which we mean that $n \to \infty$ in Equation 13.1.8. Since we know from calculus that

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n = e^r,$$

this yields

$$\begin{aligned} Q(t) &= \lim_{n\to\infty} Q_0\left(1+\tfrac{r}{n}\right)^{nt} = Q_0\left[\lim_{n\to\infty}\left(1+\tfrac{r}{n}\right)^n\right]^t \\ &= Q_0 e^{rt}. \end{aligned}$$

Observe that $Q = Q_0 e^{rt}$ is the solution of the initial value problem

$$Q' = rQ, \quad Q(0) = Q_0;$$

that is, with continuous compounding the value of the account grows exponentially.

Below is a video on exponential growth models and differential equations.



Below is a video on solving a differential equation that models exponential growth.



Example 3.1.2

If \$150 is deposited in a bank that pays $5\frac{1}{2}\%$ annual interest compounded continuously, the value of the account after t years is

$$Q(t) = 150e^{.055t}$$

dollars. (Note that it is necessary to write the interest rate as a decimal; thus, $r = .055$.) Therefore, after $t = 10$ years the value of the account is

$$Q(10) = 150e^{.55} \approx \$259.99.$$

Example 3.1.3

We wish to accumulate \$10,000 in 10 years by making a single deposit in a savings account bearing $5\frac{1}{2}\%$ annual interest compounded continuously. How much must we deposit in the account?

Solution

The value of the account at time t is

$$Q(t) = Q_0 e^{.055t}. \tag{13.1.9}$$

Since we want $Q(10)$ to be \$10,000, the initial deposit Q_0 must satisfy the equation

$$10000 = Q_0 e^{.55}, \tag{13.1.10}$$

obtained by setting $t = 10$ and $Q(10) = 10000$ in Equation 13.1.9. Solving Equation 13.1.10 for Q_0 yields

$$Q_0 = 10000e^{-.55} \approx \$5769.50.$$

Below is a video on modeling carbon 14 with a differential equation.



Below is a video on defining and differential equations that model continuous compound interest.



Mixed Growth and Decay

Example 3.1.4

A radioactive substance with decay constant k is produced at a constant rate of a units of mass per unit time.

a. Assuming that $Q(0) = Q_0$, find the mass $Q(t)$ of the substance present at time t.
b. Find $\lim_{t\to\infty} Q(t)$.

Solution a:

Here

$$Q' = \text{ rate of increase of } Q - \text{ rate of decrease of } Q.$$

The rate of increase is the constant a. Since Q is radioactive with decay constant k, the rate of decrease is kQ. Therefore

$$Q' = a - kQ.$$

This is a linear first order differential equation. Rewriting it and imposing the initial condition shows that Q is the solution of the initial value problem

$$Q' + kQ = a, \quad Q(0) = Q_0. \tag{13.1.11}$$

Since e^{-kt} is a solution of the complementary equation, the solutions of Equation 13.1.11 are of the form $Q = ue^{-kt}$, where $u'e^{-kt} = a$, so $u' = ae^{kt}$. Hence,

$$u = \frac{a}{k}e^{kt} + c$$

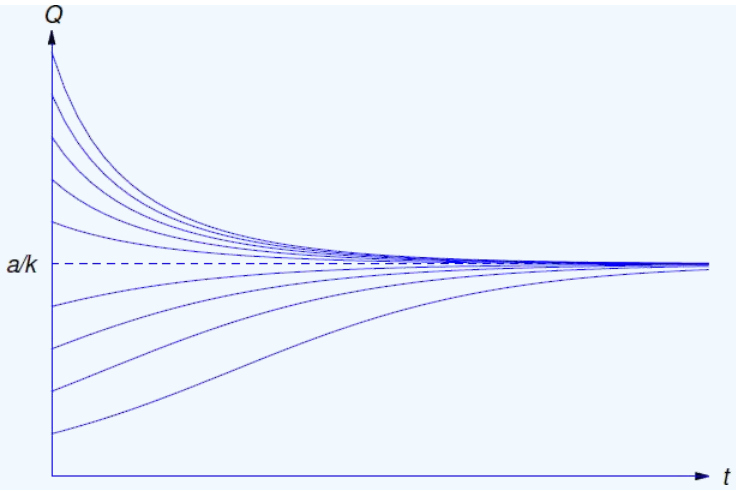


Figure 3.1.3 : $Q(t)$ approaches the steady state value $\frac{a}{k}$ as $t \to \infty$

and

$$Q = ue^{-kt} = \frac{a}{k} + ce^{-kt}.$$

Since $Q(0) = Q_0$, setting $t = 0$ here yields

$$Q_0 = \frac{a}{k} + c \quad \text{or} \quad c = Q_0 - \frac{a}{k}.$$

Therefore

$$Q = \frac{a}{k} + \left(Q_0 - \frac{a}{k}\right) e^{-kt}. \tag{13.1.12}$$

b. Since $k > 0$, $\lim_{t\to\infty} e^{-kt} = 0$, so from Equation 13.1.12

$$\lim_{t\to\infty} Q(t) = \frac{a}{k}.$$

This limit depends only on a and k, and not on Q_0. We say that a/k is the *steady state* value of Q. From Equation 13.1.12we also see that Q approaches its steady state value from above if $Q_0 > a/k$, or from below if $Q_0 < a/k$. If $Q_0 = a/k$, then Q remains constant (Figure 3.1.3).

Carbon Dating

The fact that Q approaches a steady state value in the situation discussed in Example 4 underlies the method of *carbon dating*, devised by the American chemist and Nobel Prize Winner *W.S. Libby*.

Carbon 12 is stable, but carbon-14, which is produced by cosmic bombardment of nitrogen in the upper atmosphere, is radioactive with a half-life of about 5570 years. Libby assumed that the quantity of carbon-12 in the atmosphere has been constant throughout time, and that the quantity of radioactive carbon-14 achieved its steady state value long ago as a result of its creation and decomposition over millions of years. These assumptions led Libby to conclude that the ratio of carbon-14 to carbon-12 has been nearly constant for a long time. This constant, which we denote by R, has been determined experimentally.

Living cells absorb both carbon-12 and carbon-14 in the proportion in which they are present in the environment. Therefore the ratio of carbon-14 to carbon-12 in a living cell is always R. However, when the cell dies it ceases to absorb carbon, and the ratio of carbon-14 to carbon-12 decreases exponentially as the radioactive carbon-14 decays. This is the basis for the method of carbon dating, as illustrated in the next example.

Example 3.1.5

An archaeologist investigating the site of an ancient village finds a burial ground where the amount of carbon-14 present in individual remains is between 42 and 44% of the amount present in live individuals. Estimate the age of the village and the

length of time for which it survived.

Solution

Let $Q = Q(t)$ be the quantity of carbon-14 in an individual set of remains t years after death, and let Q_0 be the quantity that would be present in live individuals. Since carbon-14 decays exponentially with half-life 5570 years, its decay constant is

$$k = \frac{\ln 2}{5570}.$$

Therefore

$$Q = Q_0 e^{-t(\ln 2)/5570}$$

if we choose our time scale so that $t_0 = 0$ is the time of death. If we know the present value of Q we can solve this equation for t, the number of years since death occurred. This yields

$$t = -5570 \frac{\ln Q/Q_0}{\ln 2}.$$

It is given that $Q = .42Q_0$ in the remains of individuals who died first. Therefore these deaths occurred about

$$t_1 = -5570 \frac{\ln .42}{\ln 2} \approx 6971$$

years ago. For the most recent deaths, $Q = .44Q_0$; hence, these deaths occurred about

$$t_2 = -5570 \frac{\ln .44}{\ln 2} \approx 6597$$

years ago. Therefore it is reasonable to conclude that the village was founded about 7000 years ago, and lasted for about 400 years.

A Savings Program

Example 3.1.6 :

A person opens a savings account with an initial deposit of \$1000 and subsequently deposits \$50 per week. Find the value $Q(t)$ of the account at time $t > 0$, assuming that the bank pays 6% interest compounded continuously.

Solution

Observe that Q isn't continuous, since there are 52 discrete deposits per year of \$50 each. To construct a mathematical model for this problem in the form of a differential equation, we make the simplifying assumption that the deposits are made continuously at a rate of \$2600 per year. This is essential, since solutions of differential equations are continuous functions. With this assumption, Q increases continuously at the rate

$$Q' = 2600 + 0.06Q$$

and therefore Q satisfies the differential equation

$$Q' - .06Q = 2600. \tag{13.1.13}$$

(Of course, we must recognize that the solution of this equation is an approximation to the true value of Q at any given time. We'll discuss this further below.) Since $e^{.06t}$ is a solution of the complementary equation, the solutions of Equation 13.1.13are of the form $Q = ue^{.06t}$, where $u'e^{.06t} = 2600$. Hence, $u' = 2600e^{-.06t}$,

$$u = -\frac{2600}{.06} e^{-0.06t} + c$$

and

$$Q = ue^{.06t} = -\frac{2600}{.06} + ce^{.06t}. \tag{13.1.14}$$

Setting $t = 0$ and $Q = 1000$ here yields

$$c = 1000 + \frac{2600}{0.06},$$

and substituting this into Equation 13.1.14yields

$$Q = 1000e^{.06t} + \frac{2600}{.06}(e^{.06t} - 1) \tag{13.1.15}$$

where the first term is the value due to the initial deposit and the second is due to the subsequent weekly deposits.

Mathematical models must be tested for validity by comparing predictions based on them with the actual outcome of experiments. Example 6 is unusual in that we can compute the exact value of the account at any specified time and compare it with the approximate value predicted by Equation 13.1.15 (See *Exercise 4.1.21*). The following table gives a comparison for a ten year period. Each exact answer corresponds to the time of the year-end deposit, and each year is assumed to have exactly 52 weeks.

Table 3.1.3

Year	Approximate Value of Q (Example 3.1.6)	Exact Value of P (*Exercise 4.1.21*)	Error $Q - P$	Percentage Error $(Q - P)/P$
1	$3741.42	$3739.87	$1.55	.0413
2	6652.36	6649.17	3.19	.0479
3	9743.30	9738.37	4.93	.0506
4	13, 025.38	13, 018.60	6.78	.0521
5	16, 510.41	16, 501.66	8.75	.0530
6	20, 210.94	20, 200.11	10.83	.0536
7	24, 140.30	24, 127.25	13.05	.0541
8	28, 312.63	28, 297.23	15.40	.0544
9	32, 742.97	32, 725.07	17.90	.0547
10	37, 447.27	37, 426.72	20.55	.0549

This page titled 13.1: Growth and Decay is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **4.1: Growth and Decay** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

13.1E: Growth and Decay (Exercises)

Q4.1.1

1. The half-life of a radioactive substance is 3200 years. Find the quantity $Q(t)$ of the substance left at time $t > 0$ if $Q(0) = 20$ g.

2. The half-life of a radioactive substance is 2 days. Find the time required for a given amount of the material to decay to 1/10 of its original mass.

3. A radioactive material loses 25% of its mass in 10 minutes. What is its half-life?

4. A tree contains a known percentage p_0 of a radioactive substance with half-life τ. When the tree dies the substance decays and isn't replaced. If the percentage of the substance in the fossilized remains of such a tree is found to be p_1, how long has the tree been dead?

5. If t_p and t_q are the times required for a radioactive material to decay to $1/p$ and $1/q$ times its original mass (respectively), how are t_p and t_q related?

6. Find the decay constant k for a radioactive substance, given that the mass of the substance is Q_1 at time t_1 and Q_2 at time t_2.

7. A process creates a radioactive substance at the rate of 2 g/hr and the substance decays at a rate proportional to its mass, with constant of proportionality $k = .1(\text{hr})^{-1}$. If $Q(t)$ is the mass of the substance at time t, find $\lim_{t\to\infty} Q(t)$.

8. A bank pays interest continuously at the rate of 6%. How long does it take for a deposit of Q_0 to grow in value to $2Q_0$?

9. At what rate of interest, compounded continuously, will a bank deposit double in value in 8 years?

10. A savings account pays 5% per annum interest compounded continuously. The initial deposit is Q_0 dollars. Assume that there are no subsequent withdrawals or deposits.

a. How long will it take for the value of the account to triple?
b. What is Q_0 if the value of the account after 10 years is \$100,000 dollars?

11. A candymaker makes 500 pounds of candy per week, while his large family eats the candy at a rate equal to $Q(t)/10$ pounds per week, where $Q(t)$ is the amount of candy present at time t.

a. Find $Q(t)$ for $t > 0$ if the candymaker has 250 pounds of candy at $t = 0$.
b. Find $\lim_{t\to\infty} Q(t)$.

12. Suppose a substance decays at a yearly rate equal to half the square of the mass of the substance present. If we start with 50 g of the substance, how long will it be until only 25 g remain?

13. A super bread dough increases in volume at a rate proportional to the volume V present. If V increases by a factor of 10 in 2 hours and $V(0) = V_0$, find V at any time t. How long will it take for V to increase to $100V_0$?

14. A radioactive substance decays at a rate proportional to the amount present, and half the original quantity Q_0 is left after 1500 years. In how many years would the original amount be reduced to $3Q_0/4$? How much will be left after 2000 years?

15. A wizard creates gold continuously at the rate of 1 ounce per hour, but an assistant steals it continuously at the rate of 5% of however much is there per hour. Let $W(t)$ be the number of ounces that the wizard has at time t. Find $W(t)$ and $\lim_{t\to\infty} W(t)$ if $W(0) = 1$.

16. A process creates a radioactive substance at the rate of 1 g/hr, and the substance decays at an hourly rate equal to 1/10 of the mass present (expressed in grams). Assuming that there are initially 20 g, find the mass $S(t)$ of the substance present at time t, and find $\lim_{t\to\infty} S(t)$.

17. A tank is empty at $t = 0$. Water is added to the tank at the rate of 10 gal/min, but it leaks out at a rate (in gallons per minute) equal to the number of gallons in the tank. What is the smallest capacity the tank can have if this process is to continue forever?

18. A person deposits \$25,000 in a bank that pays 5% per year interest, compounded continuously. The person continuously withdraws from the account at the rate of \$750 per year. Find $V(t)$, the value of the account at time t after the initial deposit.

19. A person has a fortune that grows at rate proportional to the square root of its worth. Find the worth W of the fortune as a function of t if it was \$1 million 6 months ago and is \$4 million today.

20. Let $p = p(t)$ be the quantity of a product present at time t. The product is manufactured continuously at a rate proportional to p, with proportionality constant 1/2, and it is consumed continuously at a rate proportional to p^2, with proportionality constant 1/8. Find $p(t)$ if $p(0) = 100$.

21.

a. In the situation of Example 4.1.6 find the exact value P(t) of the person's account after t years, where t is an integer. Assume that each year has exactly 52 weeks, and include the year-end deposit in the computation.

HINT: At time t the initial \$1000 has been on deposit for t years. There have been $52t$ deposits of \$50 each. The first \$50 has been on deposit for $t - 1/52$ years, the second for $t - 2/52$ years ... in general, the **j** th \$50 has been on deposit for $t - j/52$ years ($1 \le j \le 52t$). Find the present value of each \$50 deposit assuming 6% interest compounded continuously, and use the formula

$$1 + x + x^2 + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x} (x \neq 1)$$

to find their total value.

b. Let

$$p(t) = \frac{Q(t) - P(t)}{P(t)}$$

be the relative error after t years. Find

$$p(\infty) = \lim_{t \to \infty} p(t).$$

22. A homebuyer borrows P_0 dollars at an annual interest rate r, agreeing to repay the loan with equal monthly payments of M dollars per month over N years.

a. Derive a differential equation for the loan principal (amount that the homebuyer owes) $P(t)$ at time $t > 0$, making the simplifying assumption that the homebuyer repays the loan continuously rather than in discrete steps. (See Example 4.1.6.)

b. Solve the equation derived in (a).

c. Use the result of (b) to determine an approximate value for M assuming that each year has exactly 12 months of equal length.

d. It can be shown that the exact value of M is given by

$$M = \frac{rP_0}{12} \left(1 - (1 + r/12)^{-12N}\right)^{-1}.$$

Compare the value of M obtained from the answer in (c) to the exact value if (i) $P_0 = \$50,000$, \(r=7{1\over2}\)%, $N = 20$ (ii) $P_0 = \$150,000$, $r = 9.0\%$, $N = 30$.

23. Assume that the homebuyer of **Exercise 4.1.22** elects to repay the loan continuously at the rate of αM dollars per month, where α is a constant greater than 1. (This is called **accelerated payment**.)

a. Determine the time $T(\alpha)$ when the loan will be paid off and the amount $S(\alpha)$ that the homebuyer will save.
b. Suppose $P_0 = \$50,000$, $r = 8\%$, and $N = 15$. Compute the savings realized by accelerated payments with $\alpha = 1.05, 1.10$, and 1.15.

24. A benefactor wishes to establish a trust fund to pay a researcher's salary for T years. The salary is to start at S_0 dollars per year and increase at a fractional rate of a per year. Find the amount of money P_0 that the benefactor must deposit in a trust fund paying interest at a rate r per year. Assume that the researcher's salary is paid continuously, the interest is compounded continuously, and the salary increases are granted continuously.

25. A radioactive substance with decay constant k is produced at the rate of

$$\frac{at}{1 + btQ(t)}$$

units of mass per unit time, where a and b are positive constants and $Q(t)$ is the mass of the substance present at time t; thus, the rate of production is small at the start and tends to slow when Q is large.

a. Set up a differential equation for Q.

b. Choose your own positive values for a, b, k, and $Q_0 = Q(0)$. Use a numerical method to discover what happens to $Q(t)$ as $t \to \infty$. (Be precise, expressing your conclusions in terms of a, b, k. However, no proof is required.)

26. Follow the instructions of **Exercise 4.1.25**, assuming that the substance is produced at the rate of $at/(1+bt(Q(t))^2)$ units of mass per unit of time.

27. Follow the instructions of **Exercise 4.1.25**, assuming that the substance is produced at the rate of $at/(1+bt)$ units of mass per unit of time.

This page titled 13.1E: Growth and Decay (Exercises) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **4.1E: Growth and Decay (Exercises)** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

13.2: Cooling and Mixing

Newton's Law of Cooling

Newton's law of cooling states that if an object with temperature $T(t)$ at time t is in a medium with temperature $T_m(t)$, the rate of change of T at time t is proportional to $T(t)-T_m(t)$; thus, T satisfies a differential equation of the form

$$T' = -k(T - T_m). \tag{13.2.1}$$

Here $k > 0$, since the temperature of the object must decrease if $T > T_m$, or increase if $T < T_m$. We'll call k the *temperature decay constant of the medium*.

For simplicity, in this section we'll assume that the medium is maintained at a constant temperature T_m. This is another example of building a simple mathematical model for a physical phenomenon. Like most mathematical models it has its limitations. For example, it is reasonable to assume that the temperature of a room remains approximately constant if the cooling object is a cup of coffee, but perhaps not if it is a huge cauldron of molten metal. (For more on this see *Exercise 4.2.17*.)

To solve Equation 13.2.1, we rewrite it as

$$T' + kT = kT_m.$$

Since e^{-kt} is a solution of the complementary equation, the solutions of this equation are of the form $T = ue^{-kt}$, where $u'e^{-kt} = kT_m$, so $u' = kT_m e^{kt}$. Hence,

$$u = T_m e^{kt} + c,$$

so

$$T = ue^{-kt} = T_m + ce^{-kt}.$$

If $T(0) = T_0$, setting $t = 0$ here yields $c = T_0 - T_m$, so

$$T = T_m + (T_0 - T_m)e^{-kt}. \tag{13.2.2}$$

Note that $T - T_m$ decays exponentially, with decay constant k.

Below is a video on Newton's Law of Cooling.



Example 3.2.1

A ceramic insulator is baked at 400°C and cooled in a room in which the temperature is 25°C. After 4 minutes the temperature of the insulator is 200°C. What is its temperature after 8 minutes?

Solution

Here $T_0 = 400$ and $T_m = 25$, so Equation 13.2.2 becomes

$$T = 25 + 375e^{-kt}. \tag{13.2.3}$$

We determine k from the stated condition that $T(4) = 200$; that is,

$$200 = 25 + 375e^{-4k};$$

hence,

$$e^{-4k} = \frac{175}{375} = \frac{7}{15}.$$

Taking logarithms and solving for k yields

$$k = -\frac{1}{4}\ln\frac{7}{15} = \frac{1}{4}\ln\frac{15}{7}.$$

Substituting this into Equation 13.2.3 yields

$$T = 25 + 375e^{-\frac{t}{4}\ln\frac{15}{7}}$$

(Figure 3.2.1). Therefore the temperature of the insulator after 8 minutes is

$$\begin{aligned} T(8) &= 25 + 375e^{-2\ln\frac{15}{7}} \\ &= 25 + 375\left(\tfrac{7}{15}\right)^2 \approx 107^\circ\text{C}. \end{aligned}$$

Example 3.2.2

An object with temperature 72°F is placed outside, where the temperature is -20°F. At 11:05 the temperature of the object is 60°F and at 11:07 its temperature is 50°F. At what time was the object placed outside?

Solution

Let $T(t)$ be the temperature of the object at time t. For convenience, we choose the origin $t_0 = 0$ of the time scale to be 11:05 so that $T_0 = 60$. We must determine the time τ when $T(\tau) = 72$. Substituting $T_0 = 60$ and $T_m = -20$ into Equation 13.2.2 yields

$$T = -20 + \big(60 - (-20)\big)e^{-kt}$$

or

$$T = -20 + 80e^{-kt}. \tag{13.2.4}$$

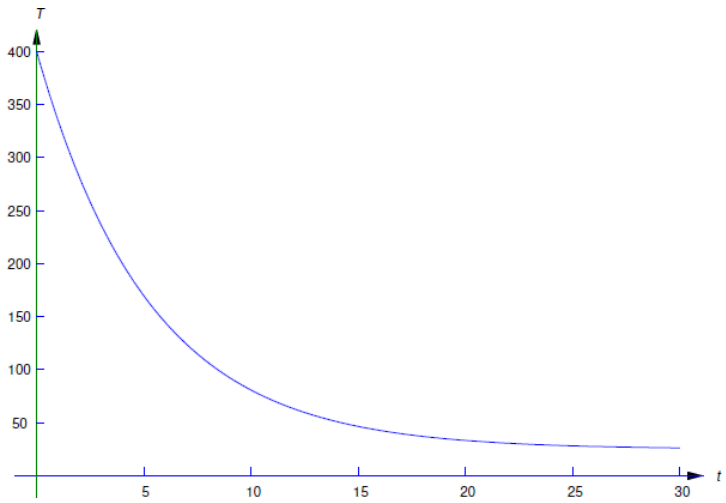


Figure 3.2.1 : $T = 25 + 375e^{-(t/4)\ln 15/7}$

We obtain k from the stated condition that the temperature of the object is 50°F at 11:07. Since 11:07 is $t = 2$ on our time scale, we can determine k by substituting $T = 50$ and $t = 2$ into Equation 13.2.4 to obtain

$$50 = -20 + 80e^{-2k}$$

This is shown in Figure 3.2.2 .

Hence,

$$e^{-2k} = \frac{70}{80} = \frac{7}{8}.$$

Taking logarithms and solving for k yields

$$k = -\frac{1}{2}\ln\frac{7}{8} = \frac{1}{2}\ln\frac{8}{7}.$$

Substituting this into Equation 13.2.4 yields

$$T = -20 + 80e^{-\frac{t}{2}\ln\frac{8}{7}},$$

and the condition $T(\tau) = 72$ implies that

$$72 = -20 + 80e^{-\frac{\tau}{2}\ln\frac{8}{7}};$$

hence,

$$e^{-\frac{\tau}{2}\ln\frac{8}{7}} = \frac{92}{80} = \frac{23}{20}.$$

Taking logarithms and solving for τ yields

$$\tau = -\frac{2\ln\frac{23}{20}}{\ln\frac{8}{7}} \approx -2.09 \text{ min}.$$

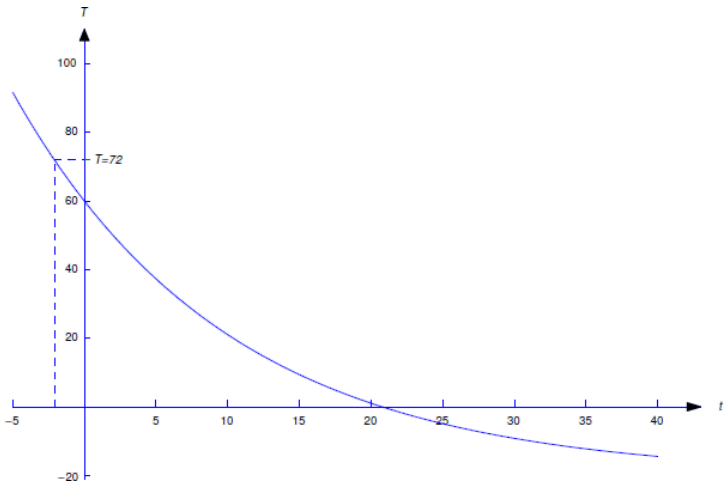


Figure 3.2.2 : $-20 + 80e^{-\frac{t}{2}\ln\frac{8}{7}}$

Therefore the object was placed outside about 2 minutes and 5 seconds before 11:05; that is, at 11:02:55.

Mixing Problems

In the next two examples a saltwater solution with a given concentration (weight of salt per unit volume of solution) is added at a specified rate to a tank that initially contains saltwater with a different concentration. The problem is to determine the quantity of salt in the tank as a function of time. This is an example of a *mixing problem*. To construct a tractable mathematical model for mixing problems we assume in our examples (and most exercises) that the mixture is stirred instantly so that the salt is always uniformly distributed throughout the mixture. *Exercises 4.2.22* and *4.2.23* deal with situations where this isn't so, but the distribution of salt becomes approximately uniform as $t \to \infty$.

Example 3.2.3

A tank initially contains 40 pounds of salt dissolved in 600 gallons of water. Starting at $t_0 = 0$, water that contains 1/2 pound of salt per gallon is poured into the tank at the rate of 4 gal/min and the mixture is drained from the tank at the same rate (Figure 3.2.3).

a. Find a differential equation for the quantity $Q(t)$ of salt in the tank at time $t > 0$, and solve the equation to determine $Q(t)$.
b. Find $\lim_{t\to\infty} Q(t)$.

Solution a

To find a differential equation for Q, we must use the given information to derive an expression for Q'. But Q' is the rate of change of the quantity of salt in the tank changes with respect to time; thus, if *rate in* denotes the rate at which salt enters the tank and *rate out* denotes the rate by which it leaves, then

$$Q' = \text{rate in} - \text{rate out}. \tag{13.2.5}$$

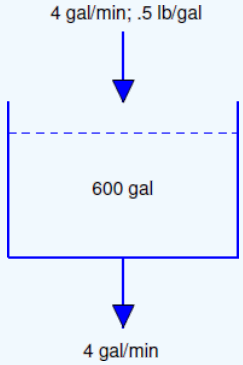


Figure 3.2.3 : A mixing problem

The rate in is

$$\left(\frac{1}{2}\ \text{lb/gal}\right) \times (4\ \text{gal/min}) = 2\ \text{lb/min}.$$

Determining the rate out requires a little more thought. We're removing 4 gallons of the mixture per minute, and there are always 600 gallons in the tank; that is, we are removing $1/150$ of the mixture per minute. Since the salt is evenly distributed in the mixture, we are also removing $1/150$ of the salt per minute. Therefore, if there are $Q(t)$ pounds of salt in the tank at time t, the rate out at any time t is $Q(t)/150$. Alternatively, we can arrive at this conclusion by arguing that

$$\begin{aligned}\text{rate out} &= (\text{concentration}) \times (\text{rate of flow out}) \\ &= (\text{lb/gal}) \times (\text{gal/min}) \\ &= \frac{Q(t)}{600} \times 4 \\ &= \frac{Q(t)}{150}.\end{aligned}$$

We can now write Equation 13.2.5 as

$$Q' = 2 - \frac{Q}{150}.$$

This first order equation can be rewritten as

$$Q' + \frac{Q}{150} = 2.$$

Since $e^{-t/150}$ is a solution of the complementary equation, the solutions of this equation are of the form $Q = ue^{-t/150}$, where $u'e^{-t/150} = 2$, so $u' = 2e^{t/150}$. Hence,

$$u = 300e^{t/150} + c,$$

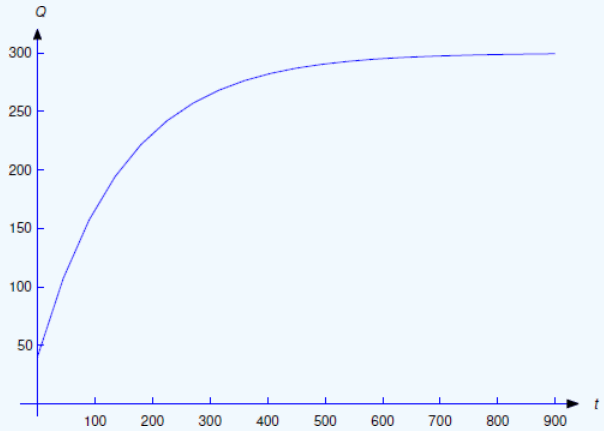


Figure 3.2.4 : $Q = 300 - 260e^{-t/150}$

so

$$Q = ue^{-t/150} = 300 + ce^{-t/150} \tag{13.2.6}$$

(Figure 3.2.4). Since $Q(0) = 40$, $c = -260$; therefore,

$$Q = 300 - 260e^{-t/150}.$$

Solution b

From Equation 13.2.6, we see that that $\lim_{t\to\infty} Q(t) = 300$ for any value of $Q(0)$. This is intuitively reasonable, since the incoming solution contains 1/2 pound of salt per gallon and there are always 600 gallons of water in the tank.

Below is a video on mixing.



Example 3.2.4

A 500-liter tank initially contains 10 g of salt dissolved in 200 liters of water. Starting at $t_0 = 0$, water that contains 1/4 g of salt per liter is poured into the tank at the rate of 4 liters/min and the mixture is drained from the tank at the rate of 2 liters/min (Figure [figure:4.2.5}). Find a differential equation for the quantity $Q(t)$ of salt in the tank at time t prior to the time when the tank overflows and find the concentration $K(t)$ (g/liter) of salt in the tank at any such time.

Solution

We first determine the amount $W(t)$ of solution in the tank at any time t prior to overflow. Since $W(0) = 200$ and we are adding 4 liters/min while removing only 2 liters/min, there's a net gain of 2 liters/min in the tank; therefore,

$$W(t) = 2t + 200.$$

Since $W(150) = 500$ liters (capacity of the tank), this formula is valid for $0 \le t \le 150$.

Now let $Q(t)$ be the number of grams of salt in the tank at time t, where $0 \le t \le 150$. As in Example 3.2.3

$$Q' = \text{rate in} - \text{rate out}. \tag{13.2.7}$$

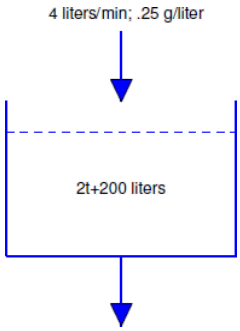


Figure 3.2.5 : Another mixing problem

The rate in is

$$\left(\frac{1}{4}\ \text{g/liter}\right) \times (4\ \text{liters/min}) = 1\ \text{g/min}. \tag{13.2.8}$$

To determine the rate out, we observe that since the mixture is being removed from the tank at the constant rate of 2 liters/min and there are $2t + 200$ liters in the tank at time t, the fraction of the mixture being removed per minute at time t is

$$\frac{2}{2t+200} = \frac{1}{t+100}.$$

We're removing this same fraction of the salt per minute. Therefore, since there are $Q(t)$ grams of salt in the tank at time t,

$$\text{rate out} = \frac{Q(t)}{t+100}. \tag{13.2.9}$$

Alternatively, we can arrive at this conclusion by arguing that

$$\begin{aligned} \text{rate out} &= (\text{concentration}) \times (\text{rate of flow out}) = (\text{g/liter}) \times (\text{liters/min}) \\ &= \tfrac{Q(t)}{2t+200} \times 2 = \tfrac{Q(t)}{t+100}. \end{aligned}$$

Substituting Equation 13.2.8 and Equation 13.2.9 into Equation 13.2.7 yields

$$Q' = 1 - \frac{Q}{t+100}, \quad \text{so} \quad Q' + \frac{1}{t+100}Q = 1. \tag{13.2.10}$$

By separation of variables, $1/(t+100)$ is a solution of the complementary equation, so the solutions of Equation 13.2.10 are of the form

$$Q = \frac{u}{t+100}, \quad \text{where} \quad \frac{u'}{t+100} = 1, \quad \text{so} \quad u' = t + 100.$$

Hence,

$$u = \frac{(t+100)^2}{2} + c. \tag{13.2.11}$$

Since $Q(0) = 10$ and $u = (t+100)Q$, Equation 13.2.11 implies that

$$(100)(10) = \frac{(100)^2}{2} + c,$$

so

$$c = 100(10) - \frac{(100)^2}{2} = -4000$$

and therefore

$$u = \frac{(t+100)^2}{2} - 4000.$$

Hence,

$$Q = \frac{u}{t+200} = \frac{t+100}{2} - \frac{4000}{t+100}.$$

Now let $K(t)$ be the concentration of salt at time t. Then

$$K(t) = \frac{1}{4} - \frac{2000}{(t+100)^2}$$

This is shown in Figure 3.2.6 .

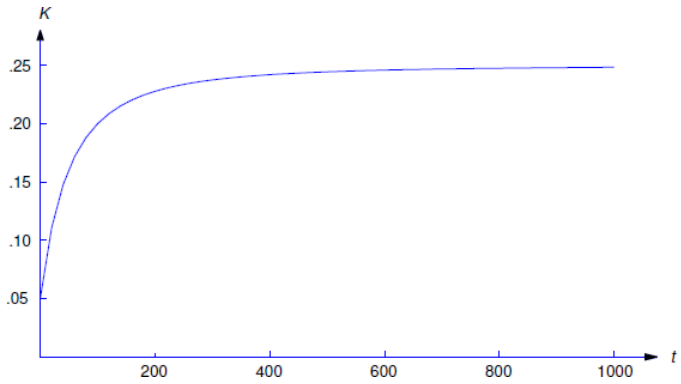


Figure 3.2.6 : $K(t) = \frac{1}{4} - \frac{2000}{(t+100)^2}$

Below is a video on salt in a tank and differential equations.



This page titled 13.2: Cooling and Mixing is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **4.2: Cooling and Mixing** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

13.2E: Cooling and Mixing (Exercises)

Q4.2.1

1. A thermometer is moved from a room where the temperature is 70°F to a freezer where the temperature is $12^\circ F$. After 30 seconds the thermometer reads 40°F. What does it read after 2 minutes?

2. A fluid initially at 100°C is placed outside on a day when the temperature is -10°C, and the temperature of the fluid drops 20°C in one minute. Find the temperature $T(t)$ of the fluid for $t > 0$.

3. At 12:00 pm a thermometer reading 10°F is placed in a room where the temperature is 70°F. It reads 56° when it is placed outside, where the temperature is 5°F, at 12:03. What does it read at 12:05 pm?

4. A thermometer initially reading 212°F is placed in a room where the temperature is 70°F. After 2 minutes the thermometer reads 125°F.

   a. What does the thermometer read after 4 minutes?
   b. When will the thermometer read 72°F?
   c. When will the thermometer read 69°F?

5. An object with initial temperature 150°C is placed outside, where the temperature is 35°C. Its temperatures at 12:15 and 12:20 are 120°C and 90°C, respectively.

   a. At what time was the object placed outside?
   b. When will its temperature be 40°C?

6. An object is placed in a room where the temperature is 20°C. The temperature of the object drops by 5°C in 4 minutes and by 7°C in 8 minutes. What was the temperature of the object when it was initially placed in the room?

7. A cup of boiling water is placed outside at 1:00 pm. One minute later the temperature of the water is 152°F. After another minute its temperature is 112°F. What is the outside temperature?

8. A tank initially contains 40 gallons of pure water. A solution with 1 gram of salt per gallon of water is added to the tank at 3 gal/min, and the resulting solution drains out at the same rate. Find the quantity $Q(t)$ of salt in the tank at time $t > 0$.

9. A tank initially contains a solution of 10 pounds of salt in 60 gallons of water. Water with $1/2$ pound of salt per gallon is added to the tank at 6 gal/min, and the resulting solution leaves at the same rate. Find the quantity $Q(t)$ of salt in the tank at time $t > 0$.

10. A tank initially contains 100 liters of a salt solution with a concentration of $.1$ g/liter. A solution with a salt concentration of $.3$ g/liter is added to the tank at 5 liters/min, and the resulting mixture is drained out at the same rate. Find the concentration $K(t)$ of salt in the tank as a function of t.

11. A 200 gallon tank initially contains 100 gallons of water with 20 pounds of salt. A salt solution with $1/4$ pound of salt per gallon is added to the tank at 4 gal/min, and the resulting mixture is drained out at 2 gal/min. Find the quantity of salt in the tank as it is about to overflow.

12. Suppose water is added to a tank at 10 gal/min, but leaks out at the rate of $1/5$ gal/min for each gallon in the tank. What is the smallest capacity the tank can have if the process is to continue indefinitely?

13. A chemical reaction in a laboratory with volume V (in ft^3) produces q_1 ft^3/min of a noxious gas as a byproduct. The gas is dangerous at concentrations greater than $\overline{c}$, but harmless at concentrations $\leq \overline{c}$. Intake fans at one end of the laboratory pull in fresh air at the rate of q_2 ft^3/min and exhaust fans at the other end exhaust the mixture of gas and air from the laboratory at the same rate. Assuming that the gas is always uniformly distributed in the room and its initial concentration c_0 is at a safe level, find the smallest value of q_2 required to maintain safe conditions in the laboratory for all time.

14. A 1200-gallon tank initially contains 40 pounds of salt dissolved in 600 gallons of water. Starting at $t_0 = 0$, water that contains $1/2$ pound of salt per gallon is added to the tank at the rate of 6 gal/min and the resulting mixture is drained from the tank at 4 gal/min. Find the quantity $Q(t)$ of salt in the tank at any time $t > 0$ prior to overflow.

15. Tank T_1 initially contain 50 gallons of pure water. Starting at $t_0 = 0$, water that contains 1 pound of salt per gallon is poured into T_1 at the rate of 2 gal/min. The mixture is drained from T_1 at the same rate into a second tank T_2, which initially contains 50

gallons of pure water. Also starting at $t_0 = 0$, a mixture from another source that contains 2 pounds of salt per gallon is poured into T_2 at the rate of 2 gal/min. The mixture is drained from T_2 at the rate of 4 gal/min.

a. Find a differential equation for the quantity $Q(t)$ of salt in tank T_2 at time $t > 0$.
b. Solve the equation derived in (a) to determine $Q(t)$.
c. Find $\lim_{t\to\infty} Q(t)$.

16. Suppose an object with initial temperature T_0 is placed in a sealed container, which is in turn placed in a medium with temperature T_m. Let the initial temperature of the container be S_0. Assume that the temperature of the object does not affect the temperature of the container, which in turn does not affect the temperature of the medium. (These assumptions are reasonable, for example, if the object is a cup of coffee, the container is a house, and the medium is the atmosphere.)

a. Assuming that the container and the medium have distinct temperature decay constants k and k_m respectively, use Newton's law of cooling to find the temperatures $S(t)$ and $T(t)$ of the container and object at time t.
b. Assuming that the container and the medium have the same temperature decay constant k, use Newton's law of cooling to find the temperatures $S(t)$ and $T(t)$ of the container and object at time t.
c. Find $\lim_{t\to\infty} S(t)$ and $\lim_{t\to\infty} T(t)$.

17. In our previous examples and exercises concerning Newton's law of cooling we assumed that the temperature of the medium remains constant. This model is adequate if the heat lost or gained by the object is insignificant compared to the heat required to cause an appreciable change in the temperature of the medium. If this isn't so, we must use a model that accounts for the heat exchanged between the object and the medium. Let $T = T(t)$ and $T_m = T_m(t)$ be the temperatures of the object and the medium, respectively, and let T_0 and T_{m0} be their initial values. Again, we assume that T and T_m are related by Newton's law of cooling,

$$T' = -k(T - T_m). \tag{A}$$

We also assume that the change in heat of the object as its temperature changes from T_0 to T is $a(T - T_0)$ and that the change in heat of the medium as its temperature changes from T_{m0} to T_m is $a_m(T_m - T_{m0})$, where a and a_m are positive constants depending upon the masses and thermal properties of the object and medium, respectively. If we assume that the total heat of the system consisting of the object and the medium remains constant (that is, energy is conserved), then

$$a(T - T_0) + a_m(T_m - T_{m0}) = 0. \tag{B}$$

a. Equation (A) involves two unknown functions T and T_m. Use (A) and (B) to derive a differential equation involving only T.
b. Find $T(t)$ and $T_m(t)$ for $t > 0$.
c. Find $\lim_{t\to\infty} T(t)$ and $\lim_{t\to\infty} T_m(t)$.

18. Control mechanisms allow fluid to flow into a tank at a rate proportional to the volume V of fluid in the tank, and to flow out at a rate proportional to V^2. Suppose $V(0) = V_0$ and the constants of proportionality are a and b, respectively. Find $V(t)$ for $t > 0$ and find $\lim_{t\to\infty} V(t)$.

19. Identical tanks T_1 and T_2 initially contain W gallons each of pure water. Starting at $t_0 = 0$, a salt solution with constant concentration c is pumped into T_1 at r gal/min and drained from T_1 into T_2 at the same rate. The resulting mixture in T_2 is also drained at the same rate. Find the concentrations $c_1(t)$ and $c_2(t)$ in tanks T_1 and T_2 for $t > 0$.

20. An infinite sequence of identical tanks T_1, T_2, …, T_n, …, initially contain W gallons each of pure water. They are hooked together so that fluid drains from T_n into T_{n+1} $(n = 1, 2, \cdots)$. A salt solution is circulated through the tanks so that it enters and leaves each tank at the constant rate of r gal/min. The solution has a concentration of c pounds of salt per gallon when it enters T_1.

a. Find the concentration $c_n(t)$ in tank T_n for $t > 0$.
b. Find $\lim_{t\to\infty} c_n(t)$ for each n.

21. Tanks T_1 and T_2 have capacities W_1 and W_2 liters, respectively. Initially they are both full of dye solutions with concentrations c_1 and c_2 grams per liter. Starting at $t_0 = 0$, the solution from T_1 is pumped into T_2 at a rate of r liters per minute, and the solution from T_2 is pumped into T_1 at the same rate.

a. Find the concentrations $c_1(t)$ and $c_2(t)$ of the dye in T_1 and T_2 for $t > 0$.
b. Find $\lim_{t\to\infty} c_1(t)$ and $\lim_{t\to\infty} c_2(t)$.

22. Consider the mixing problem of Example 4.2.3, but without the assumption that the mixture is stirred instantly so that the salt is always uniformly distributed throughout the mixture. Assume instead that the distribution approaches uniformity as $t \to \infty$. In this

case the differential equation for Q is of the form

$$Q' + \frac{a(t)}{150}Q = 2$$

where $\lim_{t\to\infty} a(t) = 1$.

a. Assuming that $Q(0) = Q_0$, can you guess the value of $\lim_{t\to\infty} Q(t)$?.
b. Use numerical methods to confirm your guess in the these cases:

$$\text{(i) } a(t) = t/(1+t) \quad \text{(ii) } a(t) = 1 - e^{-t^2} \quad \text{(iii) } a(t) = 1 - \sin(e^{-t}).$$

23. Consider the mixing problem of Example 4.2.4 in a tank with infinite capacity, but without the assumption that the mixture is stirred instantly so that the salt is always uniformly distributed throughout the mixture. Assume instead that the distribution approaches uniformity as $t \to \infty$. In this case the differential equation for Q is of the form

$$Q' + \frac{a(t)}{t+100}Q = 1$$

where $\lim_{t\to\infty} a(t) = 1$.

a. Let $K(t)$ be the concentration of salt at time t. Assuming that $Q(0) = Q_0$, can you guess the value of $\lim_{t\to\infty} K(t)$?
b. Use numerical methods to confirm your guess in the these cases:

$$\text{(i) } a(t) = t/(1+t) \quad \text{(ii) } a(t) = 1 - e^{-t^2} \quad \text{(iii) } a(t) = 1 + \sin(e^{-t}).$$

This page titled 13.2E: Cooling and Mixing (Exercises) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **4.2E: Cooling and Mixing (Exercises)** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

13.3: Elementary Mechanics

Newton's Second Law of Motion

In this section we consider an object with constant mass m moving along a line under a force F. Let $y = y(t)$ be the displacement of the object from a reference point on the line at time t, and let $v = v(t)$ and $a = a(t)$ be the velocity and acceleration of the object at time t. Thus, $v = y'$ and $a = v' = y''$, where the prime denotes differentiation with respect to t. Newton's second law of motion asserts that the force F and the acceleration a are related by the equation

$$F = ma. \tag{13.3.1}$$

Note: Units

In applications there are three main sets of units in use for length, mass, force, and time: the cgs, mks, and British systems. All three use the second as the unit of time. Table 3.3.1 shows the other units. Consistent with Equation 13.3.1, the unit of force in each system is defined to be the force required to impart an acceleration of (one unit of length)$/s^2$ to one unit of mass.

Table 3.3.1

Set	Length	Force	Mass
cgs	centimeter (cm)	dyne (d)	gram (g)
mks	meter (m)	newton (N)	kilogram (kg)
British	foot (ft)	pound (lb)	slug (sl)

If we assume that Earth is a perfect sphere with constant mass density, Newton's law of gravitation (discussed later in this section) asserts that the force exerted on an object by Earth's gravitational field is proportional to the mass of the object and inversely proportional to the square of its distance from the center of Earth. However, if the object remains sufficiently close to Earth's surface, we may assume that the gravitational force is constant and equal to its value at the surface. The magnitude of this force is mg, where g is called the *acceleration due to gravity*. (To be completely accurate, g should be called the *magnitude of the acceleration due to gravity at Earth's surface*.) This quantity has been determined experimentally. Approximate values of g are

$$\begin{aligned} g &= 980 \text{ cm/s}^2 && \text{(cgs)} \\ g &= 9.8 \text{ m/s}^2 && \text{(mks)} \\ g &= 32 \text{ ft/s}^2 && \text{(British).} \end{aligned}$$

In general, the force F in Equation 13.3.1 may depend upon t, y, and y'. Since $a = y''$, Equation 13.3.1 can be written in the form

$$my'' = F(t, y, y'), \tag{13.3.2}$$

which is a second order equation. We'll consider this equation with restrictions on F later; however, since Chapter 2 dealt only with first order equations, we consider here only problems in which Equation 13.3.2 can be recast as a first order equation. This is possible if F does not depend on y, so Equation 13.3.2 is of the form

$$my'' = F(t, y').$$

Letting $v = y'$ and $v' = y''$ yields a first order equation for v:

$$mv' = F(t, v). \tag{13.3.3}$$

Solving this equation yields v as a function of t. If we know $y(t_0)$ for some time t_0, we can integrate v to obtain y as a function of t.

Equations of the form Equation 13.3.3 occur in problems involving motion through a resisting medium.

Motion Through a Resisting Medium Under Constant Gravitational Force

Now we consider an object moving vertically in some medium. We assume that the only forces acting on the object are gravity and resistance from the medium. We also assume that the motion takes place close to Earth's surface and take the upward direction to be positive, so the gravitational force can be assumed to have the constant value $-mg$. We'll see that, under reasonable assumptions on the resisting force, the velocity approaches a limit as $t \to \infty$. We call this limit the *terminal velocity*.

Example 3.3.1

An object with mass m moves under constant gravitational force through a medium that exerts a resistance with magnitude proportional to the speed of the object. (Recall that the speed of an object is $|v|$, the absolute value of its velocity v.) Find the velocity of the object as a function of t, and find the terminal velocity. Assume that the initial velocity is v_0.

Solution

The total force acting on the object is

$$F = -mg + F_1, \tag{13.3.4}$$

where $-mg$ is the force due to gravity and F_1 is the resisting force of the medium, which has magnitude $k|v|$, where k is a positive constant. If the object is moving downward ($v \le 0$), the resisting force is upward (Figure 3.3.1a), so

$$F_1 = k|v| = k(-v) = -kv.$$

On the other hand, if the object is moving upward ($v \ge 0$), the resisting force is downward (Figure 3.3.1b , so

$$F_1 = -k|v| = -kv.$$

Thus, Equation 13.3.4 can be written as

$$F = -mg - kv, \tag{13.3.5}$$

regardless of the sign of the velocity.

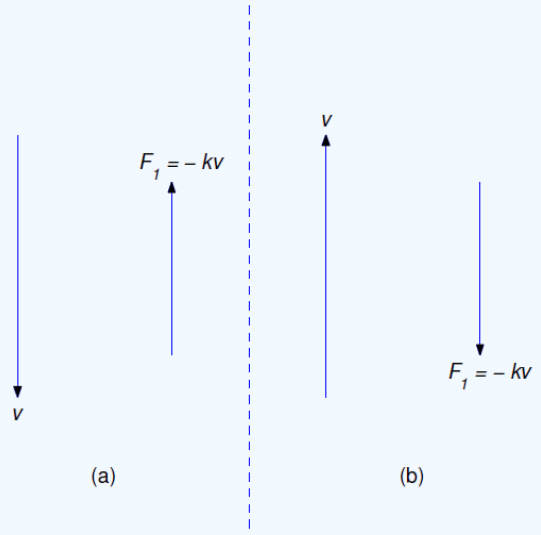


Figure 3.3.1 : Resistive forces

From Newton's second law of motion,

$$F = ma = mv',$$

so Equation 13.3.5 yields

$$mv' = -mg - kv,$$

or

$$v' + \frac{k}{m}v = -g. \tag{13.3.6}$$

Since $e^{-kt/m}$ is a solution of the complementary equation, the solutions of Equation 13.3.6 are of the form $v = ue^{-kt/m}$, where $u'e^{-kt/m} = -g$, so $u' = -ge^{kt/m}$. Hence,

$$u = -\frac{mg}{k}e^{kt/m} + c,$$

so

$$v = ue^{-kt/m} = -\frac{mg}{k} + ce^{-kt/m}. \tag{13.3.7}$$

Since $v(0) = v_0$,

$$v_0 = -\frac{mg}{k} + c,$$

so

$$c = v_0 + \frac{mg}{k}$$

and Equation 13.3.7 becomes

$$v = -\frac{mg}{k} + \left(v_0 + \frac{mg}{k}\right)e^{-kt/m}.$$

Letting $t \to \infty$ here shows that the terminal velocity is

$$\lim_{t\to\infty} v(t) = -\frac{mg}{k},$$

which is independent of the initial velocity v_0 (Figure 3.3.2).

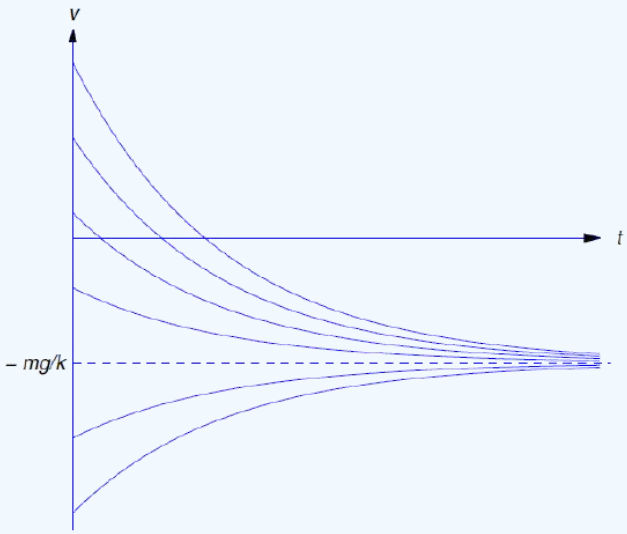


Figure 3.3.2 : Solutions of $mv' = -mg - kv$

Example 3.3.2 : Terminal Velocity

A 960-lb object is given an initial upward velocity of 60 ft/s near the surface of Earth. The atmosphere resists the motion with a force of 3 lb for each ft/s of speed. Assuming that the only other force acting on the object is constant gravity, find its velocity v as a function of t, and find its terminal velocity.

Solution

Since $mg = 960$ and $g = 32$, $m = 960/32 = 30$. The atmospheric resistance is $-3v$ lb if v is expressed in feet per second. Therefore

$$30v' = -960 - 3v,$$

which we rewrite as

$$v' + \frac{1}{10}v = -32.$$

Since $e^{-t/10}$ is a solution of the complementary equation, the solutions of this equation are of the form $v = ue^{-t/10}$, where $u'e^{-t/10} = -32$, so $u' = -32e^{t/10}$. Hence,

$$u = -320e^{t/10} + c,$$

so

$$v = ue^{-t/10} = -320 + ce^{-t/10}. \tag{13.3.8}$$

The initial velocity is 60 ft/s in the upward (positive) direction; hence, $v_0 = 60$. Substituting $t = 0$ and $v = 60$ in Equation 13.3.8 yields

$$60 = -320 + c,$$

so $c = 380$, and Equation 13.3.8 becomes

$$v = -320 + 380e^{-t/10} \text{ ft/s}$$

The terminal velocity is

$$\lim_{t\to\infty} v(t) = -320 \text{ ft/s}.$$

Example 3.3.3

A 10 kg mass is given an initial velocity $v_0 \le 0$ near Earth's surface. The only forces acting on it are gravity and atmospheric resistance proportional to the square of the speed. Assuming that the resistance is 8 N if the speed is 2 m/s, find the velocity of the object as a function of t, and find the terminal velocity.

Solution

Since the object is falling, the resistance is in the upward (positive) direction. Hence,

$$mv' = -mg + kv^2, \tag{13.3.9}$$

where k is a constant. Since the magnitude of the resistance is 8 N when $v = 2$ m/s,

$$k(2^2) = 8,$$

so $k = 2$ N-s^2/m^2. Since $m = 10$ and $g = 9.8$, Equation 13.3.9 becomes

$$10v' = -98 + 2v^2 = 2(v^2 - 49). \tag{13.3.10}$$

If $v_0 = -7$, then $v \equiv -7$ for all $t \ge 0$. If $v_0 \ne -7$, we separate variables to obtain

$$\frac{1}{v^2 - 49}v' = \frac{1}{5}, \tag{13.3.11}$$

which is convenient for the required partial fraction expansion

$$\frac{1}{v^2 - 49} = \frac{1}{(v-7)(v+7)} = \frac{1}{14}\left[\frac{1}{v-7} - \frac{1}{v+7}\right]. \tag{13.3.12}$$

Substituting Equation 13.3.12 into Equation 13.3.11 yields

$$\frac{1}{14}\left[\frac{1}{v-7}-\frac{1}{v+7}\right]v'=\frac{1}{5},$$

so

$$\left[\frac{1}{v-7}-\frac{1}{v+7}\right]v'=\frac{14}{5}.$$

Integrating this yields

$$\ln|v-7|-\ln|v+7|=14t/5+k.$$

Therefore

$$\left|\frac{v-7}{v+7}\right|=e^k e^{14t/5}.$$

Since Theorem 2.3.1 implies that $(v-7)/(v+7)$ cannot change sign (why?), we can rewrite the last equation as

$$\frac{v-7}{v+7}=ce^{14t/5}, \tag{13.3.13}$$

which is an implicit solution of Equation 13.3.10 Solving this for v yields

$$v=-7\frac{c+e^{-14t/5}}{c-e^{-14t/5}}. \tag{13.3.14}$$

Since $v(0)=v_0$, it Equation 13.3.13implies that

$$c=\frac{v_0-7}{v_0+7}.$$

Substituting this into Equation 13.3.14and simplifying yields

$$v=-7\frac{v_0(1+e^{-14t/5})-7(1-e^{-14t/5})}{v_0(1-e^{-14t/5})-7(1+e^{-14t/5})}.$$

Since $v_0\le 0$, v is defined and negative for all $t>0$. The terminal velocity is

$$\lim_{t\to\infty}v(t)=-7\text{ m/s},$$

independent of v_0. More generally, it can be shown (*Exercise 4.3.11*) that if v is any solution of Equation 13.3.9 such that $v_0\le 0$ then

$$\lim_{t\to\infty}v(t)=-\sqrt{\frac{mg}{k}}.$$

This is demonstrated in Figure 3.3.3 .

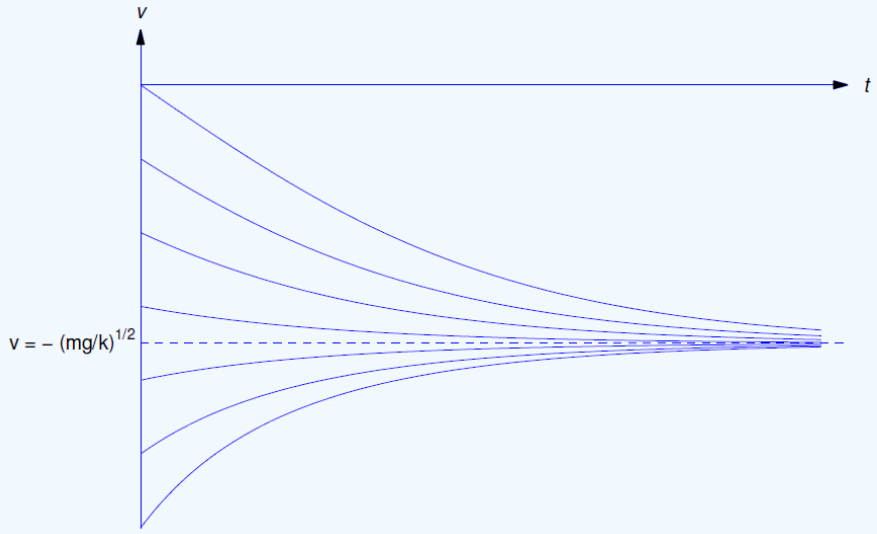


Figure 3.3.3 : Solutions of $mv' = -mg + kv^2,\ v(0) = v_0 \le 0$

Example 3.3.4

A 10-kg mass is launched vertically upward from Earth's surface with an initial velocity of v_0 m/s. The only forces acting on the mass are gravity and atmospheric resistance proportional to the square of the speed. Assuming that the atmospheric resistance is 8 N if the speed is 2 m/s, find the time T required for the mass to reach maximum altitude.

Solution

The mass will climb while $v > 0$ and reach its maximum altitude when $v = 0$. Therefore $v > 0$ for $0 \le t < T$ and $v(T) = 0$; therefore, we replace Equation 13.3.10by

$$10v' = -98 - 2v^2. \tag{13.3.15}$$

Separating variables yields

$$\frac{5}{v^2 + 49}v' = -1,$$

and integrating this yields

$$\frac{5}{7}\tan^{-1}\frac{v}{7} = -t + c.$$

(Recall that $\tan^{-1} u$ is the number θ such that $-\pi/2 < \theta < \pi/2$ and $\tan\theta = u$.) Since $v(0) = v_0$,

$$c = \frac{5}{7}\tan^{-1}\frac{v_0}{7},$$

so v is defined implicitly by

$$\frac{5}{7}\tan^{-1}\frac{v}{7} = -t + \frac{5}{7}\tan^{-1}\frac{v_0}{7}, \quad 0 \le t \le T. \tag{13.3.16}$$

Solving this for v yields

$$v = 7\tan\left(-\frac{7t}{5} + \tan^{-1}\frac{v_0}{7}\right). \tag{13.3.17}$$

Using the identity

$$\tan(A - B) = \frac{\tan A - \tan B}{1 + \tan A \tan B}$$

with $A=\tan^{-1}(v_0/7)$ and $B=7t/5$, and noting that $\tan(\tan^{-1}\theta)=\theta$, we can simplify Equation 13.3.17to

$$v=7\,\frac{v_0-7\tan(7t/5)}{7+v_0\tan(7t/5)}.$$

Since $v(T)=0$ and $\tan^{-1}(0)=0$, Equation 13.3.16implies that

$$-T+\frac{5}{7}\tan^{-1}\frac{v_0}{7}=0.$$

Therefore

$$T=\frac{5}{7}\tan^{-1}\frac{v_0}{7}.$$

Since $\tan^{-1}(v_0/7)<\pi/2$ for all v_0, the time required for the mass to reach its maximum altitude is less than

$$\frac{5\pi}{14}\approx 1.122\text{ s}$$

regardless of the initial velocity. Figure 3.3.4 shows graphs of v over $[0,T]$ for various values of v_0.

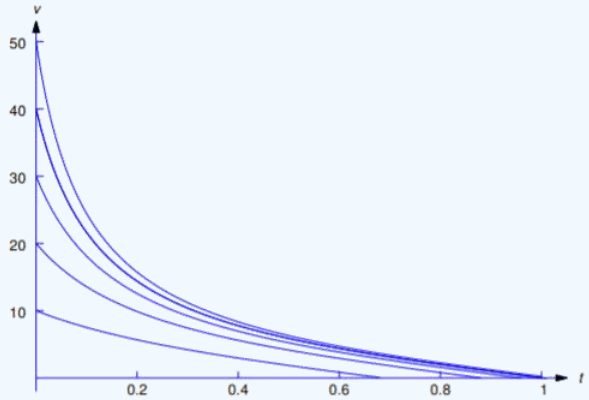


Figure 3.3.4 : Solutions of 3.3.15 for various $v_0>0$

Escape Velocity

Suppose a space vehicle is launched vertically and its fuel is exhausted when the vehicle reaches an altitude h above Earth, where h is sufficiently large so that resistance due to Earth's atmosphere can be neglected. Let $t=0$ be the time when burnout occurs. Assuming that the gravitational forces of all other celestial bodies can be neglected, the motion of the vehicle for $t>0$ is that of an object with constant mass m under the influence of Earth's gravitational force, which we now assume to vary inversely with the square of the distance from Earth's center; thus, if we take the upward direction to be positive then gravitational force on the vehicle at an altitude y above Earth is

$$F=-\frac{K}{(y+R)^2}, \tag{13.3.18}$$

where R is Earth's radius (Figure 3.3.5).

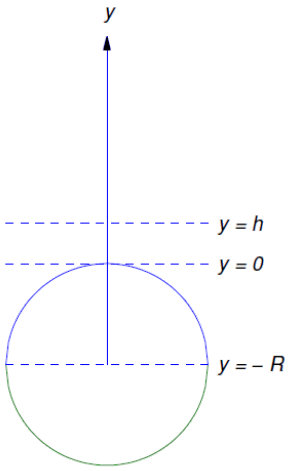


Figure 3.3.5 : Escape velocity

Since $F=-mg$ when $y=0$, setting $y=0$ in Equation 13.3.18yields

$$-mg=-\frac{K}{R^2};$$

therefore $K=mgR^2$ and Equation 13.3.18can be written more specifically as

$$F=-\frac{mgR^2}{(y+R)^2}. \tag{13.3.19}$$

From Newton's second law of motion,

$$F=m\frac{d^2y}{dt^2},$$

so Equation 13.3.19implies that

$$\frac{d^2y}{dt^2}=-\frac{gR^2}{(y+R)^2}. \tag{13.3.20}$$

We'll show that there's a number v_e, called the *escape velocity*, with these properties:

1. If $v_0\ge v_e$ then $v(t)>0$ for all $t>0$, and the vehicle continues to climb for all $t>0$; that is, it "escapes" Earth. (Is it really so obvious that $\lim_{t\to\infty}y(t)=\infty$ in this case? For a proof, see *Exercise 4.3.20*.)
2. If $v_0<v_e$ then $v(t)$ decreases to zero and becomes negative. Therefore the vehicle attains a maximum altitude y_m and falls back to Earth.

Since Equation 13.3.20is second order, we cannot solve it by methods discussed so far. However, we are concerned with v rather than y, and v is easier to find. Since $v=y'$ the chain rule implies that

$$\frac{d^2y}{dt^2}=\frac{dv}{dt}=\frac{dv}{dy}\frac{dy}{dt}=v\frac{dv}{dy}.$$

Substituting this into Equation 13.3.20yields the first order separable equation

$$v\frac{dv}{dy}=-\frac{gR^2}{(y+R)^2}. \tag{13.3.21}$$

When $t=0$, the velocity is v_0 and the altitude is h. Therefore we can obtain v as a function of y by solving the initial value problem

$$v\frac{dv}{dy}=-\frac{gR^2}{(y+R)^2},\quad v(h)=v_0.$$

Integrating Equation 13.3.21with respect to y yields

$$\frac{v^2}{2} = \frac{gR^2}{y+R} + c. \tag{13.3.22}$$

Since $v(h) = v_0$,

$$c = \frac{v_0^2}{2} - \frac{gR^2}{h+R},$$

so Equation 13.3.22becomes

$$\frac{v^2}{2} = \frac{gR^2}{y+R} + \left(\frac{v_0^2}{2} - \frac{gR^2}{h+R}\right). \tag{13.3.23}$$

If

$$v_0 \geq \left(\frac{2gR^2}{h+R}\right)^{1/2},$$

the parenthetical expression in Equation 13.3.23is nonnegative, so $v(y) > 0$ for $y > h$. This proves that there's an escape velocity v_e. We'll now prove that

$$v_e = \left(\frac{2gR^2}{h+R}\right)^{1/2}$$

by showing that the vehicle falls back to Earth if

$$v_0 < \left(\frac{2gR^2}{h+R}\right)^{1/2}. \tag{13.3.24}$$

If Equation 13.3.24holds then the parenthetical expression in Equation 13.3.23is negative and the vehicle will attain a maximum altitude $y_m > h$ that satisfies the equation

$$0 = \frac{gR^2}{y_m+R} + \left(\frac{v_0^2}{2} - \frac{gR^2}{h+R}\right).$$

The velocity will be zero at the maximum altitude, and the object will then fall to Earth under the influence of gravity.

Below is a video on solving a differential equation that models a falling object.



This page titled 13.3: Elementary Mechanics is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **4.3: Elementary Mechanics** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

13.3E: Elementary Mechanics (Exercises)

Except where directed otherwise, assume that the magnitude of the gravitational force on an object with mass m is constant and equal to mg. In exercises involving vertical motion take the upward direction to be positive.

Q4.3.1

1. A firefighter who weighs 192 lb slides down an infinitely long fire pole that exerts a frictional resistive force with magnitude proportional to his speed, with $k = 2.5$ lb-s/ft. Assuming that he starts from rest, find his velocity as a function of time and find his terminal velocity.

2. A firefighter who weighs 192 lb slides down an infinitely long fire pole that exerts a frictional resistive force with magnitude proportional to her speed, with constant of proportionality k. Find k, given that her terminal velocity is -16 ft/s, and then find her velocity v as a function of t. Assume that she starts from rest.

3. A boat weighs $64,000$ lb. Its propellor produces a constant thrust of $50,000$ lb and the water exerts a resistive force with magnitude proportional to the speed, with $k = 2000$ lb-s/ft. Assuming that the boat starts from rest, find its velocity as a function of time, and find its terminal velocity.

4. A constant horizontal force of 10 N pushes a 20 kg-mass through a medium that resists its motion with $.5$ N for every m/s of speed. The initial velocity of the mass is 7 m/s in the direction opposite to the direction of the applied force. Find the velocity of the mass for $t > 0$.

5. A stone weighing $1/2$ lb is thrown upward from an initial height of 5 ft with an initial speed of 32 ft/s. Air resistance is proportional to speed, with $k = 1/128$ lb-s/ft. Find the maximum height attained by the stone.

6. A 3200-lb car is moving at 64 ft/s down a 30-degree grade when it runs out of fuel. Find its velocity after that if friction exerts a resistive force with magnitude proportional to the square of the speed, with $k = 1$ lb-s$^2/$ft^2. Also find its terminal velocity.

7. A 96 lb weight is dropped from rest in a medium that exerts a resistive force with magnitude proportional to the speed. Find its velocity as a function of time if its terminal velocity is -128 ft/s.

8. An object with mass m moves vertically through a medium that exerts a resistive force with magnitude proportional to the speed. Let $y = y(t)$ be the altitude of the object at time t, with $y(0) = y_0$. Use the results of Example 4.3.1 to show that

$$y(t) = y_0 + \frac{m}{k}(v_0 - v - gt).$$

9. An object with mass m is launched vertically upward with initial velocity v_0 from Earth's surface ($y_0 = 0$) in a medium that exerts a resistive force with magnitude proportional to the speed. Find the time T when the object attains its maximum altitude y_m. Then use the result of **Exercise 4.3.8** to find y_m.

10. An object weighing 256 lb is dropped from rest in a medium that exerts a resistive force with magnitude proportional to the square of the speed. The magnitude of the resisting force is 1 lb when $|v| = 4$ ft/s. Find v for $t > 0$, and find its terminal velocity.

11. An object with mass m is given an initial velocity $v_0 \le 0$ in a medium that exerts a resistive force with magnitude proportional to the square of the speed. Find the velocity of the object for $t > 0$, and find its terminal velocity.

12. An object with mass m is launched vertically upward with initial velocity v_0 in a medium that exerts a resistive force with magnitude proportional to the square of the speed.

a. Find the time T when the object reaches its maximum altitude.
b. Use the result of **Exercise 4.3.11** to find the velocity of the object for $t > T$.

13. An object with mass m is given an initial velocity $v_0 \le 0$ in a medium that exerts a resistive force of the form $a|v|/(1 + |v|)$, where a is positive constant.

a. Set up a differential equation for the speed of the object.
b. Use your favorite numerical method to solve the equation you found in (a), to convince yourself that there's a unique number a_0 such that $\lim_{t\to\infty} s(t) = \infty$ if $a \le a_0$ and $\lim_{t\to\infty} s(t)$ exists (finite) if $a > a_0$. (We say that a_0 is the **bifurcation value** of a.) Try to find a_0 and $\lim_{t\to\infty} s(t)$ in the case where $a > a_0$.

14. An object of mass m falls in a medium that exerts a resistive force $f = f(s)$, where $s = |v|$ is the speed of the object. Assume that $f(0) = 0$ and f is strictly increasing and differentiable on $(0, \infty)$.

a. Write a differential equation for the speed $s = s(t)$ of the object. Take it as given that all solutions of this equation with $s(0) \geq 0$ are defined for all $t > 0$ (which makes good sense on physical grounds).
b. Show that if $\lim_{s\to\infty} f(s) \leq mg$ then $\lim_{t\to\infty} s(t) = \infty$.
c. Show that if $\lim_{s\to\infty} f(s) > mg$ then $\lim_{t\to\infty} s(t) = s_T$ (terminal speed), where $f(s_T) = mg$..

15. A 100-g mass with initial velocity $v_0 \leq 0$ falls in a medium that exerts a resistive force proportional to the fourth power of the speed. The resistance is .1 N if the speed is 3 m/s.

a. Set up the initial value problem for the velocity v of the mass for $t > 0$.
b. Use **Exercise 4.3.14** (c) to determine the terminal velocity of the object.
c. To confirm your answer to (b), use one of the numerical methods studied in Chapter 3 to compute approximate solutions on $[0, 1]$ (seconds) of the initial value problem of (a) , with initial values $v_0 = 0$, -2, -4, ..., -12. Present your results in graphical form similar to Figure 4.3.3.

16. A 64-lb object with initial velocity $v_0 \leq 0$ falls through a dense fluid that exerts a resistive force proportional to the square root of the speed. The resistance is 64 lb if the speed is 16 ft/s.

a. Set up the initial value problem for the velocity v of the mass for $t > 0$.
b. Use **Exercise 4.3.14** (c) to determine the terminal velocity of the object.
c. To confirm your answer to (b), use one of the numerical methods studied in Chapter 3 to compute approximate solutions on $[0, 4]$ (seconds) of the initial value problem of (a), with initial values $v_0 = 0$, -5, -10, ..., -30. Present your results in graphical form similar to Figure 4.3.3.

Q4.3.2

In **Exercises 4.3.17-4.3.20**, assume that the force due to gravity is given by Newton's law of gravitation. Take the upward direction to be positive.

17. A space probe is to be launched from a space station 200 miles above Earth. Determine its escape velocity in miles/s. Take Earth's radius to be 3960 miles.

18. A space vehicle is to be launched from the moon, which has a radius of about 1080 miles. The acceleration due to gravity at the surface of the moon is about 5.31 ft/s^2. Find the escape velocity in miles/s.

19.

a. Show that (Equation 4.3.27) can be rewritten as

$$v^2 = \frac{h - y}{y + R} v_e^2 + v_0^2.$$

b. Show that if $v_0 = \rho v_e$ with $0 \leq \rho < 1$, then the maximum altitude y_m attained by the space vehicle is

$$y_m = \frac{h + R\rho^2}{1 - \rho^2}.$$

c. By requiring that $v(y_m) = 0$, use (Equation 4.3.26) to deduce that if $v_0 < v_e$ then

$$|v| = v_e \left[\frac{(1 - \rho^2)(y_m - y)}{y + R} \right]^{1/2},$$

where y_m and ρ are as defined in (b) and $y \geq h$.
d. Deduce from (c) that if $v < v_e$, the vehicle takes equal times to climb from $y = h$ to $y = y_m$ and to fall back from $y = y_m$ to $y = h$.

20. In the situation considered in the discussion of escape velocity, show that $\lim_{t\to\infty} y(t) = \infty$ if $v(t) > 0$ for all $t > 0$.

This page titled 13.3E: Elementary Mechanics (Exercises) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **4.3E: Elementary Mechanics (Exercises)** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

CHAPTER OVERVIEW

14: Higher order linear ODEs

We have already studied the basics of differential equations, including separable first-order equations. In this chapter, we go a little further and look at second-order equations, which are equations containing second derivatives of the dependent variable. The solution methods we examine are different from those discussed earlier, and the solutions tend to involve trigonometric functions as well as exponential functions. Here we concentrate primarily on second-order equations with constant coefficients.

14.1: Second order linear ODEs
14.2: The Method of Undetermined Coefficients I
14.2E: The Method of Undetermined Coefficients I (Exercises)
14.3: The Method of Undetermined Coefficients II
14.3E: The Method of Undetermined Coefficients II (Exercises)
14.4: Constant coefficient second order linear ODEs
14.5: Higher order linear ODEs
14.6: Reduction of Order
14.6E: Reduction of Order (Exercises)
14.7: Variation of Parameters
14.7E: Variation of Parameters (Exercises)
14.8: Mechanical Vibrations
14.9: Nonhomogeneous Equations
14.10: Forced Oscillations and Resonance
14.E: Higher order linear ODEs (Exercises)

This page titled 14: Higher order linear ODEs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

14.1: Second order linear ODEs

Let us consider the general second order linear differential equation

$$A(x)y'' + B(x)y' + C(x)y = F(x).$$

We usually divide through by $A(x)$ to get

$$y'' + p(x)y' + q(x)y = f(x),$$

where $p(x) = \frac{B(x)}{A(x)}$, $q(x) = \frac{C(x)}{A(x)}$, and $f(x) = \frac{F(x)}{A(x)}$. The word *linear* means that the equation contains no powers nor functions of y, y', and y''.

In the special case when $f(x) = 0$ we have a so-called *homogeneous equation*

$$y'' + p(x)y' + q(x)y = 0, \tag{14.1.1}$$

We have already seen some second order linear homogeneous equations:

$$\begin{aligned} y'' + k^2 y = 0 \quad &\text{Two solutions are:} \quad y_1 = \cos(kx), \quad y_2 = \sin(kx). \\ y'' - k^2 y = 0 \quad &\text{Two solutions are:} \quad y_1 = e^{kx}, \quad y_2 = e^{-kx}. \end{aligned}$$

If we know two solutions of a linear homogeneous equation, we know a lot more of them.

Theorem: Superposition

Suppose y_1 and y_2 are two solutions of the homogeneous equation (14.1.1). Then

$$y(x) = C_1 y_1(x) + C_2 y_2(x),$$

also solves (14.1.1) for arbitrary constants C_1 and C_2.

That is, we can add solutions together and multiply them by constants to obtain new and different solutions. We call the expression $C_1 y_1 + C_2 y_2$ a *linear combination* of y_1 and y_2. Let us prove this theorem; the proof is very enlightening and illustrates how linear equations work.

Proof

Let $y = C_1 y_1 + C_2 y_2$. Then

$$\begin{aligned} y'' + py' + qy &= (C_1 y_1 + C_2 y_2)'' + p(C_1 y_1 + C_2 y_2)' + q(C_1 y_1 + C_2 y_2) \\ &= C_1 y_1'' + C_2 y_2'' + C_1 p y_1' + C_2 p y_2' + C_1 q y_1 + C_2 q y_2 \\ &= C_1(y_1'' + p y_1' + q y_1) + C_2(y_2'' + p y_2' + q y_2) \\ &= C_1 . 0 + C_2 . 0 = 0 \end{aligned}$$

The proof becomes even simpler to state if we use the operator notation. An *operator* is an object that eats functions and spits out functions (kind of like what a function, which eats numbers and spits out numbers). Define the operator L by

$$Ly = y'' + py' + qy.$$

The differential equation now becomes $Ly = 0$. The operator (and the equation) L being linear means that $L(C_1 y_1 + C_2 y_2) = C_1 L y_1 + C_2 L y_2$. The proof above becomes

$$Ly = L(C_1 y_1 + C_2 y_2) = C_1 L y_1 + C_2 L y_2 = C_1 . 0 + C_2 . 0 = 0$$

Two different solutions to the second equation $y'' - k^2 y = 0$ are $y_1 = \cosh(kx)$ and $y_2 = \sinh(kx)$. Let us remind ourselves of the definition, $\cosh x = \frac{e^x + e^{-x}}{2}$ and $\sinh x = \frac{e^x - e^{-x}}{2}$. Therefore, these are solutions by superposition as they are linear combinations of the two exponential solutions.

The functions sinh and cosh are sometimes more convenient to use than the exponential. Let us review some of their properties.

$$\cosh 0 = 1 \qquad \sinh 0 = 0,$$
$$\frac{d}{dx}[\cosh x] = \sinh x, \qquad \frac{d}{dx}[\sinh x] = \cosh x,$$
$$\cosh^2 x - \sinh^2 x = 1.$$

Below is a video on finding the constants of the exponential solution to a differential equation



Exercise

Derive these properties using the definitions of sinh and cosh in terms of exponentials.

Linear equations have nice and simple answers to the existence and uniqueness question.

Theorem: Existence and Uniqueness

Suppose $p(x)$, $q(x)$, and $f(x)$ are continuous functions on some interval I containing a with a, b_0 and b_1 constants. The equation

$$y'' + p(x)y' + q(x)y = f(x).$$

has exactly *one* solution $y(x)$ defined on the same interval I satisfying the initial conditions

$$y(a) = b_0, \quad y'(a) = b_1.$$

For example, the equation $y'' + k^2 y = 0$ with $y(0) = b_0$ and $y'(0) = b_1$ has the solution

$$y(x) = b_0 \cos(kx) + \frac{b_1}{k}\sin(kx)$$

The equation $y'' - k^2 y = 0$ with $y(0) = b_0$ and $y'(0) = b_1$ has the solution

$$y(x) = b_0 \cosh(kx) + \frac{b_1}{k}\sinh(kx)$$

Using cosh and sinh in this solution allows us to solve for the initial conditions in a cleaner way than if we have used the exponentials.

The initial conditions for a second order ODE consist of two equations. Common sense tells us that if we have two arbitrary constants and two equations, then we should be able to solve for the constants and find a solution to the differential equation satisfying the initial conditions.

Exercise

Suppose we find two different solutions y_1 and y_2 to the homogeneous equation (14.1.1). Can every solution be written (using superposition) in the form $y = C_1 y_1 + C_2 y_2$?

Answer

Answer is affirmative! Provided that y_1 and y_2 are different enough in the following sense. We will say y_1 and y_2 are linearly independent if one is not a constant multiple of the other.

Theorem

Let $p(x)$ and $q(x)$ be continuous functions and let y_1 and y_2 be two linearly independent solutions to the homogeneous equation (14.1.1). Then every other solution is of the form

$$y = C_1 y_1 + C_2 y_2.$$

That is, $y = C_1 y_1 + C_2 y_2$ is the general solution.

For example, we found the solutions $y_1 = \sin x$ and $y_2 = \cos x$ for the equation $y'' + y = 0$. It is not hard to see that sine and cosine are not constant multiples of each other. If $\sin x = A \cos x$ for some constant A, we let $x = 0$ and this would imply $A = 0$. But then $\sin x = 0$ for all x, which is preposterous. So y_1 and y_2 are linearly independent. Hence,

$$y = C_1 \cos x + C_2 \sin x$$

is the general solution to $y'' + y = 0$.

For two functions, checking linear independence is rather simple. Let us see another example. Consider $y'' - 2x^{-2}y = 0$. Then $y_1 = x^2$ and $y_2 = \frac{1}{x}$ are solutions. To see that they are linearly independent, suppose one is a multiple of the other: $y_1 = Ay_2$, we just have to find out that A cannot be a constant. In this case we have $A = \frac{y_1}{y_2} = x^3$, this most decidedly not a constant. So $y = C_1 x^2 + C_2 \frac{1}{x}$ is the general solution.

If you have one solution to a second order linear homogeneous equation, then you can find another one. This is the *reduction of order method*. The idea is that if we somehow found y_1 as a solution of $y'' + p(x)y' + q(x)y = 0$ we try a second solution of the form $y_2(x) = y_1(x)v(x)$. We just need to find v. We plug y_2 into the equation:

$$\begin{aligned} 0 = y_2'' + p(x)y_2' + q(x)y_2 \quad &= y_1''v + 2y_1'v' + y_1v'' + p(x)(y_1'v + y_1v') + q(z)y_1v \\ &= y_1v'' + (2y_1' + p(x)y_1)v' + \cancelto{0}{\left(y_1'' + p(x)y_1' + q(x)y_1\right)}v. \end{aligned} \tag{14.1.2}$$

In other words, $y_1v'' + (2y_1' + p(x)y_1)v' = 0$. Using $w = v'$ we have the first order linear equation $y_1w' + (2y_1' + p(x)y_1)w = 0$. After solving this equation for w (integrating factor), we find v by antidifferentiating w. We then form y_2 by computing y_1v. For example, suppose we somehow know $y_1 = x$ is a solution to $y'' + x^{-1}y' - x^{-2}y = 0$. The equation for w is then $xw' + 3w = 0$. We find a solution, $w = Cx^{-3}$, and we find an antiderivative $v = \frac{-C}{2x^2}$. Hence $y_2 = y_1v = \frac{-C}{2x}$. Any C works and so $C = -2$ makes $y_2 = \frac{1}{x}$. Thus, the general solution is $y = C_1x + C_2\frac{1}{x}$.

Since we have a formula for the solution to the first order linear equation, we can write a formula for y_2:

$$y_2(x) = y_1(x) \int \frac{e^{-\int p(x)\,dx}}{\left(y_1(x)\right)^2}\,dx$$

However, it is much easier to remember that we just need to try $y_2(x) = y_1(x)v(x)$ and find $v(x)$ as we did above. Also, the technique works for higher order equations too: you get to reduce the order for each solution you find. So it is better to remember how to do it rather than a specific formula.

We will study the solution of nonhomogeneous equations in Section 2.5. We will first focus on finding general solutions to homogeneous equations.

This page titled 14.1: Second order linear ODEs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **2.1: Second order linear ODEs** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

14.2: The Method of Undetermined Coefficients I

In this section we consider the constant coefficient equation

$$ay'' + by' + cy = e^{\alpha x} G(x), \tag{14.2.1}$$

where α is a constant and G is a polynomial.

From Theorem 5.3.2, the general solution of Equation 14.2.1 is $y = y_p + c_1 y_1 + c_2 y_2$, where y_p is a particular solution of Equation 14.2.1 and $\{y_1, y_2\}$ is a fundamental set of solutions of the complementary equation

$$ay'' + by' + cy = 0.$$

In Section 5.2 we showed how to find $\{y_1, y_2\}$. In this section we'll show how to find y_p. The procedure that we'll use is called *the method of undetermined coefficients*. Our first example is similar to *Exercises 5.3.16-5.3.21*.

Example 4.2.1 :

Find a particular solution of

$$y'' - 7y' + 12y = 4e^{2x}. \tag{14.2.2}$$

Then find the general solution.

Solution

Substituting $y_p = Ae^{2x}$ for y in Equation 14.2.2 will produce a constant multiple of Ae^{2x} on the left side of Equation 14.2.2, so it may be possible to choose A so that y_p is a solution of Equation 14.2.2. Let's try it; if $y_p = Ae^{2x}$ then

$$y_p'' - 7y_p' + 12y_p = 4Ae^{2x} - 14Ae^{2x} + 12Ae^{2x} = 2Ae^{2x} = 4e^{2x}$$

if $A = 2$. Therefore $y_p = 2e^{2x}$ is a particular solution of Equation 14.2.2. To find the general solution, we note that the characteristic polynomial of the complementary equation

$$y'' - 7y' + 12y = 0 \tag{14.2.3}$$

is $p(r) = r^2 - 7r + 12 = (r-3)(r-4)$, so $\{e^{3x}, e^{4x}\}$ is a fundamental set of solutions of Equation 14.2.3. Therefore the general solution of Equation 14.2.2 is

$$y = 2e^{2x} + c_1 e^{3x} + c_2 e^{4x}.$$

Below is a video on the method of undetermined coefficients



Example 4.2.2

Find a particular solution of

$$y'' - 7y' + 12y = 5e^{4x}. \tag{14.2.4}$$

Then find the general solution.

Solution

Fresh from our success in finding a particular solution of Equation 14.2.2— where we chose $y_p = Ae^{2x}$ because the right side of Equation 14.2.2 is a constant multiple of e^{2x} — it may seem reasonable to try $y_p = Ae^{4x}$ as a particular solution of Equation 14.2.4. However, this will not work, since we saw in Example 4.2.1 that e^{4x} is a solution of the complementary equation Equation 14.2.3, so substituting $y_p = Ae^{4x}$ into the left side of Equation 14.2.4 produces zero on the left, no matter how we choose A. To discover a suitable form for y_p, we use the same approach that we used in Section 5.2 to find a second solution of

$$ay'' + by' + cy = 0$$

in the case where the characteristic equation has a repeated real root: we look for solutions of Equation 14.2.4in the form $y = ue^{4x}$, where u is a function to be determined. Substituting

$$y = ue^{4x}, \quad y' = u'e^{4x} + 4ue^{4x}, \quad \text{and} \quad y'' = u''e^{4x} + 8u'e^{4x} + 16ue^{4x} \tag{14.2.5}$$

into Equation 14.2.4and canceling the common factor e^{4x} yields

$$(u'' + 8u' + 16u) - 7(u' + 4u) + 12u = 5,$$

or

$$u'' + u' = 5.$$

By inspection we see that $u_p = 5x$ is a particular solution of this equation, so $y_p = 5xe^{4x}$ is a particular solution of Equation 14.2.4 Therefore

$$y = 5xe^{4x} + c_1e^{3x} + c_2e^{4x}$$

is the general solution.

Below is a video on using the method of undetermined coefficients to solve a differential equation.



Example 4.2.3

Find a particular solution of

$$y'' - 8y' + 16y = 2e^{4x}. \tag{14.2.6}$$

Solution

Since the characteristic polynomial of the complementary equation

$$y'' - 8y' + 16y = 0 \tag{14.2.7}$$

is $p(r)=r^2-8r+16=(r-4)^2$, both $y_1=e^{4x}$ and $y_2=xe^{4x}$ are solutions of Equation 14.2.7. Therefore Equation 14.2.6 does not have a solution of the form $y_p=Ae^{4x}$ or $y_p=Axe^{4x}$. As in Example 4.2.2 , we look for solutions of Equation 14.2.6 in the form $y=ue^{4x}$, where u is a function to be determined. Substituting from Equation 14.2.5 into Equation 14.2.6 and canceling the common factor e^{4x} yields

$$(u''+8u'+16u)-8(u'+4u)+16u=2,$$

or

$$u''=2.$$

Integrating twice and taking the constants of integration to be zero shows that $u_p=x^2$ is a particular solution of this equation, so $y_p=x^2e^{4x}$ is a particular solution of Equation 14.2.4 Therefore

$$y=e^{4x}(x^2+c_1+c_2x)$$

is the general solution.

The preceding examples illustrate the following facts concerning the form of a particular solution y_p of a constant coefficent equation

$$ay''+by'+cy=ke^{\alpha x},$$

where k is a nonzero constant:

a. If $e^{\alpha x}$ isn't a solution of the complementary equation

$$ay''+by'+cy=0, \tag{14.2.8}$$

then $y_p=Ae^{\alpha x}$, where A is a constant. (See Example 4.2.1).

b. If $e^{\alpha x}$ is a solution of Equation 14.2.8 but $xe^{\alpha x}$ is not, then $y_p=Axe^{\alpha x}$, where A is a constant. (See Example 4.2.2 .)

c. If both $e^{\alpha x}$ and $xe^{\alpha x}$ are solutions of Equation 14.2.8, then $y_p=Ax^2e^{\alpha x}$, where A is a constant. (See Example 4.2.3 .)

See *Exercise 5.4.30* for the proofs of these facts.

In all three cases you can just substitute the appropriate form for y_p and its derivatives directly into

$$ay_p''+by_p'+cy_p=ke^{\alpha x},$$

and solve for the constant A, as we did in Example 4.2.1 . (See *Exercises 5.4.31-5.4.33*.) However, if the equation is

$$ay''+by'+cy=ke^{\alpha x}G(x),$$

where G is a polynomial of degree greater than zero, we recommend that you use the substitution $y=ue^{\alpha x}$ as we did in Examples 4.2.2 and 4.2.3 . The equation for u will turn out to be

$$au''+p'(\alpha)u'+p(\alpha)u=G(x), \tag{14.2.9}$$

where $p(r)=ar^2+br+c$ is the characteristic polynomial of the complementary equation and $p'(r)=2ar+b$ (*Exercise 5.4.30*); however, you shouldn't memorize this since it is easy to derive the equation for u in any particular case. Note, however, that if $e^{\alpha x}$ is a solution of the complementary equation then $p(\alpha)=0$, so Equation 14.2.9 reduces to

$$au''+p'(\alpha)u'=G(x),$$

while if both $e^{\alpha x}$ and $xe^{\alpha x}$ are solutions of the complementary equation then $p(r)=a(r-\alpha)^2$ and $p'(r)=2a(r-\alpha)$, so $p(\alpha)=p'(\alpha)=0$ and Equation 14.2.9 reduces to

$$au''=G(x).$$

Example 4.2.4

Find a particular solution of

$$y''-3y'+2y=e^{3x}(-1+2x+x^2). \tag{14.2.10}$$

Solution

Substituting

$$y = ue^{3x}, \quad y' = u'e^{3x} + 3ue^{3x}, \quad \text{and} y'' = u''e^{3x} + 6u'e^{3x} + 9ue^{3x}$$

into Equation 14.2.10 and canceling e^{3x} yields

$$(u'' + 6u' + 9u) - 3(u' + 3u) + 2u = -1 + 2x + x^2,$$

or

$$u'' + 3u' + 2u = -1 + 2x + x^2. \tag{14.2.11}$$

As in Example 5.3.2, in order to guess a form for a particular solution of Equation 14.2.11, we note that substituting a second degree polynomial $u_p = A + Bx + Cx^2$ for u in the left side of Equation 14.2.11 produces another second degree polynomial with coefficients that depend upon A, B, and C; thus,

$$\text{if} \quad u_p = A + Bx + Cx^2 \quad \text{then} \quad u_p' = B + 2Cx \quad \text{and} \quad u_p'' = 2C.$$

If u_p is to satisfy Equation 14.2.11, we must have

$$\begin{aligned} u_p'' + 3u_p' + 2u_p &= 2C + 3(B + 2Cx) + 2(A + Bx + Cx^2) \\ &= (2C + 3B + 2A) + (6C + 2B)x + 2Cx^2 = -1 + 2x + x^2. \end{aligned}$$

Equating coefficients of like powers of x on the two sides of the last equality yields

$$\begin{aligned} 2C &= 1 \\ 2B + 6C &= 2 \\ 2A + 3B + 2C &= -1. \end{aligned}$$

Solving these equations for C, B, and A (in that order) yields $C = 1/2, B = -1/2, A = -1/4$. Therefore

$$u_p = -\frac{1}{4}(1 + 2x - 2x^2)$$

is a particular solution of Equation 14.2.11, and

$$y_p = u_p e^{3x} = -\frac{e^{3x}}{4}(1 + 2x - 2x^2)$$

is a particular solution of Equation 14.2.10.

Example 4.2.5

Find a particular solution of

$$y'' - 4y' + 3y = e^{3x}(6 + 8x + 12x^2). \tag{14.2.12}$$

Solution

Substituting

$$y = ue^{3x}, \quad y' = u'e^{3x} + 3ue^{3x}, \quad \text{and } y'' = u''e^{3x} + 6u'e^{3x} + 9ue^{3x}$$

into Equation 14.2.12 and canceling e^{3x} yields

$$(u'' + 6u' + 9u) - 4(u' + 3u) + 3u = 6 + 8x + 12x^2,$$

or

$$u'' + 2u' = 6 + 8x + 12x^2. \tag{14.2.13}$$

There's no u term in this equation, since e^{3x} is a solution of the complementary equation for Equation 14.2.12. (See *Exercise 5.4.30.*) Therefore Equation 14.2.13 does not have a particular solution of the form $u_p = A + Bx + Cx^2$ that we used successfully in Example 4.2.4 , since with this choice of u_p,

$$u_p'' + 2u_p' = 2C + (B + 2Cx)$$

can't contain the last term ($12x^2$) on the right side of Equation 14.2.13. Instead, let's try $u_p = Ax + Bx^2 + Cx^3$ on the grounds that

$$u_p' = A + 2Bx + 3Cx^2 \quad \text{and} \quad u_p'' = 2B + 6Cx$$

together contain all the powers of x that appear on the right side of Equation 14.2.13.

Substituting these expressions in place of u' and u'' in Equation 14.2.13 yields

$$(2B + 6Cx) + 2(A + 2Bx + 3Cx^2) = (2B + 2A) + (6C + 4B)x + 6Cx^2 = 6 + 8x + 12x^2.$$

Comparing coefficients of like powers of x on the two sides of the last equality shows that u_p satisfies Equation 14.2.13 if

$$\begin{aligned} 6C &= 12 \\ 4B + 6C &= 8 \\ 2A + 2B &= 6. \end{aligned}$$

Solving these equations successively yields $C = 2$, $B = -1$, and $A = 4$. Therefore

$$u_p = x(4 - x + 2x^2)$$

is a particular solution of Equation 14.2.13, and

$$y_p = u_p e^{3x} = xe^{3x}(4 - x + 2x^2)$$

is a particular solution of Equation 14.2.12.

Example 4.2.6

Find a particular solution of

$$4y'' + 4y' + y = e^{-x/2}(-8 + 48x + 144x^2). \tag{14.2.14}$$

Solution

Substituting

$$y = ue^{-x/2}, \quad y' = u'e^{-x/2} - \frac{1}{2}ue^{-x/2}, \quad \text{and} \quad y'' = u''e^{-x/2} - u'e^{-x/2} + \frac{1}{4}ue^{-x/2}$$

into Equation 14.2.14 and canceling $e^{-x/2}$ yields

$$4\left(u'' - u' + \frac{u}{4}\right) + 4\left(u' - \frac{u}{2}\right) + u = 4u'' = -8 + 48x + 144x^2,$$

or

$$u'' = -2 + 12x + 36x^2, \tag{14.2.15}$$

which does not contain u or u' because $e^{-x/2}$ and $xe^{-x/2}$ are both solutions of the complementary equation. (See *Exercise 5.4.30*.) To obtain a particular solution of Equation 14.2.15 we integrate twice, taking the constants of integration to be zero; thus,

$$u_p' = -2x + 6x^2 + 12x^3 \quad \text{and} \quad u_p = -x^2 + 2x^3 + 3x^4 = x^2(-1 + 2x + 3x^2).$$

Therefore

$$y_p = u_p e^{-x/2} = x^2 e^{-x/2}(-1 + 2x + 3x^2)$$

is a particular solution of Equation 14.2.14.

Below is a video on using the method of undetermined coefficients to solve a nonhomogeneous differential equation.



Summary

The preceding examples illustrate the following facts concerning particular solutions of a constant coefficent equation of the form

$$ay'' + by' + cy = e^{\alpha x} G(x),$$

where G is a polynomial (see *Exercise 5.4.30*):

a. If $e^{\alpha x}$ isn't a solution of the complementary equation

$$ay'' + by' + cy = 0, \tag{14.2.16}$$

then $y_p = e^{\alpha x} Q(x)$, where Q is a polynomial of the same degree as G. (See Example 4.2.4).

b. If $e^{\alpha x}$ is a solution of Equation 14.2.16but $xe^{\alpha x}$ is not, then $y_p = xe^{\alpha x} Q(x)$, where Q is a polynomial of the same degree as G. (See Example 4.2.5 .)

c. If both $e^{\alpha x}$ and $xe^{\alpha x}$ are solutions of Equation 14.2.16, then $y_p = x^2 e^{\alpha x} Q(x)$, where Q is a polynomial of the same degree as G. (See Example 4.2.6 .)

In all three cases, you can just substitute the appropriate form for y_p and its derivatives directly into

$$ay_p'' + by_p' + cy_p = e^{\alpha x} G(x),$$

and solve for the coefficients of the polynomial Q. However, if you try this you will see that the computations are more tedious than those that you encounter by making the substitution $y = ue^{\alpha x}$ and finding a particular solution of the resulting equation for u. (See *Exercises 5.4.34-5.4.36*.) In Case (a) the equation for u will be of the form

$$au'' + p'(\alpha)u' + p(\alpha)u = G(x),$$

with a particular solution of the form $u_p = Q(x)$, a polynomial of the same degree as G, whose coefficients can be found by the method used in Example 4.2.4 . In Case (b) the equation for u will be of the form

$$au'' + p'(\alpha)u' = G(x)$$

(no u term on the left), with a particular solution of the form $u_p = xQ(x)$, where Q is a polynomial of the same degree as G whose coefficents can be found by the method used in Example 4.2.5 . In Case (c), the equation for u will be of the form

$$au'' = G(x)$$

with a particular solution of the form $u_p = x^2 Q(x)$ that can be obtained by integrating $G(x)/a$ twice and taking the constants of integration to be zero, as in Example 4.2.6 .

Using the Principle of Superposition

The next example shows how to combine the method of undetermined coefficients and Theorem 5.3.3, the principle of superposition.

Example 4.2.7

Find a particular solution of

$$y'' - 7y' + 12y = 4e^{2x} + 5e^{4x}. \tag{14.2.17}$$

Solution

In Example 4.2.1 we found that $y_{p_1} = 2e^{2x}$ is a particular solution of

$$y'' - 7y' + 12y = 4e^{2x},$$

and in Example 4.2.2 we found that $y_{p_2} = 5xe^{4x}$ is a particular solution of

$$y'' - 7y' + 12y = 5e^{4x}.$$

Therefore the principle of superposition implies that $y_p = 2e^{2x} + 5xe^{4x}$ is a particular solution of Equation 14.2.17).

This page titled 14.2: The Method of Undetermined Coefficients I is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **5.4: The Method of Undetermined Coefficients I** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

14.2E: The Method of Undetermined Coefficients I (Exercises)

Q5.4.1

In **Exercises 5.4.1-5.4.14** find a particular solution.

1. $y''-3y'+2y=e^{3x}(1+x)$
2. $y''-6y'+5y=e^{-3x}(35-8x)$
3. $y''-2y'-3y=e^{x}(-8+3x)$
4. $y''+2y'+y=e^{2x}(-7-15x+9x^2)$
5. $y''+4y=e^{-x}(7-4x+5x^2)$
6. $y''-y'-2y=e^{x}(9+2x-4x^2)$
7. $y''-4y'-5y=-6xe^{-x}$
8. $y''-3y'+2y=e^{x}(3-4x)$
9. $y''+y'-12y=e^{3x}(-6+7x)$
10. $2y''-3y'-2y=e^{2x}(-6+10x)$
11. $y''+2y'+y=e^{-x}(2+3x)$
12. $y''-2y'+y=e^{x}(1-6x)$
13. $y''-4y'+4y=e^{2x}(1-3x+6x^2)$
14. $9y''+6y'+y=e^{-x/3}(2-4x+4x^2)$

Q5.4.2

In **Exercises 5.4.15-5.4.19** find the general solution.

15. $y''-3y'+2y=e^{3x}(1+x)$
16. $y''-6y'+8y=e^{x}(11-6x)$
17. $y''+6y'+9y=e^{2x}(3-5x)$
18. $y''+2y'-3y=-16xe^{x}$
19. $y''-2y'+y=e^{x}(2-12x)$

Q5.4.3

In **Exercises 5.4.20-5.4.23** solve the initial value problem and plot the solution.

20. $y''-4y'-5y=9e^{2x}(1+x),\quad y(0)=0,\quad y'(0)=-10$
21. $y''+3y'-4y=e^{2x}(7+6x),\quad y(0)=2,\quad y'(0)=8$
22. $y''+4y'+3y=-e^{-x}(2+8x),\quad y(0)=1,\quad y'(0)=2$
23. $y''-3y'-10y=7e^{-2x},\quad y(0)=1,\quad y'(0)=-17$

Q5.4.4

In **Exercises 5.4.24-5.4.29** use the principle of superposition to find a particular solution.

24. $y''+y'+y=xe^{x}+e^{-x}(1+2x)$
25. $y''-7y'+12y=-e^{x}(17-42x)-e^{3x}$
26. $y''-8y'+16y=6xe^{4x}+2+16x+16x^2$
27. $y''-3y'+2y=-e^{2x}(3+4x)-e^{x}$

28. $y''-2y'+2y=e^x(1+x)+e^{-x}(2-8x+5x^2)$
29. $y''+y=e^{-x}(2-4x+2x^2)+e^{3x}(8-12x-10x^2)$

Q5.4.5

30.

a. Prove that y is a solution of the constant coefficient equation

$$ay''+by'+cy=e^{\alpha x}G(x) \tag{A}$$

if and only if $y=ue^{\alpha x}$, where u satisfies

$$au''+p'(\alpha)u'+p(\alpha)u=G(x) \tag{B}$$

and $p(r)=ar^2+br+c$ is the characteristic polynomial of the complementary equation

$$ay''+by'+cy=0.$$

For the rest of this exercise, let G be a polynomial. Give the requested proofs for the case where

$$G(x)=g_0+g_1x+g_2x^2+g_3x^3.$$

b. Prove that if $e^{\alpha x}$ isn't a solution of the complementary equation then (B) has a particular solution of the form $u_p=A(x)$, where A is a polynomial of the same degree as G, as in Example 5.4.4. Conclude that (A) has a particular solution of the form $y_p=e^{\alpha x}A(x)$.

c. Show that if $e^{\alpha x}$ is a solution of the complementary equation and $xe^{\alpha x}$ isn't , then (B) has a particular solution of the form $u_p=xA(x)$, where A is a polynomial of the same degree as G, as in Example 5.4.5. Conclude that (A) has a particular solution of the form $y_p=xe^{\alpha x}A(x)$.

d. Show that if $e^{\alpha x}$ and $xe^{\alpha x}$ are both solutions of the complementary equation then (B) has a particular solution of the form $u_p=x^2A(x)$, where A is a polynomial of the same degree as G, and $x^2A(x)$ can be obtained by integrating G/a twice, taking the constants of integration to be zero, as in Example 5.4.6. Conclude that (A) has a particular solution of the form $y_p=x^2e^{\alpha x}A(x)$.

Q5.4.6

Exercises 5.4.31–5.4.36 treat the equations considered in Examples 5.4.1–5.4.6. Substitute the suggested form of y_p into the equation and equate the resulting coefficients of like functions on the two sides of the resulting equation to derive a set of simultaneous equations for the coefficients in y_p. Then solve for the coefficients to obtain y_p. Compare the work you've done with the work required to obtain the same results in Examples 5.4.1–5.4.6.

31. Compare with Example 5.4.1:

$$y''-7y'+12y=4e^{2x};\quad y_p=Ae^{2x}$$

32. Compare with Example 5.4.2:

$$y''-7y'+12y=5e^{4x};\quad y_p=Axe^{4x}$$

33. Compare with Example 5.4.3:

$$y''-8y'+16y=2e^{4x};\quad y_p=Ax^2e^{4x}$$

34. Compare with Example 5.4.4:

$$y''-3y'+2y=e^{3x}(-1+2x+x^2),\quad y_p=e^{3x}(A+Bx+Cx^2)$$

35. Compare with Example 5.4.5:

$$y''-4y'+3y=e^{3x}(6+8x+12x^2),\quad y_p=e^{3x}(Ax+Bx^2+Cx^3)$$

36. Compare with Example 5.4.6:

$$4y''+4y'+y=e^{-x/2}(-8+48x+144x^2),\quad y_p=e^{-x/2}(Ax^2+Bx^3+Cx^4)$$

Q5.4.7

37. Write $y = ue^{\alpha x}$ to find the general solution.

a. $y'' + 2y' + y = \frac{e^{-x}}{\sqrt{x}}$
b. $y'' + 6y' + 9y = e^{-3x} \ln x$
c. $y'' - 4y' + 4y = \frac{e^{2x}}{1+x}$
d. $4y'' + 4y' + y = 4e^{-x/2} \left(\frac{1}{x} + x\right)$

38. Suppose $\alpha \neq 0$ and k is a positive integer. In most calculus books integrals like $\int x^k e^{\alpha x}\, dx$ are evaluated by integrating by parts k times. This exercise presents another method. Let

$$y = \int e^{\alpha x} P(x)\, dx$$

with

$$P(x) = p_0 + p_1 x + \cdots + p_k x^k$$

(where $p_k \neq 0$).

a. Show that $y = e^{\alpha x} u$, where

$$u' + \alpha u = P(x). \tag{A}$$

b. Show that (A) has a particular solution of the form

$$u_p = A_0 + A_1 x + \cdots + A_k x^k,$$

where A_k, A_{k-1}, ..., A_0 can be computed successively by equating coefficients of $x^k, x^{k-1}, \ldots, 1$ on both sides of the equation

$$u_p' + \alpha u_p = P(x).$$

c. Conclude that

$$\int e^{\alpha x} P(x)\, dx = \left(A_0 + A_1 x + \cdots + A_k x^k\right) e^{\alpha x} + c,$$

where c is a constant of integration.

39. Use the method of **Exercise 5.4.38** to evaluate the integral.

a. $\int e^x (4 + x) dx$
b. $\int e^{-x} (-1 + x^2) dx$
c. $\int x^3 e^{-2x} dx$
d. $\int e^x (1 + x)^2 dx$
e. $\int e^{3x} (-14 + 30x + 27x^2) dx$
f. $\int e^{-x} (1 + 6x^2 - 14x^3 + 3x^4) dx$

40. Use the method suggested in **Exercise 5.4.38** to evaluate $\int x^k e^{\alpha x}\, dx$, where k is an arbitrary positive integer and $\alpha \neq 0$.

This page titled 14.2E: The Method of Undetermined Coefficients I (Exercises) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **5.4E: The Method of Undetermined Coefficients I (Exercises)** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

14.3: The Method of Undetermined Coefficients II

In this section we consider the constant coefficient equation

$$ay'' + by' + cy = e^{\lambda x}\left(P(x)\cos\omega x + Q(x)\sin\omega x\right) \tag{14.3.1}$$

where λ and ω are real numbers, $\omega \neq 0$, and P and Q are polynomials. We want to find a particular solution of Equation 14.3.1. As in Section 5.4, the procedure that we will use is called *the method of undetermined coefficients*.

Forcing Functions Without Exponential Factors

We begin with the case where $\lambda = 0$ in Equation 14.3.1; thus, we we want to find a particular solution of

$$ay'' + by' + cy = P(x)\cos\omega x + Q(x)\sin\omega x, \tag{14.3.2}$$

where P and Q are polynomials.

Differentiating $x^r\cos\omega x$ and $x^r\sin\omega x$ yields

$$\frac{d}{dx}x^r\cos\omega x = -\omega x^r\sin\omega x + rx^{r-1}\cos\omega x$$

and

$$\frac{d}{dx}x^r\sin\omega x = \quad \omega x^r\cos\omega x + rx^{r-1}\sin\omega x.$$

This implies that if

$$y_p = A(x)\cos\omega x + B(x)\sin\omega x$$

where A and B are polynomials, then

$$ay_p'' + by_p' + cy_p = F(x)\cos\omega x + G(x)\sin\omega x,$$

where F and G are polynomials with coefficients that can be expressed in terms of the coefficients of A and B. This suggests that we try to choose A and B so that $F = P$ and $G = Q$, respectively. Then y_p will be a particular solution of Equation 14.3.2. The next theorem tells us how to choose the proper form for y_p. For the proof see *Exercise 5.5.37*.

Below is a video writing the form of the particular solution to a differential equation.



Theorem 4.3.1

Suppose ω is a positive number and P and Q are polynomials. Let k be the larger of the degrees of P and Q. Then the equation

$$ay'' + by' + cy = P(x)\cos\omega x + Q(x)\sin\omega x$$

has a particular solution

$$y_p = A(x)\cos\omega x + B(x)\sin\omega x, \tag{14.3.3}$$

where

$$A(x) = A_0 + A_1x + \cdots + A_kx^k \quad \text{and} \quad B(x) = B_0 + B_1x + \cdots + B_kx^k,$$

provided that $\cos\omega x$ and $\sin\omega x$ are not solutions of the complementary equation. The solutions of

$$a(y''+\omega^2 y)=P(x)\cos\omega x+Q(x)\sin\omega x$$

for which $\cos\omega x$ and $\sin\omega x$ are solutions of the complementary equation are of the form of Equation 14.3.3, where

$$A(x)=A_0x+A_1x^2+\cdots+A_kx^{k+1}\quad\text{and}\quad B(x)=B_0x+B_1x^2+\cdots+B_kx^{k+1}.$$

For an analog of this theorem that's applicable to Equation 14.3.1, see *Exercise 5.5.38.*

Below is a video on writing the form of the particular solution to a differential equation.



Example 4.3.1

Find a particular solution of

$$y''-2y'+y=5\cos 2x+10\sin 2x. \tag{14.3.4}$$

Solution

In Equation 14.3.4 the coefficients of $\cos 2x$ and $\sin 2x$ are both zero degree polynomials (constants). Therefore Theorem 4.3.1 implies that Equation 14.3.4has a particular solution

$$y_p=A\cos 2x+B\sin 2x.$$

Since

$$y_p'=-2A\sin 2x+2B\cos 2x\quad\text{and}\quad y_p''=-4(A\cos 2x+B\sin 2x),$$

replacing y by y_p in Equation 14.3.4yields

$$\begin{aligned}y_p''-2y_p'+y_p &=-4(A\cos 2x+B\sin 2x)-4(-A\sin 2x+B\cos 2x)\\ &\qquad +(A\cos 2x+B\sin 2x)\\ &=(-3A-4B)\cos 2x+(4A-3B)\sin 2x.\end{aligned}$$

Equating the coefficients of $\cos 2x$ and $\sin 2x$ here with the corresponding coefficients on the right side of Equation 14.3.4 shows that y_p is a solution of Equation 14.3.4if

$$\begin{aligned}-3A-4B&=\ \ 5\\ 4A-3B&=10.\end{aligned}$$

Solving these equations yields $A=1$, $B=-2$. Therefore

$$y_p=\cos 2x-2\sin 2x$$

is a particular solution of Equation 14.3.4.

Below is a video on solving an initial value problem using the particular solution and the method of undetermined coefficients.



Below is a video on finding the finding the general solution to a nonhomogeneous differential equation using undetermined coefficients.



Example 4.3.2

Find a particular solution of

$$y'' + 4y = 8\cos 2x + 12\sin 2x. \tag{14.3.5}$$

Solution

The procedure used in Example 4.3.1 doesn't work here; substituting $y_p = A\cos 2x + B\sin 2x$ for y in Equation 14.3.5 yields

$$y_p'' + 4y_p = -4(A\cos 2x + B\sin 2x) + 4(A\cos 2x + B\sin 2x) = 0$$

for any choice of A and B, since $\cos 2x$ and $\sin 2x$ are both solutions of the complementary equation for Equation 14.3.5. We're dealing with the second case mentioned in Theorem 4.3.1 , and should therefore try a particular solution of the form

$$y_p = x(A\cos 2x + B\sin 2x). \tag{14.3.6}$$

Then

$$\begin{aligned} y_p' &= A\cos 2x + B\sin 2x + 2x(-A\sin 2x + B\cos 2x) \\ \text{and} y_p'' &= -4A\sin 2x + 4B\cos 2x - 4x(A\cos 2x + B\sin 2x) \\ &= -4A\sin 2x + 4B\cos 2x - 4y_p \text{ (see (14.3.6))}, \end{aligned}$$

so

$$y_p'' + 4y_p = -4A\sin 2x + 4B\cos 2x.$$

Therefore y_p is a solution of Equation 14.3.5 if

$$-4A\sin 2x + 4B\cos 2x = 8\cos 2x + 12\sin 2x,$$

which holds if $A = -3$ and $B = 2$. Therefore

$$y_p = -x(3\cos 2x - 2\sin 2x)$$

is a particular solution of Equation 14.3.5.

Below is a video on using the method of undetermined coefficients to solve a nonhomogeneous differential equation.



Example 4.3.3

Find a particular solution of

$$y''+3y'+2y=(16+20x)\cos x+10\sin x. \tag{14.3.7}$$

Solution

The coefficients of $\cos x$ and $\sin x$ in Equation 14.3.7 are polynomials of degree one and zero, respectively. Therefore Theorem 4.3.1 tells us to look for a particular solution of Equation 14.3.7 of the form

$$y_p=(A_0+A_1x)\cos x+(B_0+B_1x)\sin x. \tag{14.3.8}$$

Then

$$y_p'=(A_1+B_0+B_1x)\cos x+(B_1-A_0-A_1x)\sin x \tag{14.3.9}$$

and

$$y_p''=(2B_1-A_0-A_1x)\cos x-(2A_1+B_0+B_1x)\sin x, \tag{14.3.10}$$

so

$$y_p''+3y_p'+2y_p = [A_0+3A_1+3B_0+2B_1+(A_1+3B_1)x]\cos x+[B_0+3B_1-3A_0-2A_1+(B_1-3A_1)x]\sin x. \tag{14.3.11}$$

Comparing the coefficients of $x\cos x$, $x\sin x$, $\cos x$, and $\sin x$ here with the corresponding coefficients in Equation 14.3.7 shows that y_p is a solution of Equation 14.3.7 if

$$\begin{aligned} A_1+3B_1 &= 20\\ -3A_1+B_1 &= 0\\ A_0+3B_0+3A_1+2B_1 &= 16\\ -3A_0+B_0-2A_1+3B_1 &= 10. \end{aligned}$$

Solving the first two equations yields $A_1=2$, $B_1=6$. Substituting these into the last two equations yields

$$\begin{aligned} A_0+3B_0 &= 16-3A_1-2B_1=-2\\ -3A_0+B_0 &= 10+2A_1-3B_1=-4. \end{aligned}$$

Solving these equations yields $A_0=1$, $B_0=-1$. Substituting $A_0=1$, $A_1=2$, $B_0=-1$, $B_1=6$ into Equation 14.3.8 shows that

$$y_p=(1+2x)\cos x-(1-6x)\sin x$$

is a particular solution of Equation 14.3.7.

A Useful Observation

In Equations 14.3.9, 14.3.10, and 14.3.11 the polynomials multiplying $\sin x$ can be obtained by replacing A_0, A_1, B_0, and B_1 by B_0, B_1, $-A_0$, and $-A_1$, respectively, in the polynomials mutiplying $\cos x$. An analogous result applies in general, as follows (*Exercise 5.5.36*).

Theorem 4.3.2

If

$$y_p = A(x)\cos\omega x + B(x)\sin\omega x,$$

where $A(x)$ and $B(x)$ are polynomials with coefficients A_0 …, A_k and B_0, …, B_k, then the polynomials multiplying $\sin\omega x$ in

$$y_p', \quad y_p'', \quad ay_p'' + by_p' + cy_p \quad \text{and} \quad y_p'' + \omega^2 y_p$$

can be obtained by replacing A_0, …, A_k by B_0, …, B_k and B_0, …, B_k by $-A_0$, …, $-A_k$ in the corresponding polynomials multiplying $\cos\omega x$.

We will not use this theorem in our examples, but we recommend that you use it to check your manipulations when you work the exercises.

Example 4.3.4

Find a particular solution of

$$y'' + y = (8-4x)\cos x - (8+8x)\sin x. \tag{14.3.12}$$

Solution

According to Theorem 4.3.1 , we should look for a particular solution of the form

$$y_p = (A_0x + A_1x^2)\cos x + (B_0x + B_1x^2)\sin x, \tag{14.3.13}$$

since $\cos x$ and $\sin x$ are solutions of the complementary equation. However, let's try

$$y_p = (A_0 + A_1x)\cos x + (B_0 + B_1x)\sin x \tag{14.3.14}$$

first, so you can see why it doesn't work. From Equation 14.3.10,

$$y_p'' = (2B_1 - A_0 - A_1x)\cos x - (2A_1 + B_0 + B_1x)\sin x,$$

which together with Equation 14.3.14implies that

$$y_p'' + y_p = 2B_1\cos x - 2A_1\sin x.$$

Since the right side of this equation does not contain $x\cos x$ or $x\sin x$, Equation 14.3.14can't satisfy Equation 14.3.12no matter how we choose A_0, A_1, B_0, and B_1.

Now let y_p be as in Equation 14.3.13. Then

$$\begin{aligned} y_p' &= \left[A_0 + (2A_1 + B_0)x + B_1x^2\right]\cos x \\ &\quad + \left[B_0 + (2B_1 - A_0)x - A_1x^2\right]\sin x \end{aligned}$$

and

$$\begin{aligned} y_p'' &= \left[2A_1 + 2B_0 - (A_0 - 4B_1)x - A_1x^2\right]\cos x \\ &\quad + \left[2B_1 - 2A_0 - (B_0 + 4A_1)x - B_1x^2\right]\sin x, \end{aligned}$$

so

$$y_p'' + y_p = (2A_1 + 2B_0 + 4B_1x)\cos x + (2B_1 - 2A_0 - 4A_1x)\sin x.$$

Comparing the coefficients of $\cos x$ and $\sin x$ here with the corresponding coefficients in Equation 14.3.12shows that y_p is a solution of Equation 14.3.12if

$$\begin{aligned} 4B_1 &= -4 \\ -4A_1 &= -8 \\ 2B_0 + 2A_1 &= 8 \\ -2A_0 + 2B_1 &= -8. \end{aligned}$$

The solution of this system is $A_1 = 2$, $B_1 = -1$, $A_0 = 3$, $B_0 = 2$. Therefore

$$y_p = x\left[(3+2x)\cos x + (2-x)\sin x\right]$$

is a particular solution of Equation 14.3.12.

Forcing Functions with Exponential Factors

To find a particular solution of

$$ay'' + by' + cy = e^{\lambda x}\left(P(x)\cos\omega x + Q(x)\sin\omega x\right) \tag{14.3.15}$$

when $\lambda \neq 0$, we recall from Section 5.4 that substituting $y = ue^{\lambda x}$ into Equation 14.3.15will produce a constant coefficient equation for u with the forcing function $P(x)\cos\omega x + Q(x)\sin\omega x$. We can find a particular solution u_p of this equation by the procedure that we used in Examples 4.3.1

-4.3.4 . Then $y_p = u_p e^{\lambda x}$ is a particular solution of Equation 14.3.15

Example 4.3.5

Find a particular solution of

$$y'' - 3y' + 2y = e^{-2x}\left[2\cos 3x - (34 - 150x)\sin 3x\right]. \tag{14.3.16}$$

Let $y = ue^{-2x}$. Then

$$\begin{aligned} y'' - 3y' + 2y &= e^{-2x}\left[(u'' - 4u' + 4u) - 3(u' - 2u) + 2u\right] \\ &= e^{-2x}(u'' - 7u' + 12u) \\ &= e^{-2x}\left[2\cos 3x - (34 - 150x)\sin 3x\right] \end{aligned}$$

if

$$u'' - 7u' + 12u = 2\cos 3x - (34 - 150x)\sin 3x. \tag{14.3.17}$$

Since $\cos 3x$ and $\sin 3x$ aren't solutions of the complementary equation

$$u'' - 7u' + 12u = 0,$$

Theorem 4.3.1 tells us to look for a particular solution of Equation 14.3.17of the form

$$u_p = (A_0 + A_1 x)\cos 3x + (B_0 + B_1 x)\sin 3x. \tag{14.3.18}$$

Then

$$\begin{aligned} u_p' &= (A_1 + 3B_0 + 3B_1 x)\cos 3x + (B_1 - 3A_0 - 3A_1 x)\sin 3x \\ \text{and} \qquad u_p'' &= (-9A_0 + 6B_1 - 9A_1 x)\cos 3x - (9B_0 + 6A_1 + 9B_1 x)\sin 3x, \end{aligned}$$

so

$$\begin{aligned} u_p'' - 7u_p' + 12u_p &= [3A_0 - 21B_0 - 7A_1 + 6B_1 + (3A_1 - 21B_1)x]\cos 3x \\ &\quad + [21A_0 + 3B_0 - 6A_1 - 7B_1 + (21A_1 + 3B_1)x]\sin 3x. \end{aligned}$$

Comparing the coefficients of $x\cos 3x$, $x\sin 3x$, $\cos 3x$, and $\sin 3x$ here with the corresponding coefficients on the right side of Equation 14.3.17shows that u_p is a solution of Equation 14.3.17if

$$\begin{aligned} 3A_1 - 21B_1 &= 0 \\ 21A_1 + 3B_1 &= 150 \\ 3A_0 - 21B_0 - 7A_1 + 6B_1 &= 2 \\ 21A_0 + 3B_0 - 6A_1 - 7B_1 &= -34. \end{aligned} \tag{14.3.19}$$

Solving the first two equations yields $A_1 = 7$, $B_1 = 1$. Substituting these values into the last two equations of Equation 14.3.19yields

$$\begin{aligned} 3A_0 - 21B_0 &= 2 + 7A_1 - 6B_1 = 45 \\ 21A_0 + 3B_0 &= -34 + 6A_1 + 7B_1 = 15. \end{aligned}$$

Solving this system yields $A_0 = 1$, $B_0 = -2$. Substituting $A_0 = 1$, $A_1 = 7$, $B_0 = -2$, and $B_1 = 1$ into Equation 14.3.18shows that

$$u_p = (1 + 7x)\cos 3x - (2 - x)\sin 3x$$

is a particular solution of Equation 14.3.17 Therefore

$$y_p = e^{-2x}\left[(1 + 7x)\cos 3x - (2 - x)\sin 3x\right]$$

is a particular solution of Equation 14.3.16

Example 4.3.6

Find a particular solution of

$$y'' + 2y' + 5y = e^{-x}\left[(6 - 16x)\cos 2x - (8 + 8x)\sin 2x\right]. \tag{14.3.20}$$

Solution

Let $y = ue^{-x}$. Then

$$\begin{aligned} y'' + 2y' + 5y &= e^{-x}\left[(u'' - 2u' + u) + 2(u' - u) + 5u\right] \\ &= e^{-x}(u'' + 4u) \\ &= e^{-x}\left[(6 - 16x)\cos 2x - (8 + 8x)\sin 2x\right] \end{aligned}$$

if

$$u''+4u=(6-16x)\cos 2x-(8+8x)\sin 2x. \tag{14.3.21}$$

Since $\cos 2x$ and $\sin 2x$ are solutions of the complementary equation

$$u''+4u=0,$$

Theorem 4.3.1 tells us to look for a particular solution of Equation 14.3.21of the form

$$u_p=(A_0x+A_1x^2)\cos 2x+(B_0x+B_1x^2)\sin 2x.$$

Then

$$\begin{aligned} u_p' &= \left[A_0+(2A_1+2B_0)x+2B_1x^2\right]\cos 2x \\ &\quad +\left[B_0+(2B_1-2A_0)x-2A_1x^2\right]\sin 2x \end{aligned}$$

and

$$\begin{aligned} u_p'' &= \left[2A_1+4B_0-(4A_0-8B_1)x-4A_1x^2\right]\cos 2x \\ &\quad +\left[2B_1-4A_0-(4B_0+8A_1)x-4B_1x^2\right]\sin 2x, \end{aligned}$$

so

$$u_p''+4u_p=(2A_1+4B_0+8B_1x)\cos 2x+(2B_1-4A_0-8A_1x)\sin 2x.$$

Equating the coefficients of $x\cos 2x$, $x\sin 2x$, $\cos 2x$, and $\sin 2x$ here with the corresponding coefficients on the right side of Equation 14.3.21 shows that u_p is a solution of Equation 14.3.21if

$$\begin{aligned} 8B_1 &= -16 \\ -8A_1 &= -\ 8 \\ 4B_0+2A_1 &= 6 \\ -4A_0+2B_1 &= -8. \end{aligned} \tag{14.3.22}$$

The solution of this system is $A_1=1$, $B_1=-2$, $B_0=1$, $A_0=1$. Therefore

$$u_p=x[(1+x)\cos 2x+(1-2x)\sin 2x]$$

is a particular solution of Equation 14.3.21, and

$$y_p=xe^{-x}\left[(1+x)\cos 2x+(1-2x)\sin 2x\right]$$

is a particular solution of Equation 14.3.20

You can also find a particular solution of Equation 14.3.20by substituting

$$y_p=xe^{-x}\left[(A_0+A_1x)\cos 2x+(B_0+B_1x)\sin 2x\right]$$

for y in Equation 14.3.20and equating the coefficients of $xe^{-x}\cos 2x$, $xe^{-x}\sin 2x$, $e^{-x}\cos 2x$, and $e^{-x}\sin 2x$ in the resulting expression for

$$y_p''+2y_p'+5y_p$$

with the corresponding coefficients on the right side of Equation 14.3.20 (See *Exercise 5.5.38*). This leads to the same system Equation 14.3.22 of equations for A_0, A_1, B_0, and B_1 that we obtained in Example 4.3.6 . However, if you try this approach you'll see that deriving Equation 14.3.22this way is much more tedious than the way we did it in Example 4.3.6 .

Below is a video on using the method of undetermined coefficients to find the general solution to a differential equation.



This page titled 14.3: The Method of Undetermined Coefficients II is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **5.5: The Method of Undetermined Coefficients II** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

14.3E: The Method of Undetermined Coefficients II (Exercises)

Q5.5.1

In **Exercises 5.5.1-5.5.17** find a particular solution.

1. $y''+3y'+2y=7\cos x-\sin x$
2. $y''+3y'+y=(2-6x)\cos x-9\sin x$
3. $y''+2y'+y=e^{x}(6\cos x+17\sin x)$
4. $y''+3y'-2y=-e^{2x}(5\cos 2x+9\sin 2x)$
5. $y''-y'+y=e^{x}(2+x)\sin x$
6. $y''+3y'-2y=e^{-2x}\left[(4+20x)\cos 3x+(26-32x)\sin 3x\right]$
7. $y''+4y=-12\cos 2x-4\sin 2x$
8. $y''+y=(-4+8x)\cos x+(8-4x)\sin x$
9. $4y''+y=-4\cos x/2-8x\sin x/2$
10. $y''+2y'+2y=e^{-x}(8\cos x-6\sin x)$
11. $y''-2y'+5y=e^{x}\left[(6+8x)\cos 2x+(6-8x)\sin 2x\right]$
12. $y''+2y'+y=8x^{2}\cos x-4x\sin x$
13. $y''+3y'+2y=(12+20x+10x^{2})\cos x+8x\sin x$
14. $y''+3y'+2y=(1-x-4x^{2})\cos 2x-(1+7x+2x^{2})\sin 2x$
15. $y''-5y'+6y=-e^{x}\left[(4+6x-x^{2})\cos x-(2-4x+3x^{2})\sin x\right]$
16. $y''-2y'+y=-e^{x}\left[(3+4x-x^{2})\cos x+(3-4x-x^{2})\sin x\right]$
17. $y''-2y'+2y=e^{x}\left[(2-2x-6x^{2})\cos x+(2-10x+6x^{2})\sin x\right]$

Q5.5.2

In **Exercises 5.5.18-5.5.21** find a particular solution and graph it.

18. $y''+2y'+y=e^{-x}\left[(5-2x)\cos x-(3+3x)\sin x\right]$
19. $y''+9y=-6\cos 3x-12\sin 3x$
20. $y''+3y'+2y=(1-x-4x^{2})\cos 2x-(1+7x+2x^{2})\sin 2x$
21. $y''+4y'+3y=e^{-x}\left[(2+x+x^{2})\cos x+(5+4x+2x^{2})\sin x\right]$

Q5.5.3

In **Exercises 5.5.22-5.5.26** solve the initial value problem.

22. $y''-7y'+6y=-e^{x}(17\cos x-7\sin x),\quad y(0)=4,\ y'(0)=2$
23. $y''-2y'+2y=-e^{x}(6\cos x+4\sin x),\quad y(0)=1,\ y'(0)=4$
24. $y''+6y'+10y=-40e^{x}\sin x,\quad y(0)=2,\quad y'(0)=-3$
25. $y''-6y'+10y=-e^{3x}(6\cos x+4\sin x),\quad y(0)=2,\quad y'(0)=7$
26. $y''-3y'+2y=e^{3x}\left[21\cos x-(11+10x)\sin x\right],\ y(0)=0,\quad y'(0)=6$

Q5.5.4

In **Exercises 5.5.27-5.5.32** use the principle of superposition to find a particular solution. Where indicated, solve the initial value problem.

27. $y''-2y'-3y=4e^{3x}+e^{x}(\cos x-2\sin x)$

28. $y''+y=4\cos x-2\sin x+xe^x+e^{-x}$

29. $y''-3y'+2y=xe^x+2e^{2x}+\sin x$

30. $y''-2y'+2y=4xe^x\cos x+xe^{-x}+1+x^2$

31. $y''-4y'+4y=e^{2x}(1+x)+e^{2x}(\cos x-\sin x)+3e^{3x}+1+x$

32. $y''-4y'+4y=6e^{2x}+25\sin x,\quad y(0)=5,\ y'(0)=3$

Q5.5.5

In **Exercises 5.5.33-5.5.35** solve the initial value problem and graph the solution.

33. $y''+4y=-e^{-2x}\left[(4-7x)\cos x+(2-4x)\sin x\right],\ y(0)=3,\quad y'(0)=1$

34. $y''+4y'+4y=2\cos 2x+3\sin 2x+e^{-x},\quad y(0)=-1,\ y'(0)=2$

35. $y''+4y=e^x(11+15x)+8\cos 2x-12\sin 2x,\quad y(0)=3,\ y'(0)=5$

Q5.5.6

36.

a. Verify that if

$$y_p=A(x)\cos\omega x+B(x)\sin\omega x$$

where A and B are twice differentiable, then

$$\begin{aligned} y_p' &= (A'+\omega B)\cos\omega x+(B'-\omega A)\sin\omega x \quad \text{and}\\ y_p'' &= (A''+2\omega B'-\omega^2A)\cos\omega x+(B''-2\omega A'-\omega^2B)\sin\omega x. \end{aligned}$$

b. Use the results of (a) to verify that

$$\begin{aligned} ay_p''+by_p'+cy_p= &\left[(c-a\omega^2)A+b\omega B+2a\omega B'+bA'+aA''\right]\cos\omega x+\\ &\left[-b\omega A+(c-a\omega^2)B-2a\omega A'+bB'+aB''\right]\sin\omega x. \end{aligned}$$

c. Use the results of (a) to verify that

$$y_p''+\omega^2y_p=(A''+2\omega B')\cos\omega x+(B''-2\omega A')\sin\omega x.$$

d. Prove Theorem 5.5.2.

37. Let a, b, c, and ω be constants, with $a\neq 0$ and $\omega>0$, and let

$$P(x)=p_0+p_1x+\cdots+p_kx^k\quad\text{and}\quad Q(x)=q_0+q_1x+\cdots+q_kx^k,$$

where at least one of the coefficients p_k, q_k is nonzero, so k is the larger of the degrees of P and Q.

a. Show that if $\cos\omega x$ and $\sin\omega x$ are not solutions of the complementary equation

$$ay''+by'+cy=0,$$

then there are polynomials

$$A(x)=A_0+A_1x+\cdots+A_kx^k\quad\text{and}\quad B(x)=B_0+B_1x+\cdots+B_kx^k \tag{A}$$

such that

$$\begin{aligned} (c-a\omega^2)A+b\omega B+2a\omega B'+bA'+aA'' &= P\\ -b\omega A+(c-a\omega^2)B-2a\omega A'+bB'+aB'' &= Q, \end{aligned}$$

where (A_k,B_k), (A_{k-1},B_{k-1}), ...,(A_0,B_0) can be computed successively by solving the systems

$$(c-a\omega^2)A_k+b\omega B_k \quad = p_k$$
$$-b\omega A_k+(c-a\omega^2)B_k \quad = q_k,$$

and, if $1\le r\le k$,

$$(c-a\omega^2)A_{k-r}+b\omega B_{k-r} \quad = p_{k-r}+\cdots$$
$$-b\omega A_{k-r}+(c-a\omega^2)B_{k-r} \quad = q_{k-r}+\cdots,$$

where the terms indicated by "$\cdots$" depend upon the previously computed coefficients with subscripts greater than $k-r$. Conclude from this and **Exercise 5.5.36b** that

$$y_p=A(x)\cos\omega x+B(x)\sin\omega x \tag{B}$$

is a particular solution of

$$ay''+by'+cy=P(x)\cos\omega x+Q(x)\sin\omega x.$$

b. Conclude from **Exercise 5.5.36c** that the equation

$$a(y''+\omega^2 y)=P(x)\cos\omega x+Q(x)\sin\omega x \tag{C}$$

does not have a solution of the form (B) with A and B as in (A). Then show that there are polynomials

$$A(x)=A_0x+A_1x^2+\cdots+A_kx^{k+1} \quad \text{and} \quad B(x)=B_0x+B_1x^2+\cdots+B_kx^{k+1}$$

such that

$$a(A''+2\omega B') \quad = P$$
$$a(B''-2\omega A') \quad = Q,$$

where the pairs (A_k,B_k), (A_{k-1},B_{k-1}), ..., (A_0,B_0) can be computed successively as follows:

$$A_k=-\frac{q_k}{2a\omega(k+1)}$$
$$B_k=\frac{p_k}{2a\omega(k+1)},$$

and, if $k\ge 1$,

$$A_{k-j} \;=-\frac{1}{2\omega}\left[\frac{q_{k-j}}{a(k-j+1)}-(k-j+2)B_{k-j+1}\right]$$
$$B_{k-j} \;=\frac{1}{2\omega}\left[\frac{p_{k-j}}{a(k-j+1)}-(k-j+2)A_{k-j+1}\right]$$

for $1\le j\le k$. Conclude that (B) with this choice of the polynomials A and B is a particular solution of (C).

38. Show that Theorem 5.5.1 implies the next theorem:

Theorem 5.5E.1

Suppose ω is a positive number and P and Q are polynomials. Let k be the larger of the degrees of P and Q. Then the equation

$$ay''+by'+cy=e^{\lambda x}\left(P(x)\cos\omega x+Q(x)\sin\omega x\right)$$

has a particular solution

$$y_p=e^{\lambda x}\left(A(x)\cos\omega x+B(x)\sin\omega x\right), \tag{A}$$

where

$$A(x)=A_0+A_1x+\cdots+A_kx^k \quad \text{and} \quad B(x)=B_0+B_1x+\cdots+B_kx^k,$$

provided that $e^{\lambda x}\cos\omega x$ and $e^{\lambda x}\sin\omega x$ are not solutions of the complementary equation. The equation

$$a\left[y''-2\lambda y'+(\lambda^2+\omega^2)y\right]=e^{\lambda x}\left(P(x)\cos\omega x+Q(x)\sin\omega x\right)$$

(for which $e^{\lambda x}\cos\omega x$ and $e^{\lambda x}\sin\omega x$ are solutions of the complementary equation) has a particular solution of the form (A), where

$$A(x)=A_0x+A_1x^2+\cdots+A_kx^{k+1}\quad\text{and}\quad B(x)=B_0x+B_1x^2+\cdots+B_kx^{k+1}.$$

39. This exercise presents a method for evaluating the integral

$$y=\int e^{\lambda x}\left(P(x)\cos\omega x+Q(x)\sin\omega x\right)\,dx$$

where $\omega\neq0$ and

$$P(x)=p_0+p_1x+\cdots+p_kx^k,\quad Q(x)=q_0+q_1x+\cdots+q_kx^k.$$

a. Show that $y=e^{\lambda x}u$, where

$$u'+\lambda u=P(x)\cos\omega x+Q(x)\sin\omega x. \tag{A}$$

b. Show that (A) has a particular solution of the form

$$u_p=A(x)\cos\omega x+B(x)\sin\omega x,$$

where

$$A(x)=A_0+A_1x+\cdots+A_kx^k,\quad B(x)=B_0+B_1x+\cdots+B_kx^k,$$

and the pairs of coefficients (A_k,B_k), (A_{k-1},B_{k-1}), …,(A_0,B_0) can be computed successively as the solutions of pairs of equations obtained by equating the coefficients of $x^r\cos\omega x$ and $x^r\sin\omega x$ for $r=k$, $k-1$, …, 0.

c. Conclude that

$$\int e^{\lambda x}\left(P(x)\cos\omega x+Q(x)\sin\omega x\right)\,dx=e^{\lambda x}\left(A(x)\cos\omega x+B(x)\sin\omega x\right)+c,$$

where c is a constant of integration.

40. Use the method of **Exercise 5.5.39** to evaluate the integral.

a. $\int x^2\cos x\,dx$
b. $\int x^2e^x\cos x\,dx$
c. $\int xe^{-x}\sin2x\,dx$
d. $\int x^2e^{-x}\sin x\,dx$
e. $\int x^3e^x\sin x\,dx$
f. $\int e^x[x\cos x-(1+3x)\sin x]dx$
g. $\int e^{-x}[(1+x^2)\cos x+(1^2_x)\sin x]dx$

This page titled 14.3E: The Method of Undetermined Coefficients II (Exercises) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **5.5E: The Method of Undetermined Coefficients II (Exercises)** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

14.4: Constant coefficient second order linear ODEs

14.4.1: Solving Constant Coefficient Equations

Suppose we have the problem

$$y'' - 6y' + 8y = 0, y(0) = -2, y'(0) = 6$$

This is a second order linear homogeneous equation with constant coefficients. Constant coefficients means that the functions in front of y'', y', and y are constants and do not depend on x.

To guess a solution, think of a function that you know stays essentially the same when we differentiate it, so that we can take the function and its derivatives, add some multiples of these together, and end up with zero.

Let us try[1] a solution of the form $y = e^{rx}$. Then $y' = re^{rx}$ and $y'' = r^2 e^{rx}$. Plug in to get

$$\begin{aligned} y'' - 6y' + 8y &= 0, \\ \underbrace{r^2 e^{rx}}_{y''} - 6\underbrace{re^{rx}}_{y'} + 8\underbrace{e^{rx}}_{y} &= 0, \\ r^2 - 6r + 8 &= 0 \qquad (\text{divide through by } e^{rx}), \\ (r-2)(r-4) &= 0. \end{aligned} \tag{14.4.1}$$

Hence, if $r = 2$ or $r = 4$, then e^{rx} is a solution. So let $y_1 = e^{2x}$ and $y_2 = e^{4x}$.

? Exercise 14.4.1

Check that y_1 and y_2 are solutions.

Solution

The functions e^{2x} and e^{4x} are linearly independent. If they were not linearly independent we could write $e^{4x} = Ce^{2x}$ for some constant C, implying that $e^{2x} = C$ for all x, which is clearly not possible. Hence, we can write the general solution as

$$y = C_1 e^{2x} + C_2 e^{4x}$$

We need to solve for C_1 and C_2. To apply the initial conditions we first find $y' = 2C_1 e^{2x} + 4C_2 e^{4x}$. We plug in $x = 0$ and solve.

$$\begin{aligned} -2 &= y(0) = C_1 + C_2 \\ 6 &= y'(0) = 2C_1 + 4C_2 \end{aligned} \tag{14.4.2}$$

Either apply some matrix algebra, or just solve these by high school math. For example, divide the second equation by 2 to obtain $3 = C_1 + 2C_2$, and subtract the two equations to get $5 = C_2$. Then $C_1 = -7$ as $-2 = C_1 + 5$. Hence, the solution we are looking for is

$$y = -7e^{2x} + 5e^{4x}$$

Let us generalize this example into a method. Suppose that we have an equation

$$ay'' + by' + cy = 0, \tag{14.4.3}$$

where a, b, c are constants. Try the solution $y = e^{rx}$ to obtain

$$ar^2 e^{rx} + bre^{rx} + ce^{rx} = 0$$

Divide by e^{rx} to obtain the so-called *characteristic equation* of the ODE:

$$ar^2 + br + c = 0$$

Solve for the r by using the quadratic formula.

$$r_1, r_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Therefore, we have $e^{r_1 x}$ and $e^{r_2 x}$ as solutions. There is still a difficulty if $r_1 = r_2$, but it is not hard to overcome.

Theorem 14.4.1

Suppose that r_1 and r_2 are the roots of the characteristic equation.

If r_1 and r_2 are distinct and real (when $b^2 - 4ac > 0$), then (14.4.3) has the general solution

$$y = C_1 e^{r_1 x} + C_2 e^{r_2 x}$$

If $r_1 = r_2$ (happens when $b^2 - 4ac = 0$), then (14.4.3) has the general solution

$$y = (C_1 + C_2 x) e^{r_1 x}$$

For another example of the first case, take the equation $y'' - k^2 y = 0$. Here the characteristic equation is $r^2 - k^2 = 0$ or $(r-k)(r+k) = 0$. Consequently, e^{-kx} and e^{kx} are the two linearly independent solutions.

Below is a video on the characteristic equation of a differential equation.



Example 14.4.1

Solve

$$y'' - k^2 y = 0.$$

Solution

The characteristic equation is $r^2 - k^2 = 0$ or $(r-k)(r+k) = 0$. Consequently, e^{-kx} and e^{kx} are the two linearly independent solutions, and the general solution is

$$y = C_1 e^{kx} + C_2 e_{-kx}.$$

Since $\cosh s = \frac{e^s + e^{-s}}{2}$ and $\sinh s = \frac{e^s - e^{-s}}{2}$, we can also write the general solution as

$$y = D_1 \cosh(kx) + D_2 \sinh(kx).$$

Below is a video on finding the finding the general solution to a differential equation.



Example 14.4.2:

Find the general solution of

$$y'' - 8y' + 16y = 0$$

Solution

The characteristic equation is $r^2 - 8r + 16 = (r-4)^2 = 0$. The equation has a double root $r_1 = r_2 = 4$. The general solution is, therefore,

$$y = (C_1 + C_2 x)e^{4x} = C_1 e^{4x} + C_2 x e^{4x}$$

Below is a video on finding the general solution to a differential equation involving two real irrational roots.



Exercise 14.4.2: Linear Independence

Check that e^{4x} and xe^{4x} *are linearly independent.*

Answer

That e^{4x} solves the equation is clear. If xe^{4x} solves the equation, then we know we are done. Let us compute $y' = e^{4x} + 4xe^{4x}$ and $y'' = 8e^{4x} + 16xe^{4x}$. Plug in

$$y'' - 8y' + 16y = 8e^{4x} + 16xe^{4x} - 8(e^{4x} + 4xe^{4x}) + 16xe^{4x} = 0$$

We should note that in practice, doubled root rarely happens. If coefficients are picked truly randomly we are very unlikely to get a doubled root.

Below is a video on finding the the general solution to a differential equation.



Let us give a short proof for why the solution xe^{rx} works when the root is doubled. This case is really a limiting case of when the two roots are distinct and very close. Note that $\frac{e^{r}2^{x}-e^{x}1^{x}}{r_2-r_1}$ is a solution when the roots are distinct. When we take the limit as r_1 goes to r_2, we are really taking the derivative of e^{rx} using r as the variable. Therefore, the limit is xe^{rx}, and hence this is a solution in the doubled root case.

Below is a video on finding the solution to a differential equation involving repeated roots.



14.4.2: Complex numbers and Euler's formula

It may happen that a polynomial has some complex roots. For example, the equation $r^2+1=0$ has no real roots, but it does have two complex roots. Here we review some properties of complex numbers.

Complex numbers may seem a strange concept, especially because of the terminology. There is nothing imaginary or really complicated about complex numbers. A complex number is simply a pair of real numbers, (a,b). We can think of a complex number as a point in the plane. We add complex numbers in the straightforward way, $(a,b)+(c,d)=(a+c,b+d)$. We define multiplication by

$$(a,b)\times(c,d)\stackrel{\text{def}}{=}(ac-bd,ad+bc).$$

It turns out that with this multiplication rule, all the standard properties of arithmetic hold. Further, and most importantly $(0,1)\times(0,1)=(-1,0)$.

Generally we just write (a,b) as $(a+ib)$, and we treat i as if it were an unknown. We do arithmetic with complex numbers just as we would with polynomials. The property we just mentioned becomes $i^2=-1$. So whenever we see i^2, we replace it by -1. The numbers i and $-i$ are the two roots of $r^2+1=0$.

Note that engineers often use the letter j instead of i for the square root of -1. We will use the mathematicians' convention and use i.

Exercise 14.4.3

Make sure you understand (that you can justify) the following identities:

a. $i^2 = -1, i^3 = -1, i^4 = 1$,
b. $\frac{1}{i} = -i$,
c. $(3-7i)(-2-9i) = \cdots = -69-13i$,
d. $(3-2i)(3+2i) = 3^2-(2i)^2 = 3^2+2^2 = 13$,
e. $\frac{1}{3-2i} = \frac{1}{3-2i}\frac{3+2i}{3+2i} = \frac{3+2i}{13} = \frac{3}{13}+\frac{2}{13}i$.

We can also define the exponential e^{a+ib} of a complex number. We do this by writing down the Taylor series and plugging in the complex number. Because most properties of the exponential can be proved by looking at the Taylor series, these properties still hold for the complex exponential. For example the very important property: $e^{x+y} = e^x e^y$. This means that $e^{a+ib} = e^a e^{ib}$. Hence if we can compute e^{ib}, we can compute e^{a+ib}. For e^{ib} we use the so-called Euler's formula.

Theorem 14.4.2

Euler's Formula

$$e^{i\theta} = \cos\theta + i\sin\theta \quad \textit{and} \quad e^{-i\theta} = \cos\theta - i\sin\theta$$

In other words, $e^{a+ib} = e^a(\cos(b)+i\sin(b)) = e^a\cos(b)+ie^a\sin(b)$.

Exercise 14.4.4:

Using Euler's formula, check the identities:

$$\cos\theta = \frac{e^{i\theta}+e^{-i\theta}}{2} \quad \text{and} \quad \sin\theta = \frac{e^{i\theta}-e^{-i\theta}}{2i}$$

Exercise 14.4.5

Double angle identities: Start with $e^{i(2\theta)} = \left(e^{i\theta}\right)^2$. Use Euler on each side and deduce:

Answer

$$\cos(2\theta) = \cos^2\theta - \sin^2\theta \quad \text{and} \quad \sin(2\theta) = 2\sin\theta\cos\theta$$

For a complex number $a+ib$ we call a the real part and b the imaginary part of the number. Often the following notation is used,

$$\mathrm{Re}(a+ib) = a \quad \text{and} \quad \mathrm{Im}(a+ib) = b$$

14.4.3: Complex roots

Suppose that the equation $ay''+by'+cy=0$ has the characteristic equation $ar^2+br+c=0$ that has complex roots. By the quadratic formula, the roots are $\frac{-b\pm\sqrt{b^2-4ac}}{2a}$. These roots are complex if $b^2-4ac<0$. In this case the roots are

$$r_1, r_2 = \frac{-b}{2a} \pm i\frac{\sqrt{4ac-b^2}}{2a}$$

As you can see, we always get a pair of roots of the form $\alpha \pm i\beta$. In this case we can still write the solution as

$$y = C_1 e^{(\alpha+i\beta)x} + C_2 e^{(\alpha-i\beta)x}$$

However, the exponential is now complex valued. We would need to allow C_1 and C_2 to be complex numbers to obtain a real-valued solution (which is what we are after). While there is nothing particularly wrong with this approach, it can make calculations harder and it is generally preferred to find two real-valued solutions.

Here we can use Euler's formula. Let

$$y_1 = e^{(\alpha+i\beta)x} \quad \text{and} \quad y_2 = e^{(\alpha-i\beta)x}$$

Then note that

$$\begin{aligned} y_1 &= e^{ax}\cos(\beta x) + ie^{ax}\sin(\beta x) \\ y_2 &= e^{ax}\cos(\beta x) - ie^{ax}\sin(\beta x) \end{aligned} \tag{14.4.4}$$

Linear combinations of solutions are also solutions. Hence,

$$\begin{aligned} y_3 &= \frac{y_1+y_2}{2} = e^{ax}\cos(\beta x) \\ y_4 &= \frac{y_1-y_2}{2i} = e^{ax}\sin(\beta x) \end{aligned} \tag{14.4.5}$$

are also solutions. Furthermore, they are real-valued. It is not hard to see that they are linearly independent (not multiples of each other). Therefore, we have the following theorem.

Theorem 14.4.3

For the homegneous second order ODE

$$ay'' + by' + cy = 0$$

If the characteristic equation has the roots $\alpha \pm i\beta$ (when $b^2 - 4ac < 0$), then the general solution is

$$y = C_1 e^{ax}\cos(\beta x) + C_2 e^{ax}\sin(\beta x)$$

Below is a video on finding the solution to a differential equation using the principal of superposition.



Example 14.4.3

Find the general solution of $y'' + k^2 y = 0$, for a constant $k > 0$.

Solution

The characteristic equation is $r^2 + k^2 = 0$. Therefore, the roots are $r = \pm ik$ and by the theorem we have the general solution

$$y = C_1\cos(kx) + C_2\sin(kx)$$

Below is a video on finding the solution to a differential equation involving complex roots.



Example 14.4.4

Find the solution of $y'' - 6y' + 13y = 0, y(0) = 0, y'(0) = 10$.

Solution

The characteristic equation is $r^2 - 6r + 13 = 0$. By completing the square we get $(r-3)^2 + 2^2 = 0$ and hence the roots are $r = 3 \pm 2i$. By the theorem we have the general solution

$$y = C_1 e^{3x} \cos(2x) + C_2 e^{3x} \sin(2x)$$

To find the solution satisfying the initial conditions, we first plug in zero to get

$$0 = y(0) = C_1 e^0 \cos 0 + C_2 e^0 \sin 0 = C_1$$

Hence $C_1 = 0$ and $y = C_2 e^{3x} \sin(2x)$. We differentiate

$$y' = 3C_2 e^{3x} \sin(2x) + 2C_2 e^{3x} \cos(2x)$$

We again plug in the initial condition and obtain $10 = y'(0) = 2C_2$, or $C_2 = 5$. Hence the solution we are seeking is

$$y = 5e^{3x} \sin(2x)$$

Below is a video on finding the solution to an initial value problem.



Below is a video on finding the solution to a differential equation given initial values.



Below is another video on finding the solution to a differential equation given initial values.



14.4.4: Footnotes

[1] Making an educated guess with some parameters to solve for is such a central technique in differential equations, that people sometimes use a fancy name for such a guess: *ansatz*, German for “initial placement of a tool at a work piece.” Yes, the Germans have a word for that.

This page titled 14.4: Constant coefficient second order linear ODEs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **2.2: Constant coefficient second order linear ODEs** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

14.5: Higher order linear ODEs

Equations that appear in applications tend to be second order, although higher order equations do appear from time to time. Hence, it is a generally assumed that the world is "second order" from a modern physics perspective. The basic results about linear ODEs of higher order are essentially the same as for second order equations, with 2 replaced by n. The important concept of linear independence is somewhat more complicated when more than two functions are involved.

For higher order constant coefficient ODEs, the methods are also somewhat harder to apply, but we will not dwell on these complications. We can always use the methods for systems of linear equations to solve higher order constant coefficient equations. So let us start with a general homogeneous linear equation:

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \ldots + p_1(x)y' + p_o(x)y = f(x) \tag{14.5.1}$$

Theorem 14.5.1

Superposition

Suppose $y_1, y_2, \ldots, y_n$ are solutions of the homogeneous equation (Equation 14.5.1). Then

$$y(x) = C_1y_1(x) + C_2y_2(x) + \ldots + C_ny_n(x)$$

also solves Equation 14.5.1 for arbitrary constants $C_1, \ldots . C_n$.

In other words, a linear combination of solutions to Equation 14.5.1 is also a solution to Equation 14.5.1. We also have the existence and uniqueness theorem for nonhomogeneous linear equations.

Theorem 14.5.2

Existence and Uniqueness

Suppose p_o through p_{n-1}, and f are continuous functions on some interval I, a is a number in I, and $b_0, b_1, \ldots, b_{n-1}$ are constants. The equation

$$y^{(n)} + p_{n-1}(x)y^{(n-1)} + \ldots + p_1(x)y' + p_o(x)y = f(x)$$

has exactly one solution $y(x)$ defined on the same interval I satisfying the initial conditions

$$y(a) = b_0, \quad y'(a) = b_1, \quad \ldots, \quad y^{(n-1)}(a) = b_{n-1}$$

14.5.1: Linear Independence

When we had two functions y_1 and y_2 we said they were linearly independent if one was not the multiple of the other. Same idea holds for n functions. In this case it is easier to state as follows. The functions $y_1, y_2, \ldots, y_n$ are linearly independent if

$$c_1y_1 + c_2y_2 + \cdots + c_ny_n = 0$$

has only the trivial solution $c_1 = c_2 = \cdots = c_n = 0$, where the equation must hold for all x. If we can solve equation with some constants where for example $c_1 \neq 0$, then we can solve for y_1 as a linear combination of the others. If the functions are not linearly independent, they are linearly dependent.

Example 14.5.1

Show that e^x, e^{2x}, and e^{3x} are linearly independent functions.

Solution

Let us give several ways to show this fact. Many textbooks introduce Wronskians, but that is really not necessary to solve this example. Let us write down

$$c_1e^x + c_2e^{2x} + c_3e^{3x} = 0$$

We use rules of exponentials and write $z = e^x$. Then we have

$$c_1 z + c_2 z^2 + c_3 z^3 = 0$$

The left hand side is a third degree polynomial in z. It can either be identically zero, or it can have at most 3 zeros. Therefore, it is identically zero, $c_1 = c_2 = c_3 = 0$, and the functions are linearly independent.

Let us try another way. As before we write

$$c_1 e^x + c_2 e^{2x} + c_3 e^{3x} = 0$$

This equation has to hold for all x. What we could do is divide through by e^{3x} to get

$$c_1 e^{-2x} + c_2 e^{-x} + c_3 = 0$$

As the equation is true for all x, let $x \to \infty$. After taking the limit we see that $c_3 = 0$. Hence our equation becomes

$$c_1 e^x + c_2 e^{2x} = 0$$

Rinse, repeat!

How about yet another way. We again write

$$c_1 e^x + c_2 e^{2x} + c_3 e^{3x} = 0$$

We can evaluate the equation and its derivatives at different values of x to obtain equations for c_1, c_2, and c_3. Let us first divide by e^x for simplicity.

$$c_1 + c_2 e^x + c_3 e^{2x} = 0$$

We set $x = 0$ to get the equation $c_1 + c_2 + c_3 = 0$. Now differentiate both sides

$$c_2 e^x + 2c_3 e^{2x} = 0$$

We set $x = 0$ to get $c_2 + 2c_3 = 0$. We divide by e^x again and differentiate to get $2c_3 e^x = 0$. It is clear that c_3 is zero. Then c_2 must be zero as $c_2 = -2c_3$, and c_1 must be zero because $c_1 + c_2 + c_3 = 0$.

There is no one best way to do it. All of these methods are perfectly valid. The important thing is to understand why the functions are linearly independent.

Example 14.5.2

On the other hand, the functions e^x, e^{-x} and $\cosh x$ are linearly dependent. Simply apply definition of the hyperbolic cosine:

$$\cosh x = \frac{e^x + e^{-x}}{2} \quad \text{or} \quad 2\cosh x - e^x - e^{-x} = 0$$

14.5.2: Constant Coefficient Higher Order ODEs

When we have a higher order constant coefficient homogeneous linear equation, the song and dance is exactly the same as it was for second order. We just need to find more solutions. If the equation is n^{th} order we need to find n linearly independent solutions. It is best seen by example.

Example 14.5.3: Third order ODE with Constant Coefficients

Find the general solution to

$$y''' - 3'' - y' + 3y = 0 \tag{14.5.2}$$

Solution

Try: $y = e^{rx}$. We plug in and get

$$\underbrace{r^3 e^{rx}}_{y'''} - 3\underbrace{r^2 e^{rx}}_{y''} - \underbrace{r e^{rx}}_{y'} + 3\underbrace{e^{rx}}_{y} = 0.$$

We divide through by e^{rx}. Then

$$r^3 - 3r^2 - r + 3 = 0$$

The trick now is to find the roots. There is a formula for the roots of degree 3 and 4 polynomials, but it is very complicated. There is no formula for higher degree polynomials. That does not mean that the roots do not exist. There are always n roots for an nth degree polynomial. They may be repeated and they may be complex. Computers are pretty good at finding roots approximately for reasonable size polynomials.

A good place to start is to plot the polynomial and check where it is zero. We can also simply try plugging in. We just start plugging in numbers $r = -2, -1, 0, 1, 2, \ldots$ and see if we get a hit (we can also try complex numbers). Even if we do not get a hit, we may get an indication of where the root is. For example, we plug $r = -2$ into our polynomial and get -15; we plug in $r = 0$ and get 3. That means there is a root between $r = -2$ and $r = 0$, because the sign changed. If we find one root, say r_1, then we know $(r - r_1)$ is a factor of our polynomial. Polynomial long division can then be used.

A good strategy is to begin with $r = -1$, 1, or 0. These are easy to compute. Our polynomial happens to have two such roots, $r_1 = -1$ and $r_2 = 1$ and. There should be three roots and the last root is reasonably easy to find. The constant term in a monic [1] polynomial such as this is the multiple of the negations of all the roots because $r^3 - 3r^2 - r + 3 = (r - r_1)(r - r_2)(r - r_3)$. So

$$3 = (-r_1)(-r_2)(-r_3) = (1)(-1)(-r_3) = r_3$$

You should check that $r_3 = 3$ really is a root. Hence we know that e^{-x}, e^x, and e^{3x} are solutions to (14.5.2). They are linearly independent as can easily be checked, and there are three of them, which happens to be exactly the number we need. Hence the general solution is

$$y = C_1 e^{-x} + C_2 e^x + C_3 e^{3x}$$

Suppose we were given some initial conditions $y(0) = 1, y'(0) = 2$, and $y''(0) = 3$. Then

$$\begin{aligned} 1 &= y(0) = C_1 + C_2 + C_3 \\ 2 &= y'(0) = -C_1 + C_2 + 3C_3 \\ 3 &= y''(0) = C_1 + C_2 + 9C_3 \end{aligned} \tag{14.5.3}$$

It is possible to find the solution by high school algebra, but it would be a pain. The sensible way to solve a system of equations such as this is to use matrix algebra, see Section 3.2 or Appendix A. For now we note that the solution is $C_1 = -\frac{1}{4}$, $C_2 = 1$, and $C_3 = \frac{1}{4}$. The specific solution to the ODE is

$$y = -\frac{1}{4}e^{-x} + e^x + \frac{1}{4}e^{3x}$$

Next, suppose that we have real roots, but they are repeated. Let us say we have a root r repeated k times. In the spirit of the second order solution, and for the same reasons, we have the solutions

$$e^{rx}, xe^{rx}, x^2e^{rx}, \ldots, x^{k-1}e^{rx}$$

We take a linear combination of these solutions to find the general solution.

Example 14.5.4

Solve

$$y^{(4)} - 3y''' + 3y'' - y' = 0$$

Solution

We note that the characteristic equation is

$$r^4 - 3r^3 + 3r^2 - r = 0$$

By inspection we note that $r^4 - 3r^3 + 3r^2 - r = r(r-1)^3$. Hence the roots given with multiplicity are $r = 0, 1, 1, 1$. Thus the general solution is

$$y = \underbrace{(C_1 + C_2 + C_3 x^2)e^x}_{\text{terms coming from r = 1}} + \underbrace{C_4}_{\text{from r = 0}}$$

The case of complex roots is similar to second order equations. Complex roots always come in pairs $r = \alpha \pm i\beta$. Suppose we have two such complex roots, each repeated k times. The corresponding solution is

$$(C_0 + C_1 x + \cdots + C_{k-1}x^{k-1})e^{ax}\cos(\beta x) + (D_0 + D_1 x + \cdots + D_{k-1}x^{k-1})e^{ax}\sin(\beta x)$$

where $C_0, \ldots, C_{k-1}, D_0, \ldots, D_{k-1}$ are arbitrary constants.

Below is a video on finding the solution to a differential equation given initial values.



Example 14.5.5

Solve

$$y^{(4)} - 4y''' + 8y'' - 8y' + 4y = 0$$

Solution

The characteristic equation is

$$\begin{aligned} r^4 - 4r^3 + 8r^2 - 8r + 4 &= 0 \\ (r^2 - 2r + 2)^2 &= 0 \\ ((r-1)^2 + 1)^2 &= 0 \end{aligned} \tag{14.5.4}$$

Hence the roots are $1 \pm i$, both with multiplicity 2. Hence the general solution to the ODE is

$$y = (C_1 + C_2 x)e^x \cos x + (C_3 + C_4 x)e^x \sin x$$

The way we solved the characteristic equation above is really by guessing or by inspection. It is not so easy in general. We could also have asked a computer or an advanced calculator for the roots.

14.5.3: Footnotes

[1] The word monic means that the coefficient of the top degree r^d, in our case r^3, is 1.

14.5.4: Outside Links

- After reading this lecture, it may be good to try Project III from the IODE website: www.math.uiuc.edu/iode/.

This page titled 14.5: Higher order linear ODEs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **2.3: Higher order linear ODEs** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

14.6: Reduction of Order

In this section we give a method for finding the general solution of

$$P_0(x)y'' + P_1(x)y' + P_2(x)y = F(x) \tag{14.6.1}$$

if we know a nontrivial solution y_1 of the complementary equation

$$P_0(x)y'' + P_1(x)y' + P_2(x)y = 0. \tag{14.6.2}$$

The method is called *reduction of order* because it reduces the task of solving Equation 14.6.1 to solving a first order equation. Unlike the method of undetermined coefficients, it does not require P_0, P_1, and P_2 to be constants, or F to be of any special form.

By now you shoudn't be surprised that we look for solutions of Equation 14.6.1 in the form

$$y = uy_1 \tag{14.6.3}$$

where u is to be determined so that y satisfies Equation 14.6.1. Substituting Equation 14.6.3 and

$$\begin{aligned} y' &= u'y_1 + uy_1' \\ y'' &= u''y_1 + 2u'y_1' + uy_1'' \end{aligned}$$

into Equation 14.6.1 yields

$$P_0(x)(u''y_1 + 2u'y_1' + uy_1'') + P_1(x)(u'y_1 + uy_1') + P_2(x)uy_1 = F(x).$$

Collecting the coefficients of u, u', and u'' yields

$$(P_0y_1)u'' + (2P_0y_1' + P_1y_1)u' + (P_0y_1'' + P_1y_1' + P_2y_1)u = F. \tag{14.6.4}$$

However, the coefficient of u is zero, since y_1 satisfies Equation 14.6.2. Therefore Equation 14.6.4 reduces to

$$Q_0(x)u'' + Q_1(x)u' = F, \tag{14.6.5}$$

with

$$Q_0 = P_0y_1 \quad \text{and} \quad Q_1 = 2P_0y_1' + P_1y_1.$$

(It isn't worthwhile to memorize the formulas for Q_0 and Q_1!) Since Equation 14.6.5 is a linear first order equation in u', we can solve it for u' by variation of parameters as in Section 1.2, integrate the solution to obtain u, and then obtain y from Equation 14.6.3.

Here is a video on using reduction of order to solve a differential equation.



Example 4.6.1

a. Find the general solution of

$$xy'' - (2x+1)y' + (x+1)y = x^2, \tag{14.6.6}$$

given that $y_1 = e^x$ is a solution of the complementary equation

$$xy'' - (2x+1)y' + (x+1)y = 0. \tag{14.6.7}$$

b. As a byproduct of (a), find a fundamental set of solutions of Equation 14.6.7.

Solution

a. If $y = ue^x$, then $y' = u'e^x + ue^x$ and $y'' = u''e^x + 2u'e^x + ue^x$, so

$$\begin{aligned} xy'' - (2x+1)y' + (x+1)y &= x(u''e^x + 2u'e^x + ue^x) - (2x+1)(u'e^x + ue^x) + (x+1)ue^x \\ &= (xu'' - u')e^x. \end{aligned}$$

Therefore $y = ue^x$ is a solution of Equation 14.6.6 if and only if

$$(xu'' - u')e^x = x^2,$$

which is a first order equation in u'. We rewrite it as

$$u'' - \frac{u'}{x} = xe^{-x}. \tag{14.6.8}$$

To focus on how we apply variation of parameters to this equation, we temporarily write $z = u'$, so that Equation 14.6.8 becomes

$$z' - \frac{z}{x} = xe^{-x}. \tag{14.6.9}$$

We leave it to you to show (by separation of variables) that $z_1 = x$ is a solution of the complementary equation

$$z' - \frac{z}{x} = 0$$

for Equation 14.6.9. By applying variation of parameters as in Section 1.2, we can now see that every solution of Equation 14.6.9 is of the form

$$z = vx \quad \text{where} \quad v'x = xe^{-x}, \quad \text{so} \quad v' = e^{-x} \quad \text{and} \quad v = -e^{-x} + C_1.$$

Since $u' = z = vx$, u is a solution of Equation 14.6.8 if and only if

$$u' = vx = -xe^{-x} + C_1 x.$$

Integrating this yields

$$u = (x+1)e^{-x} + \frac{C_1}{2}x^2 + C_2.$$

Therefore the general solution of Equation 14.6.6 is

$$y = ue^x = x + 1 + \frac{C_1}{2}x^2e^x + C_2e^x. \tag{14.6.10}$$

b. By letting $C_1 = C_2 = 0$ in Equation 14.6.10, we see that $y_{p_1} = x+1$ is a solution of Equation 14.6.6. By letting $C_1 = 2$ and $C_2 = 0$, we see that $y_{p_2} = x + 1 + x^2e^x$ is also a solution of Equation 14.6.6. Since the difference of two solutions of Equation 14.6.6 is a solution of Equation 14.6.7, $y_2 = y_{p_1} - y_{p_2} = x^2e^x$ is a solution of Equation 14.6.7. Since y_2/y_1 is nonconstant and we already know that $y_1 = e^x$ is a solution of Equation 14.6.6, Theorem 5.1.6 implies that $\{e^x, x^2e^x\}$ is a fundamental set of solutions of Equation 14.6.7.

Although Equation 14.6.10 is a correct form for the general solution of Equation 14.6.6, it is silly to leave the arbitrary coefficient of x^2e^x as $C_1/2$ where C_1 is an arbitrary constant. Moreover, it is sensible to make the subscripts of the coefficients of $y_1 = e^x$ and $y_2 = x^2e^x$ consistent with the subscripts of the functions themselves. Therefore we rewrite Equation 14.6.10 as

$$y = x + 1 + c_1e^x + c_2x^2e^x$$

by simply renaming the arbitrary constants. We'll also do this in the next two examples, and in the answers to the exercises.

Here is a video on using reduction of order to solve a differential equation given one solution.



Example 4.6.2

a. Find the general solution of

$$x^2y'' + xy' - y = x^2 + 1,$$

given that $y_1 = x$ is a solution of the complementary equation

$$x^2y'' + xy' - y = 0. \tag{14.6.11}$$

As a byproduct of this result, find a fundamental set of solutions of Equation 14.6.11.

b. Solve the initial value problem

$$x^2y'' + xy' - y = x^2 + 1, \quad y(1) = 2,\ y'(1) = -3. \tag{14.6.12}$$

Solution

a. If $y = ux$, then $y' = u'x + u$ and $y'' = u''x + 2u'$, so

$$\begin{aligned} x^2y'' + xy' - y &= x^2(u''x + 2u') + x(u'x + u) - ux \\ &= x^3u'' + 3x^2u'. \end{aligned}$$

Therefore $y = ux$ is a solution of Equation 14.6.12 if and only if

$$x^3u'' + 3x^2u' = x^2 + 1,$$

which is a first order equation in u'. We rewrite it as

$$u'' + \frac{3}{x}u' = \frac{1}{x} + \frac{1}{x^3}. \tag{14.6.13}$$

To focus on how we apply variation of parameters to this equation, we temporarily write $z = u'$, so that Equation 14.6.13 becomes

$$z' + \frac{3}{x}z = \frac{1}{x} + \frac{1}{x^3}. \tag{14.6.14}$$

We leave it to you to show by separation of variables that $z_1 = 1/x^3$ is a solution of the complementary equation

$$z' + \frac{3}{x}z = 0$$

for Equation 14.6.14. By variation of parameters, every solution of Equation 14.6.14 is of the form

$$z=\frac{v}{x^3}\quad\text{where}\quad\frac{v'}{x^3}=\frac{1}{x}+\frac{1}{x^3},\quad\text{so}\quad v'=x^2+1\quad\text{and}\quad v=\frac{x^3}{3}+x+C_1.$$

Since $u'=z=v/x^3$, u is a solution of Equation 14.6.14if and only if

$$u'=\frac{v}{x^3}=\frac{1}{3}+\frac{1}{x^2}+\frac{C_1}{x^3}.$$

Integrating this yields

$$u=\frac{x}{3}-\frac{1}{x}-\frac{C_1}{2x^2}+C_2.$$

Therefore the general solution of Equation 14.6.12is

$$y=ux=\frac{x^2}{3}-1-\frac{C_1}{2x}+C_2x. \tag{14.6.15}$$

Reasoning as in the solution of Example $14.6.1a$, we conclude that $y_1=x$ and $y_2=1/x$ form a fundamental set of solutions for Equation 14.6.11.

As we explained above, we rename the constants in Equation 14.6.15and rewrite it as

$$y=\frac{x^2}{3}-1+c_1x+\frac{c_2}{x}. \tag{14.6.16}$$

b. Differentiating Equation 14.6.16yields

$$y'=\frac{2x}{3}+c_1-\frac{c_2}{x^2}. \tag{14.6.17}$$

Setting $x=1$ in Equation 14.6.16and Equation 14.6.17and imposing the initial conditions $y(1)=2$ and $y'(1)=-3$ yields

$$\begin{aligned}c_1+c_2&=\frac{8}{3}\\ c_1-c_2&=-\frac{11}{3}.\end{aligned}$$

Solving these equations yields $c_1=-1/2$, $c_2=19/6$. Therefore the solution of Equation 14.6.12is

$$y=\frac{x^2}{3}-1-\frac{x}{2}+\frac{19}{6x}.$$

Here is a video on using reduction of order to solve a differential equation given one solution.



Using reduction of order to find the general solution of a homogeneous linear second order equation leads to a homogeneous linear first order equation in u' that can be solved by separation of variables. The next example illustrates this.

Example 4.6.3

Find the general solution and a fundamental set of solutions of

$$x^2y'' - 3xy' + 3y = 0, \tag{14.6.18}$$

given that $y_1 = x$ is a solution.

Solution

If $y = ux$ then $y' = u'x + u$ and $y'' = u''x + 2u'$, so

$$\begin{aligned} x^2y'' - 3xy' + 3y &= x^2(u''x + 2u') - 3x(u'x + u) + 3ux \\ &= x^3u'' - x^2u'. \end{aligned}$$

Therefore $y = ux$ is a solution of Equation 14.6.18 if and only if

$$x^3u'' - x^2u' = 0.$$

Separating the variables u' and x yields

$$\frac{u''}{u'} = \frac{1}{x},$$

so

$$\ln|u'| = \ln|x| + k, \quad \text{or equivalently} \quad u' = C_1x.$$

Therefore

$$u = \frac{C_1}{2}x^2 + C_2,$$

so the general solution of Equation 14.6.18 is

$$y = ux = \frac{C_1}{2}x^3 + C_2x,$$

which we rewrite as

$$y = c_1x + c_2x^3.$$

Therefore $\{x, x^3\}$ is a fundamental set of solutions of Equation 14.6.18.

Here is a video on using reduction of order to solve a differential equation given one solution.



This page titled 14.6: Reduction of Order is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **5.6: Reduction of Order** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

14.6E: Reduction of Order (Exercises)

Q5.6.1

In **Exercises 5.6.1-5.6.17** find the general solution, given that y_1 satisfies the complementary equation. As a byproduct, find a fundamental set of solutions of the complementary equation.

1. $(2x+1)y''-2y'-(2x+3)y=(2x+1)^2;\quad y_1=e^{-x}$
2. $x^2y''+xy'-y=\frac{4}{x^2};\quad y_1=x$
3. $x^2y''-xy'+y=x;\quad y_1=x$
4. $y''-3y'+2y=\frac{1}{1+e^{-x}};\quad y_1=e^{2x}$
5. $y''-2y'+y=7x^{3/2}e^x;\quad y_1=e^x$
6. $4x^2y''+(4x-8x^2)y'+(4x^2-4x-1)y=4x^{1/2}e^x(1+4x);\quad y_1=x^{1/2}e^x$
7. $y''-2y'+2y=e^x\sec x;\quad y_1=e^x\cos x$
8. $y''+4xy'+(4x^2+2)y=8e^{-x(x+2)};\quad y_1=e^{-x^2}$
9. $x^2y''+xy'-4y=-6x-4;\quad y_1=x^2$
10. $x^2y''+2x(x-1)y'+(x^2-2x+2)y=x^3e^{2x};\quad y_1=xe^{-x}$
11. $x^2y''-x(2x-1)y'+(x^2-x-1)y=x^2e^x;\quad y_1=xe^x$
12. $(1-2x)y''+2y'+(2x-3)y=(1-4x+4x^2)e^x;\quad y_1=e^x$
13. $x^2y''-3xy'+4y=4x^4;\quad y_1=x^2$
14. $2xy''+(4x+1)y'+(2x+1)y=3x^{1/2}e^{-x};\quad y_1=e^{-x}$
15. $xy''-(2x+1)y'+(x+1)y=-e^x;\quad y_1=e^x$
16. $4x^2y''-4x(x+1)y'+(2x+3)y=4x^{5/2}e^{2x};\quad y_1=x^{1/2}$
17. $x^2y''-5xy'+8y=4x^2;\quad y_1=x^2$

Q5.6.2

In **Exercises 5.6.18-5.6.30** find a fundamental set of solutions, given that y_1 is a solution.

18. $xy''+(2-2x)y'+(x-2)y=0;\quad y_1=e^x$
19. $x^2y''-4xy'+6y=0;\quad y_1=x^2$
20. $x^2(\ln|x|)^2y''-(2x\ln|x|)y'+(2+\ln|x|)y=0;\quad y_1=\ln|x|$
21. $4xy''+2y'+y=0;\quad y_1=\sin\sqrt{x}$
22. $xy''-(2x+2)y'+(x+2)y=0;\quad y_1=e^x$
23. $x^2y''-(2a-1)xy'+a^2y=0;\quad y_1=x^a$
24. $x^2y''-2xy'+(x^2+2)y=0;\quad y_1=x\sin x$
25. $xy''-(4x+1)y'+(4x+2)y=0;\quad y_1=e^{2x}$
26. $4x^2(\sin x)y''-4x(x\cos x+\sin x)y'+(2x\cos x+3\sin x)y=0;\quad y_1=x^{1/2}$
27. $4x^2y''-4xy'+(3-16x^2)y=0;\quad y_1=x^{1/2}e^{2x}$
28. $(2x+1)xy''-2(2x^2-1)y'-4(x+1)y=0;\quad y_1=1/x$
29. $(x^2-2x)y''+(2-x^2)y'+(2x-2)y=0;\quad y_1=e^x$
30. $xy''-(4x+1)y'+(4x+2)y=0;\quad y_1=e^{2x}$

Q5.6.3

In **Exercises 5.6.31-5.6.33** solve the initial value problem, given that y_1 satisfies the complementary equation.

31. $x^2y''-3xy'+4y=4x^4, \quad y(-1)=7, \quad y'(-1)=-8; \quad y_1=x^2$

32. $(3x-1)y''-(3x+2)y'-(6x-8)y=0, \quad y(0)=2,\ y'(0)=3; \quad y_1=e^{2x}$

33. $(x+1)^2y''-2(x+1)y'-(x^2+2x-1)y=(x+1)^3e^x, \quad y(0)=1, \quad y'(0)=-1; \quad y_1=(x+1)e^x$

Q5.6.4

In **Exercises 5.6.34** and **5.6.35** solve the initial value problem and graph the solution, given that y_1 satisfies the complementary equation.

34. $x^2y''+2xy'-2y=x^2, \quad y(1)=\frac{5}{4},\ y'(1)=\frac{3}{2}; \quad y_1=x$

35. $(x^2-4)y''+4xy'+2y=x+2, \quad y(0)=-\frac{1}{3}, \quad y'(0)=-1; \quad y_1=\frac{1}{x-2}$

Q5.6.5

36. Suppose p_1 and p_2 are continuous on (a,b). Let y_1 be a solution of

$$y''+p_1(x)y'+p_2(x)y=0 \tag{A}$$

that has no zeros on (a,b), and let x_0 be in (a,b). Use reduction of order to show that y_1 and

$$y_2(x)=y_1(x)\int_{x_0}^{x}\frac{1}{y_1^2(t)}\exp\left(-\int_{x_0}^{t}p_1(s)\,ds\right)dt$$

form a fundamental set of solutions of (A) on (a,b).

37. The nonlinear first order equation

$$y'+y^2+p(x)y+q(x)=0 \tag{A}$$

is a **Riccati equation**. (See **Exercise 2.4.55**.) Assume that p and q are continuous.

a. Show that y is a solution of (A) if and only if $y=z'/z$, where

$$z''+p(x)z'+q(x)z=0. \tag{B}$$

b. Show that the general solution of (A) is

$$y=\frac{c_1z_1'+c_2z_2'}{c_1z_1+c_2z_2}, \tag{C}$$

where $\{z_1,z_2\}$ is a fundamental set of solutions of (B) and c_1 and c_2 are arbitrary constants.

c. Does the formula (C) imply that the first order equation (A) has a two–parameter family of solutions? Explain your answer.

38. Use a method suggested by *Exercise 5.6.37* to find all solutions. of the equation.

a. $y'+y^2+k^2=0$
b. $y'+y^2-3y+2=0$
c. $y'+y^2+5y-6=0$
d. $y'+y^2+8y+7=0$
e. $y'+y^2+14y+50=0$
f. $6y'+6y^2-y-1=0$
g. $36y'+36y^2-12y+1=0$

39. Use a method suggested by **Exercise 5.6.37** and reduction of order to find all solutions of the equation, given that y_1 is a solution.

a. $x^2(y'+y^2)-x(x+2)y+x+2=0; \quad y_1=1/x$
b. $y'+y^2+4xy+4x^2+2=0; \quad y_1=-2x$
c. $(2x+1)(y'+y^2)-2y-(2x+3)=0; \quad y_1=-1$

d. $(3x-1)(y'+y^2)-(3x+2)y-6x+8=0; \quad y_1=2$
e. $x^2(y'+y^2)+xy+x^2-\frac{1}{4}=0; \quad y_1=-\tan x-\frac{1}{2x}$
f. $x^2(y'+y^2)-7xy+7=0; \quad y_1=1/x$

40. The nonlinear first order equation

$$y'+r(x)y^2+p(x)y+q(x)=0 \tag{A}$$

is the **generalized Riccati equation**. (See **Exercise 2.4.55**.) Assume that p and q are continuous and r is differentiable.

a. Show that y is a solution of (A) if and only if $y=z'/rz$, where

$$z''+\left[p(x)-\frac{r'(x)}{r(x)}\right]z'+r(x)q(x)z=0. \tag{B}$$

b. Show that the general solution of (A) is

$$y=\frac{c_1z_1'+c_2z_2'}{r(c_1z_1+c_2z_2)},$$

where $\{z_1,z_2\}$ is a fundamental set of solutions of (B) and c_1 and c_2 are arbitrary constants.

This page titled 14.6E: Reduction of Order (Exercises) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **5.6E: Reduction of Order (Exercises)** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

14.7: Variation of Parameters

In this section we give a method called *variation of parameters* for finding a particular solution of

$$P_0(x)y'' + P_1(x)y' + P_2(x)y = F(x) \tag{14.7.1}$$

if we know a fundamental set $\{y_1, y_2\}$ of solutions of the complementary equation

$$P_0(x)y'' + P_1(x)y' + P_2(x)y = 0. \tag{14.7.2}$$

Having found a particular solution y_p by this method, we can write the general solution of Equation 14.7.1 as

$$y = y_p + c_1y_1 + c_2y_2.$$

Since we need only one nontrivial solution of Equation 14.7.2 to find the general solution of Equation 14.7.1 by reduction of order, it is natural to ask why we are interested in variation of parameters, which requires two linearly independent solutions of Equation 14.7.2 to achieve the same goal. Here's the answer:

- If we already know two linearly independent solutions of Equation 14.7.2 then variation of parameters will probably be simpler than reduction of order.
- Variation of parameters generalizes naturally to a method for finding particular solutions of higher order linear equations (Section 9.4) and linear systems of equations (Section 10.7), while reduction of order doesn't.
- Variation of parameters is a powerful theoretical tool used by researchers in differential equations. Although a detailed discussion of this is beyond the scope of this book, you can get an idea of what it means from *Exercises 5.7.37–5.7.39.*

We'll now derive the method. As usual, we consider solutions of Equation 14.7.1 and Equation 14.7.2 on an interval (a, b) where P_0, P_1, P_2, and F are continuous and P_0 has no zeros. Suppose that $\{y_1, y_2\}$ is a fundamental set of solutions of the complementary equation Equation 14.7.2. We look for a particular solution of Equation 14.7.1 in the form

$$y_p = u_1y_1 + u_2y_2 \tag{14.7.3}$$

where u_1 and u_2 are functions to be determined so that y_p satisfies Equation 14.7.1. You may not think this is a good idea, since there are now two unknown functions to be determined, rather than one. However, since u_1 and u_2 have to satisfy only one condition (that y_p is a solution of Equation 14.7.1), we can impose a second condition that produces a convenient simplification, as follows.

Differentiating Equation 14.7.3 yields

$$y_p' = u_1y_1' + u_2y_2' + u_1'y_1 + u_2'y_2. \tag{14.7.4}$$

As our second condition on u_1 and u_2 we require that

$$u_1'y_1 + u_2'y_2 = 0. \tag{14.7.5}$$

Then Equation 14.7.4 becomes

$$y_p' = u_1y_1' + u_2y_2'; \tag{14.7.6}$$

that is, Equation 14.7.5 permits us to differentiate y_p (once!) as if u_1 and u_2 are constants. Differentiating Equation 14.7.4 yields

$$y_p'' = u_1y_1'' + u_2y_2'' + u_1'y_1' + u_2'y_2'. \tag{14.7.7}$$

(There are no terms involving u_1'' and u_2'' here, as there would be if we hadn't required Equation 14.7.5.) Substituting Equation 14.7.3, Equation 14.7.6, and Equation 14.7.7 into Equation 14.7.1 and collecting the coefficients of u_1 and u_2 yields

$$u_1(P_0y_1'' + P_1y_1' + P_2y_1) + u_2(P_0y_2'' + P_1y_2' + P_2y_2) + P_0(u_1'y_1' + u_2'y_2') = F.$$

As in the derivation of the method of reduction of order, the coefficients of u_1 and u_2 here are both zero because y_1 and y_2 satisfy the complementary equation. Hence, we can rewrite the last equation as

$$P_0(u_1'y_1' + u_2'y_2') = F. \tag{14.7.8}$$

Therefore y_p in Equation 14.7.3 satisfies Equation 14.7.1 if

$$\begin{aligned} u_1'y_1 + u_2'y_2 &= 0 \\ u_1'y_1' + u_2'y_2' &= \tfrac{F}{P_0}, \end{aligned} \tag{14.7.9}$$

where the first equation is the same as Equation 14.7.5 and the second is from Equation 14.7.8.

We'll now show that you can always solve Equation 14.7.9 for u_1' and u_2'. (The method that we use here will always work, but simpler methods usually work when you're dealing with specific equations.) To obtain u_1', multiply the first equation in Equation 14.7.9 by y_2' and the second equation by y_2. This yields

$$\begin{aligned} u_1'y_1y_2' + u_2'y_2y_2' &= 0 \\ u_1'y_1'y_2 + u_2'y_2'y_2 &= \frac{Fy_2}{P_0}. \end{aligned}$$

Subtracting the second equation from the first yields

$$u_1'(y_1y_2' - y_1'y_2) = -\frac{Fy_2}{P_0}. \tag{14.7.10}$$

Since $\{y_1, y_2\}$ is a fundamental set of solutions of Equation 14.7.2 on (a, b), Theorem 5.1.6 implies that the Wronskian $y_1y_2' - y_1'y_2$ has no zeros on (a, b). Therefore we can solve Equation 14.7.10 for u_1', to obtain

$$u_1' = -\frac{Fy_2}{P_0(y_1y_2' - y_1'y_2)}. \tag{14.7.11}$$

We leave it to you to start from Equation 14.7.9 and show by a similar argument that

$$u_2' = \frac{Fy_1}{P_0(y_1y_2' - y_1'y_2)}. \tag{14.7.12}$$

We can now obtain u_1 and u_2 by integrating u_1' and u_2'. The constants of integration can be taken to be zero, since any choice of u_1 and u_2 in Equation 14.7.3 will suffice.

You should not memorize Equation 14.7.11 and Equation 14.7.12 On the other hand, you don't want to rederive the whole procedure for every specific problem. We recommend the a compromise:

a. Write

$$y_p = u_1y_1 + u_2y_2 \tag{14.7.13}$$

to remind yourself of what you're doing.

b. Write the system

$$\begin{aligned} u_1'y_1 + u_2'y_2 &= 0 \\ u_1'y_1' + u_2'y_2' &= \tfrac{F}{P_0} \end{aligned} \tag{14.7.14}$$

for the specific problem you're trying to solve.

c. Solve Equation 14.7.14 for u_1' and u_2' by any convenient method.

d. Obtain u_1 and u_2 by integrating u_1' and u_2', taking the constants of integration to be zero.

e. Substitute u_1 and u_2 into Equation 14.7.13 to obtain y_p.

Below is a video on using variation of parameters to solve a differential equation.



Below is another video on using variation of parameters to solve a differential equation.



Example 4.7.1

Find a particular solution y_p of

$$x^2y'' - 2xy' + 2y = x^{9/2}, \tag{14.7.15}$$

given that $y_1 = x$ and $y_2 = x^2$ are solutions of the complementary equation

$$x^2y'' - 2xy' + 2y = 0.$$

Then find the general solution of Equation 14.7.15

Solution

We set

$$y_p = u_1x + u_2x^2,$$

where

$$\begin{aligned} u_1'x + \ \ u_2'x^2 \ &= 0 \\ u_1' \ \ + 2u_2'x \ &= \frac{x^{9/2}}{x^2} = x^{5/2}. \end{aligned}$$

From the first equation, $u_1' = -u_2'x$. Substituting this into the second equation yields $u_2'x = x^{5/2}$, so $u_2' = x^{3/2}$ and therefore $u_1' = -u_2'x = -x^{5/2}$. Integrating and taking the constants of integration to be zero yields

$$u_1 = -\frac{2}{7}x^{7/2} \quad \text{and} \quad u_2 = \frac{2}{5}x^{5/2}.$$

Therefore

$$y_p = u_1 x + u_2 x^2 = -\frac{2}{7}x^{7/2}x + \frac{2}{5}x^{5/2}x^2 = \frac{4}{35}x^{9/2},$$

and the general solution of Equation 14.7.15is

$$y = \frac{4}{35}x^{9/2} + c_1 x + c_2 x^2.$$

Below is a video on using variation of parameters to solve a differential equation.



Example 4.7.2

Find a particular solution y_p of

$$(x-1)y'' - xy' + y = (x-1)^2, \tag{14.7.16}$$

given that $y_1 = x$ and $y_2 = e^x$ are solutions of the complementary equation

$$(x-1)y'' - xy' + y = 0.$$

Then find the general solution of Equation 14.7.16

Solution

We set

$$y_p = u_1 x + u_2 e^x,$$

where

$$\begin{aligned} u_1' x + u_2' e^x &= 0 \\ u_1' \quad + u_2' e^x &= \frac{(x-1)^2}{x-1} = x-1. \end{aligned}$$

Subtracting the first equation from the second yields $-u_1'(x-1) = x-1$, so $u_1' = -1$. From this and the first equation, $u_2' = -xe^{-x}u_1' = xe^{-x}$. Integrating and taking the constants of integration to be zero yields

$$u_1 = -x \quad \text{and} \quad u_2 = -(x+1)e^{-x}.$$

Therefore

$$y_p = u_1 x + u_2 e^x = (-x)x + (-(x+1)e^{-x})e^x = -x^2 - x - 1,$$

so the general solution of Equation 14.7.16is

$$y = y_p + c_1 x + c_2 e^x = -x^2 - x - 1 + c_1 x + c_2 e^x = -x^2 - 1 + (c_1 - 1)x + c_2 e^x. \tag{14.7.17}$$

However, since c_1 is an arbitrary constant, so is $c_1 - 1$; therefore, we improve the appearance of this result by renaming the constant and writing the general solution as

$$y = -x^2 - 1 + c_1 x + c_2 e^x. \tag{14.7.18}$$

There's nothing *wrong* with leaving the general solution of Equation 14.7.16 in the form Equation 14.7.17; however, we think you'll agree that Equation 14.7.18 is preferable. We can also view the transition from Equation 14.7.17 to Equation 14.7.18 differently. In this example the particular solution $y_p = -x^2 - x - 1$ contained the term $-x$, which satisfies the complementary equation. We can drop this term and redefine $y_p = -x^2 - 1$, since $-x^2 - x - 1$ is a solution of Equation 14.7.16 and x is a solution of the complementary equation; hence, $-x^2 - 1 = (-x^2 - x - 1) + x$ is also a solution of Equation 14.7.16. In general, it is always legitimate to drop linear combinations of $\{y_1, y_2\}$ from particular solutions obtained by variation of parameters. (See *Exercise 5.7.36* for a general discussion of this question.) We'll do this in the following examples and in the answers to exercises that ask for a particular solution. Therefore, don't be concerned if your answer to such an exercise differs from ours only by a solution of the complementary equation.

Below is a video on verifying a solution to a differential equation and then solving it using variation of parameters.



Example 4.7.3

Find a particular solution of

$$y'' + 3y' + 2y = \frac{1}{1 + e^x}. \tag{14.7.19}$$

Then find the general solution.

Solution

The characteristic polynomial of the complementary equation

$$y'' + 3y' + 2y = 0 \tag{14.7.20}$$

is $p(r) = r^2 + 3r + 2 = (r+1)(r+2)$, so $y_1 = e^{-x}$ and $y_2 = e^{-2x}$ form a fundamental set of solutions of Equation 14.7.20. We look for a particular solution of Equation 14.7.19 in the form

$$y_p = u_1 e^{-x} + u_2 e^{-2x},$$

where

$$\begin{aligned} u_1' e^{-x} + u_2' e^{-2x} &= 0 \\ -u_1' e^{-x} - 2u_2' e^{-2x} &= \frac{1}{1 + e^x}. \end{aligned}$$

Adding these two equations yields

$$-u_2' e^{-2x} = \frac{1}{1+e^x}, \quad \text{so} \quad u_2' = -\frac{e^{2x}}{1+e^x}.$$

From the first equation,

$$u_1' = -u_2' e^{-x} = \frac{e^x}{1+e^x}.$$

Integrating by means of the substitution $v = e^x$ and taking the constants of integration to be zero yields

$$u_1 = \int \frac{e^x}{1+e^x}\,dx = \int \frac{dv}{1+v} = \ln(1+v) = \ln(1+e^x)$$

and

$$\begin{aligned} u_2 &= -\int \frac{e^{2x}}{1+e^x}\,dx = -\int \frac{v}{1+v}\,dv = \int \left[\frac{1}{1+v} - 1\right] dv \\ &= \ln(1+v) - v = \ln(1+e^x) - e^x. \end{aligned}$$

Therefore

$$\begin{aligned} y_p &= u_1 e^{-x} + u_2 e^{-2x} \\ &= [\ln(1+e^x)]e^{-x} + [\ln(1+e^x) - e^x]\, e^{-2x}, \end{aligned}$$

so

$$y_p = \left(e^{-x} + e^{-2x}\right) \ln(1+e^x) - e^{-x}.$$

Since the last term on the right satisfies the complementary equation, we drop it and redefine

$$y_p = \left(e^{-x} + e^{-2x}\right) \ln(1+e^x).$$

The general solution of Equation 14.7.19is

$$y = y_p + c_1 e^{-x} + c_2 e^{-2x} = \left(e^{-x} + e^{-2x}\right) \ln(1+e^x) + c_1 e^{-x} + c_2 e^{-2x}.$$

Example 4.7.4

Solve the initial value problem

$$(x^2-1)y'' + 4xy' + 2y = \frac{2}{x+1}, \quad y(0) = -1, \quad y'(0) = -5, \tag{14.7.21}$$

given that

$$y_1 = \frac{1}{x-1} \quad \text{and} \quad y_2 = \frac{1}{x+1}$$

are solutions of the complementary equation

$$(x^2-1)y'' + 4xy' + 2y = 0.$$

Solution

We first use variation of parameters to find a particular solution of

$$(x^2-1)y'' + 4xy' + 2y = \frac{2}{x+1}$$

on $(-1, 1)$ in the form

$$y_p = \frac{u_1}{x-1} + \frac{u_2}{x+1},$$

where

$$\frac{u_1'}{x-1}+\frac{u_2'}{x+1}=0 \tag{14.7.22}$$

$$-\frac{u_1'}{(x-1)^2}-\frac{u_2'}{(x+1)^2}=\frac{2}{(x+1)(x^2-1)}$$

Multiplying the first equation by $1/(x-1)$ and adding the result to the second equation yields

$$\left[\frac{1}{x^2-1}-\frac{1}{(x+1)^2}\right]u_2'=\frac{2}{(x+1)(x^2-1)}. \tag{14.7.23}$$

Since

$$\left[\frac{1}{x^2-1}-\frac{1}{(x+1)^2}\right]=\frac{(x+1)-(x-1)}{(x+1)(x^2-1)}=\frac{2}{(x+1)(x^2-1)},$$

Equation 14.7.23implies that $u_2'=1$. From Equation 14.7.22,

$$u_1'=-\frac{x-1}{x+1}u_2'=-\frac{x-1}{x+1}.$$

Integrating and taking the constants of integration to be zero yields

$$\begin{aligned}u_1&=-\int\frac{x-1}{x+1}\,dx=-\int\frac{x+1-2}{x+1}\,dx\\&=\int\left[\frac{2}{x+1}-1\right]dx=2\ln(x+1)-x\end{aligned}$$

and

$$u_2=\int dx=x.$$

Therefore

$$\begin{aligned}y_p&=\frac{u_1}{x-1}+\frac{u_2}{x+1}=[2\ln(x+1)-x]\frac{1}{x-1}+x\frac{1}{x+1}\\&=\frac{2\ln(x+1)}{x-1}+x\left[\frac{1}{x+1}-\frac{1}{x-1}\right]=\frac{2\ln(x+1)}{x-1}-\frac{2x}{(x+1)(x-1)}.\end{aligned}$$

However, since

$$\frac{2x}{(x+1)(x-1)}=\left[\frac{1}{x+1}+\frac{1}{x-1}\right]$$

is a solution of the complementary equation, we redefine

$$y_p=\frac{2\ln(x+1)}{x-1}.$$

Therefore the general solution of Equation 14.7.24is

$$y=\frac{2\ln(x+1)}{x-1}+\frac{c_1}{x-1}+\frac{c_2}{x+1}. \tag{14.7.24}$$

Differentiating this yields

$$y'=\frac{2}{x^2-1}-\frac{2\ln(x+1)}{(x-1)^2}-\frac{c_1}{(x-1)^2}-\frac{c_2}{(x+1)^2}.$$

Setting $x=0$ in the last two equations and imposing the initial conditions $y(0)=-1$ and $y'(0)=-5$ yields the system

$$\begin{aligned}-c_1+c_2&=-1\\-2-c_1-c_2&=-5.\end{aligned}$$

The solution of this system is $c_1 = 2,\ c_2 = 1$. Substituting these into Equation 14.7.24yields

$$\begin{aligned} y &= \frac{2\ln(x+1)}{x-1} + \frac{2}{x-1} + \frac{1}{x+1} \\ &= \frac{2\ln(x+1)}{x-1} + \frac{3x+1}{x^2-1} \end{aligned}$$

as the solution of Equation 14.7.21. Figure 4.7.1 is a graph of the solution.

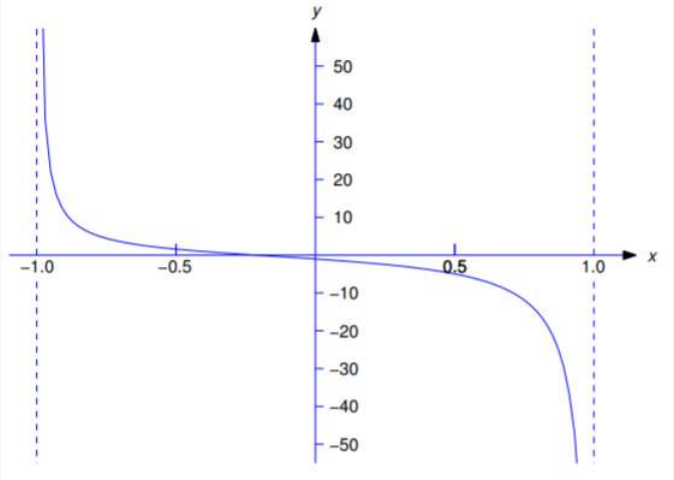


Figure 4.7.1 : $y = \frac{2\ln(x+1)}{x-1} + \frac{3x+1}{x^2-1}$

Below is a video on using variation of parameters to solve a differential equation.



We've now considered three methods for solving nonhomogeneous linear equations: undetermined coefficients, reduction of order, and variation of parameters. It's natural to ask which method is best for a given problem. The method of undetermined coefficients should be used for constant coefficient equations with forcing functions that are linear combinations of polynomials multiplied by functions of the form $e^{\alpha x}$, $e^{\lambda x}\cos\omega x$, or $e^{\lambda x}\sin\omega x$. Although the other two methods can be used to solve such problems, they will be more difficult except in the most trivial cases, because of the integrations involved.

If the equation isn't a constant coefficient equation or the forcing function isn't of the form just specified, the method of undetermined coefficients does not apply and the choice is necessarily between the other two methods. The case could be made that reduction of order is better because it requires only one solution of the complementary equation while variation of parameters requires two. However, variation of parameters will probably be easier if you already know a fundamental set of solutions of the complementary equation.

This page titled 14.7: Variation of Parameters is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **5.7: Variation of Parameters** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

14.7E: Variation of Parameters (Exercises)

Q5.7.1

In **Exercises 5.7.1–5.7.6** use variation of parameters to find a particular solution.

1. $y''+9y=\tan 3x$
2. $y''+4y=\sin 2x\sec^2 2x$
3. $y''-3y'+2y=\frac{4}{1+e^{-x}}$
4. $y''-2y'+2y=3e^x\sec x$
5. $y''-2y'+y=14x^{3/2}e^x$
6. $y''-y=\frac{4e^{-x}}{1-e^{-2x}}$

Q5.7.2

In **Exercises 5.7.7-5.7.29** use variation of parameters to find a particular solution, given the solutions y_1, y_2 of the complementary equation.

7. $x^2y''+xy'-y=2x^2+2;\quad y_1=x,\quad y_2=\frac{1}{x}$
8. $xy''+(2-2x)y'+(x-2)y=e^{2x};\quad y_1=e^x,\quad y_2=\frac{e^x}{x}$
9. $4x^2y''+(4x-8x^2)y'+(4x^2-4x-1)y=4x^{1/2}e^x,\quad x>0\ ;y_1=x^{1/2}e^x,\ y_2=x^{-1/2}e^x$
10. $y''+4xy'+(4x^2+2)y=4e^{-x(x+2)};\quad y_1=e^{-x^2},\quad y_2=xe^{-x^2}$
11. $x^2y''-4xy'+6y=x^{5/2},\ x>0;\quad y_1=x^2,\ y_2=x^3$
12. $x^2y''-3xy'+3y=2x^4\sin x;\quad y_1=x,\ y_2=x^3$
13. $(2x+1)y''-2y'-(2x+3)y=(2x+1)^2e^{-x};\quad y_1=e^{-x},\quad y_2=xe^x$
14. $4xy''+2y'+y=\sin\sqrt{x};\quad y_1=\cos\sqrt{x},\quad y_2=\sin\sqrt{x}$
15. $xy''-(2x+2)y'+(x+2)y=6x^3e^x;\quad y_1=e^x,\quad y_2=x^3e^x$
16. $x^2y''-(2a-1)xy'+a^2y=x^{a+1};\quad y_1=x^a,\quad y_2=x^a\ln x$
17. $x^2y''-2xy'+(x^2+2)y=x^3\cos x;\quad y_1=x\cos x,\quad y_2=x\sin x$
18. $xy''-y'-4x^3y=8x^5;\quad y_1=e^{x^2},\ y_2=e^{-x^2}$
19. $(\sin x)y''+(2\sin x-\cos x)y'+(\sin x-\cos x)y=e^{-x};\quad y_1=e^{-x},\quad y_2=e^{-x}\cos x$
20. $4x^2y''-4xy'+(3-16x^2)y=8x^{5/2};\quad y_1=\sqrt{x}e^{2x},\ y_2=\sqrt{x}e^{-2x}$
21. $4x^2y''-4xy'+(4x^2+3)y=x^{7/2};\quad y_1=\sqrt{x}\sin x,\ y_2=\sqrt{x}\cos x$
22. $x^2y''-2xy'-(x^2-2)y=3x^4;\quad y_1=xe^x,\ y_2=xe^{-x}$
23. $x^2y''-2x(x+1)y'+(x^2+2x+2)y=x^3e^x;\quad y_1=xe^x,\quad y_2=x^2e^x$
24. $x^2y''-xy'-3y=x^{3/2};\quad y_1=1/x,\quad y_2=x^3$
25. $x^2y''-x(x+4)y'+2(x+3)y=x^4e^x;\quad y_1=x^2,\quad y_2=x^2e^x$
26. $x^2y''-2x(x+2)y'+(x^2+4x+6)y=2xe^x;\quad y_1=x^2e^x,\quad y_2=x^3e^x$
27. $x^2y''-4xy'+(x^2+6)y=x^4;\quad y_1=x^2\cos x,\quad y_2=x^2\sin x$
28. $(x-1)y''-xy'+y=2(x-1)^2e^x;\quad y_1=x,\quad y_2=e^x$
29. $4x^2y''-4x(x+1)y'+(2x+3)y=x^{5/2}e^x;\quad y_1=\sqrt{x},\quad y_2=\sqrt{x}e^x$

Q5.7.3

In **Exercises 5.7.30–5.7.32** use variation of parameters to solve the initial value problem, given y_1, y_2 are solutions of the complementary equation.

30. $(3x-1)y''-(3x+2)y'-(6x-8)y=(3x-1)^2e^{2x}, \quad y(0)=1,\ y'(0)=2 \ ; y_1=e^{2x},\ y_2=xe^{-x}$

31. $(x-1)^2y''-2(x-1)y'+2y=(x-1)^2, \quad y(0)=3, \quad y'(0)=-6 \ ;$

$y_1=x-1\,, y_2=x^2-1$

32. $(x-1)^2y''-(x^2-1)y'+(x+1)y=(x-1)^3e^x, \quad y(0)=4, \quad y'(0)=-6 \ ;$

$y_1=(x-1)e^x, \quad y_2=x-1$

Q5.7.4

In **Exercises 5.7.33-5.7.35** use variation of parameters to solve the initial value problem and graph the solution, given that y_1, y_2 are solutions of the complementary equation.

33. $(x^2-1)y''+4xy'+2y=2x, \quad y(0)=0,\ y'(0)=-2; \quad y_1=\frac{1}{x-1},\ y_2=\frac{1}{x+1}$

34. $x^2y''+2xy'-2y=-2x^2, \quad y(1)=1,\ y'(1)=-1; \quad y_1=x,\ y_2=\frac{1}{x^2}$

35. $(x+1)(2x+3)y''+2(x+2)y'-2y=(2x+3)^2, \quad y(0)=0, \quad y'(0)=0 \ ; y_1=x+2, \quad y_2=\frac{1}{x+1}$

Q5.7.5

36. Suppose

$$y_p=\overline{y}+a_1y_1+a_2y_2$$

is a particular solution of

$$P_0(x)y''+P_1(x)y'+P_2(x)y=F(x), \tag{A}$$

where y_1 and y_2 are solutions of the complementary equation

$$P_0(x)y''+P_1(x)y'+P_2(x)y=0.$$

Show that $\overline{y}$ is also a solution of (A).

37. Suppose p, q, and f are continuous on (a,b) and let x_0 be in (a,b). Let y_1 and y_2 be the solutions of

$$y''+p(x)y'+q(x)y=0$$

such that

$$y_1(x_0)=1, \quad y_1'(x_0)=0, \quad y_2(x_0)=0, \quad y_2'(x_0)=1.$$

Use variation of parameters to show that the solution of the initial value problem

$$y''+p(x)y'+q(x)y=f(x), \quad y(x_0)=k_0,\ y'(x_0)=k_1,$$

is

$$\begin{aligned} y(x) &= k_0y_1(x)+k_1y_2(x) \\ &\quad +\int_{x_0}^{x}\left(y_1(t)y_2(x)-y_1(x)y_2(t)\right)f(t)\exp\left(\int_{x_0}^{t}p(s)\,ds\right)dt. \end{aligned}$$

HINT: **Use Abel's formula for the Wronskian of $\{y_1,y_2\}$, and integrate u_1' and u_2' from x_0 to x.**

Show also that

$$\begin{aligned} y'(x) &= k_0y_1'(x)+k_1y_2'(x) \\ &\quad +\int_{x_0}^{x}\left(y_1(t)y_2'(x)-y_1'(x)y_2(t)\right)f(t)\exp\left(\int_{x_0}^{t}p(s)\,ds\right)dt. \end{aligned}$$

38. Suppose f is continuous on an open interval that contains $x_0 = 0$. Use variation of parameters to find a formula for the solution of the initial value problem

$$y'' - y = f(x), \quad y(0) = k_0, \quad y'(0) = k_1.$$

39. Suppose f is continuous on (a, ∞), where $a < 0$, so $x_0 = 0$ is in (a, ∞).

a. Use variation of parameters to find a formula for the solution of the initial value problem

$$y'' + y = f(x), \quad y(0) = k_0, \quad y'(0) = k_1.$$

HINT: **You will need the addition formulas for the sine and cosine.**

$$\begin{aligned} \sin(A+B) &= \sin A \cos B + \cos A \sin B \\ \cos(A+B) &= \cos A \cos B - \sin A \sin B \end{aligned}$$

For the rest of this exercise assume that the improper integral $\int_0^\infty f(t)\,dt$ is absolutely convergent.

b. Show that if y is a solution of

$$y'' + y = f(x) \tag{A}$$

on (a, ∞), then

$$\lim_{x\to\infty} (y(x) - A_0 \cos x - A_1 \sin x) = 0 \tag{B}$$

and

$$\lim_{x\to\infty} (y'(x) + A_0 \sin x - A_1 \cos x) = 0, \tag{C}$$

where

$$A_0 = k_0 - \int_0^\infty f(t) \sin t\, dt \quad \text{and} \quad A_1 = k_1 + \int_0^\infty f(t) \cos t\, dt.$$

HINT: Recall from calculus that if $\int_0^\infty f(t)dt$ *converges absolutely, then* $\lim_{x\to\infty} \int_x^\infty |f(t)|dt = 0$.

c. Show that if A_0 and A_1 are arbitrary constants, then there's a unique solution of $y'' + y = f(x)$ on (a, ∞) that satisfies (B) and (C).

This page titled 14.7E: Variation of Parameters (Exercises) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by William F. Trench.

- **5.7E: Variation of Parameters (Exercises)** by William F. Trench is licensed CC BY-NC-SA 3.0. Original source: https://digitalcommons.trinity.edu/mono/9.

14.8: Mechanical Vibrations

Let us look at some applications of linear second order constant coefficient equations.

14.8.1: Some examples

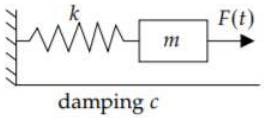


Figure 14.8.1

Our first example is a mass on a spring. Suppose we have a mass $m > 0$ (in kilograms) connected by a spring with spring constant $k > 0$ (in newtons per meter) to a fixed wall. There may be some external force $F(t)$ (in newtons) acting on the mass. Finally, there is some friction measured by $c \geq 0$ (in newton-seconds per meter) as the mass slides along the floor (or perhaps there is a damper connected).

Let x be the displacement of the mass ($x = 0$ is the rest position), with x growing to the right (away from the wall). The force exerted by the spring is proportional to the compression of the spring by Hooke's law. Therefore, it is kx in the negative direction. Similarly the amount of force exerted by friction is proportional to the velocity of the mass. By Newton's second law we know that force equals mass times acceleration and hence $mx'' = F(t) - cx' - kx$ or

$$mx'' + cx' + kx = F(t)$$

This is a linear second order constant coefficient ODE. We set up some terminology about this equation. We say the motion is

i. forced, if $F \not\equiv 0$ (if F is not identically zero),
ii. unforced or free, if $F \equiv 0$ (if F is identically zero),
iii. damped, if $c > 0$, and
iv. undamped, if $c = 0$.

This system appears in lots of applications even if it does not at first seem like it. Many real-world scenarios can be simplified to a mass on a spring. For example, a bungee jump setup is essentially a mass and spring system (you are the mass). It would be good if someone did the math before you jump off the bridge, right? Let us give two other examples.

Here is an example for electrical engineers. Consider the pictured RLC circuit. There is a resistor with a resistance of R ohms, an inductor with an inductance of L henries, and a capacitor with a capacitance of C farads. There is also an electric source (such as a battery) giving a voltage of $E(t)$ volts at time t (measured in seconds). Let $Q(t)$ be the charge in coulombs on the capacitor and $I(t)$ be the current in the circuit. The relation between the two is $Q' = I$. By elementary principles we find $LI' + RI + \frac{Q}{C} = E$. We differentiate to get

$$LI''(t) + RI'(t) + \frac{1}{C}I(t) = E'(t).$$

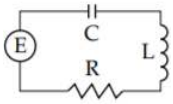


Figure 14.8.2

This is a nonhomogeneous second order constant coefficient linear equation. As L, R, and C are all positive, this system behaves just like the mass and spring system. Position of the mass is replaced by current. Mass is replaced by inductance, damping is replaced by resistance, and the spring constant is replaced by one over the capacitance. The change in voltage becomes the forcing function—for constant voltage this is an unforced motion.

Our next example behaves like a mass and spring system only approximately. Suppose a mass m hangs on a pendulum of length L. We seek an equation for the angle $\theta(t)$ (in radians). Let g be the force of gravity. Elementary physics mandates that the equation is

$$\theta'' + \frac{g}{L}\sin\theta = 0.$$

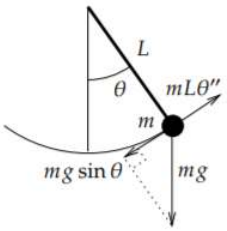


Figure 14.8.3

Let us derive this equation using Newton's second law: force equals mass times acceleration. The acceleration is $L\theta''$ and mass is m. So $mL\theta''$ has to be equal to the tangential component of the force given by the gravity, which is $mg\sin\theta$ in the opposite direction. So $mL\theta'' = -mg\sin\theta$. The m curiously cancels from the equation.

Now we make our approximation. For small θ we have that approximately $\sin\theta \approx \theta$. This can be seen by looking at the graph. In Figure 14.8.4we can see that for approximately $-0.5 < \theta < 0.5$ (in radians) the graphs of $\sin\theta$ and θ are almost the same.

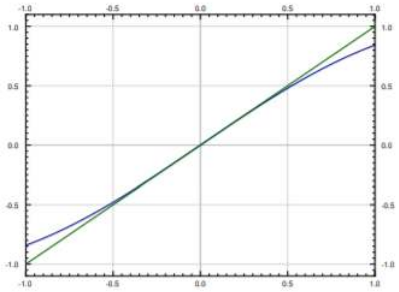


Figure 14.8.4: The graphs of $\sin\theta$ and θ (in radians).

Therefore, when the swings are small, θ is small and we can model the behavior by the simpler linear equation

$$\theta'' + \frac{g}{L}\theta = 0.$$

The errors from this approximation build up. So after a long time, the state of the real-world system might be substantially different from our solution. Also we will see that in a mass-spring system, the amplitude is independent of the period. This is not true for a pendulum. Nevertheless, for reasonably short periods of time and small swings (that is, only small angles θ), the approximation is reasonably good.

In real-world problems it is often necessary to make these types of simplifications. We must understand both the mathematics and the physics of the situation to see if the simplification is valid in the context of the questions we are trying to answer.

14.8.2: Free Undamped Motion

In this section we will only consider free or unforced motion, as we cannot yet solve nonhomogeneous equations. Let us start with undamped motion where $c = 0$. We have the equation

$$mx'' + kx = 0$$

If we divide by m and let $w_0 = \sqrt{\frac{k}{m}}$, then we can write the equation as

$$x'' + w_0^2 x = 0$$

The general solution to this equation is

$$x(t) = A\cos(w_0 t) + B\sin(w_0 t)$$

By a trigonometric identity, we have that for two different constants C and γ, we have

$$A\cos(w_0 t) + B\sin(w_0 t) = C\cos(w_0 t - \gamma)$$

It is not hard to compute that $C = \sqrt{A^2 + B^2}$ and $\tan\gamma = \dfrac{B}{A}$. Therefore, we let C and γ be our arbitrary constants and write $x(t) = C\cos(w_0 t - \gamma)$.

Below is a video on solving a differential equation that comes from a free undamped spring system.



Exercise 14.8.1

Justify the above identity and verify the equations for C and γ. Hint: Start with $\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta)$ and multiply by C. Then think what should α and β be.

While it is generally easier to use the first form with A and B to solve for the initial conditions, the second form is much more natural. The constants C and γ have very nice interpretation. We look at the form of the solution

$$x(t) = C\cos(w_0 t - \gamma)$$

We can see that the amplitude is C, w_0 is the (angular) frequency, and γ is the so-called phase shift. The phase shift just shifts the graph left or right. We call w_0 the natural (angular) frequency. This entire setup is usually called simple harmonic motion.

Let us pause to explain the word angular before the word frequency. The units of w_0 are radians per unit time, not cycles per unit time as is the usual measure of frequency. Because we know one cycle is 2π radians, the usual frequency is given by $\dfrac{w_0}{2\pi}$. It is simply a matter of where we put the constant 2π, and that is a matter of taste.

The period of the motion is one over the frequency (in cycles per unit time) and hence $\dfrac{2\pi}{w_0}$. That is the amount of time it takes to complete one full oscillation.

Below is a video on solving a initial value problem that comes from a free undamped spring system.



Example 14.8.1

Suppose that $m = 2kg$ and $k = 8\dfrac{N}{m}$. The whole mass and spring setup is sitting on a truck that was traveling at $1\dfrac{m}{s}$. The truck crashes and hence stops. The mass was held in place 0.5 meters forward from the rest position. During the crash the mass gets loose. That is, the mass is now moving forward at $1\dfrac{m}{s}$, while the other end of the spring is held in place. The mass therefore starts oscillating. What is the frequency of the resulting oscillation and what is the amplitude. The units are the mks units (meters-kilograms-seconds).

The setup means that the mass was at half a meter in the positive direction during the crash and relative to the wall the spring is mounted to, the mass was moving forward (in the positive direction) at $1\dfrac{m}{s}$. This gives us the initial conditions.

So the equation with initial conditions is

$$2x'' + 8x = 0, \quad x(0) = 0.5, \quad x'(0) = 1$$

We can directly compute $w_0 = \sqrt{\dfrac{k}{m}} = \sqrt{4} = 2$. Hence the angular frequency is 2. The usual frequency in Hertz (cycles per second) is $\dfrac{2}{2\pi} = \dfrac{1}{\pi} \approx 0.318$.

The general solution is

$$x(t) = A\cos(2t) + B\sin(2t)$$

Letting $x(0) = 0.5$ means $A = 0.5$. Then $x'(t) = -2(0.5)\sin(2t) + 2B\cos(2t)$. Letting $x'(0) = 1$ we get $B = 0.5$. Therefore, the amplitude is $C = \sqrt{A^2 + B^2} = \sqrt{0.25 + 0.25} = \sqrt{0.5} \approx 0.707$. The solution is

$$x(t) = 0.5\cos(2t) + 0.5\sin(2t)$$

A plot of $x(t)$ is shown in Figure 14.8.5

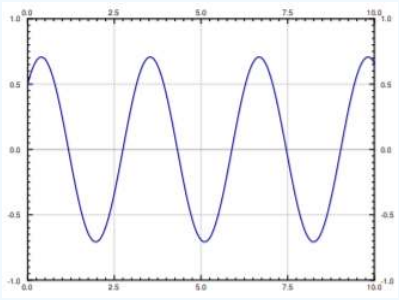


Figure 14.8.5: Simple undamped oscillation.

In general, for free undamped motion, a solution of the form

$$x(t) = A\cos(w_0 t) + B\sin(w_0 t)$$

corresponds to the initial conditions $x(0) = A$ and $x'(0) = w_0 B$. Therefore, it is easy to figure out A and B from the initial conditions. The amplitude and the phase shift can then be computed from A and B. In the example, we have already found the amplitude C. Let us compute the phase shift. We know that $\tan\gamma = \dfrac{B}{A} = 1$. We take the arctangent of 1 and get approximately 0.785. We still need to check if this γ is in the correct quadrant (and add π to γ if it is not). Since both A and B are positive, then γ should be in the first quadrant, and 0.785 radians really is in the first quadrant.

Note

Many calculators and computer software do not only have the atan function for arctangent, but also what is sometimes called `atan2` . This function takes two arguments, B and A, and returns a γ in the correct quadrant for you.

Below is a video on solving a differential equation that comes from an undamped spring system.



14.8.3: Free Damped Motion

Let us now focus on damped motion. Let us rewrite the equation

$$mx'' + cx' + kx = 0$$

as

$$x'' + 2px' + w_0^2 x = 0$$

where

$$w_0 = \sqrt{\frac{k}{m}}, \quad p = \frac{c}{2m}$$

The characteristic equation is

$$r^2 + 2pr + w_0^2 = 0$$

Using the quadratic formula we get that the roots are

$$r = -p \pm \sqrt{p^2 - w_0^2}$$

The form of the solution depends on whether we get complex or real roots. We get real roots if and only if the following number is nonnegative:

$$p^2 - w_0^2 = \left(\frac{c}{2m}\right)^2 - \frac{k}{m} = \frac{c^2 - 4km}{4m^2}$$

The sign of $p^2 - w_0^2$ is the same as the sign of $c^2 - 4km$. Thus we get real roots if and only if $c^2 - 4km$ is nonnegative, or in other words if $c^2 \geq 4km$.

14.8.3.1: Overdamping

When $c^2 - 4km > 0$, we say the system is overdamped. In this case, there are two distinct real roots r_1 and r_2. Notice that both roots are negative. As $\sqrt{p^2 - w_0^2}$ is always less than P, then $-P \pm \sqrt{P^2 - w_0^2}$ is negative.

The solution is

$$x(t) = C_1 e^{r_1 t} + C_2 e^{r_2 t}$$

Since r_1, r_2 are negative, $x(t) \to 0$ as $t \to \infty$. Thus the mass will tend towards the rest position as time goes to infinity. For a few sample plots for different initial conditions (Figure 14.8.6).

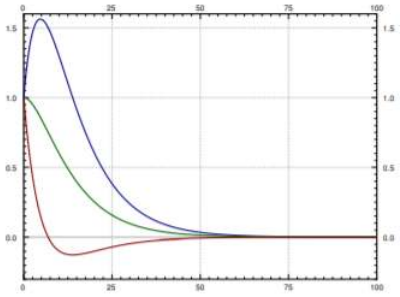


Figure 14.8.6: Overdamped motion for several different initial conditions.

Do note that no oscillation happens. In fact, the graph will cross the x axis at most once. To see why, we try to solve $0 = C_1 e^{r_1 t} + C_2 e^{r_2 t}$. Therefore, $C_1 e^{r_1 t} = -C_2 e^{r_2 t}$ and using laws of exponents we obtain

$$\frac{-C_1}{C_2} = e^{(r_2 - r_1)t}$$

This equation has at most one solution $t \geq 0$. For some initial conditions the graph will never cross the x axis, as is evident from the sample graphs.

Example 14.8.2

Suppose the mass is released from rest. That is $x(0) = x_0$ and $x'(0) = 0$. Then

$$x(t) = \frac{x_0}{r_1 - r_2}(r_1 e^{r_2 t} - r_2 e^{r_1 t})$$

It is not hard to see that this satisfies the initial conditions.

Below is a video on solving a differential equation that comes from a damped spring system.



14.8.3.2: Critical damping

When $c^2 - 4km = 0$, we say the system is critically damped. In this case, there is one root of multiplicity 2 and this root is $-P$. Therefore, our solution is

$$x(t) = C_1 e^{-pt} + C_2 t e^{-pt}$$

The behavior of a critically damped system is very similar to an overdamped system. After all a critically damped system is in some sense a limit of overdamped systems. Since these equations are really only an approximation to the real world, in reality we are never critically damped, it is a place we can only reach in theory. We are always a little bit underdamped or a little bit overdamped. It is better not to dwell on critical damping.

14.8.3.3: Underdamping

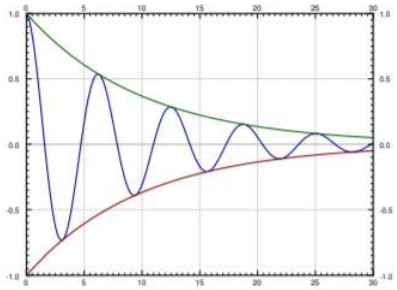


Figure 14.8.7: Underdamped motion with the envelope curves shown.

When $c^2 - 4km < 0$, we say the system is underdamped. In this case, the roots are complex.

$$\begin{aligned} r &= -p \pm \sqrt{p^2 - w_0^2} \\ &= -p \pm \sqrt{-1}\sqrt{w_0^2 - p^2} \\ &= -p \pm iw_1 \end{aligned} \tag{14.8.1}$$

where $w_1 = \sqrt{w_0^2 - p^2}$. Our solution is

$$x(t) = e^{-pt}(A\cos(w_1 t) + B\sin(w_1 t)$$

or

$$x(t) = Ce^{-pt}\cos(w_1 t - \gamma)$$

An example plot is given in Figure 14.8.7. Note that we still have that $x(t) \to 0$ as $t \to \infty$.

In the figure we also show the *envelope curves* Ce^{-pt} and $-Ce^{pt}$. The solution is the oscillating line between the two envelope curves. The envelope curves give the maximum amplitude of the oscillation at any given point in time. For example if you are bungee jumping, you are really interested in computing the envelope curve so that you do not hit the concrete with your head.

The phase shift γ just shifts the graph left or right but within the envelope curves (the envelope curves do not change if γ changes).

Finally note that the angular pseudo-frequency[1] (we do not call it a frequency since the solution is not really a periodic function) w_1 becomes smaller when the damping c (and hence P) becomes larger. This makes sense. When we change the damping just a little bit, we do not expect the behavior of the solution to change dramatically. If we keep making c larger, then at some point the solution should start looking like the solution for critical damping or overdamping, where no oscillation happens. So if c^2 approaches $4km$, we want w_1 to approach 0.

On the other hand when c becomes smaller, w_1 approaches w_0 (w_1 is always smaller than w_0), and the solution looks more and more like the steady periodic motion of the undamped case. The envelope curves become flatter and flatter as c (and hence P) goes to 0.

14.8.4: Footnotes

[1] We do not call ω_1 a frequency since the solution is not really a periodic function.

This page titled 14.8: Mechanical Vibrations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **2.4: Mechanical Vibrations** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

14.9: Nonhomogeneous Equations

14.9.1: Solving Nonhomogeneous Equations

We have solved linear constant coefficient homogeneous equations. What about nonhomogeneous linear ODEs? For example, the equations for forced mechanical vibrations. That is, suppose we have an equation such as

$$y'' + 5y' + 6y = 2x + 1 \tag{14.9.1}$$

We will write $Ly = 2x + 1$ when the exact form of the operator is not important. We solve (Equation 14.9.1) in the following manner. First, we find the general solution y_c to the associated homogeneous equation

$$y'' + 5y' + 6y = 0 \tag{14.9.2}$$

We call y_c the *complementary solution*. Next, we find a single particular solution y_p to (14.9.1) in some way. Then

$$y = y_c + y_p$$

is the general solution to (14.9.1). We have $Ly_c = 0$ and $Ly_p = 2x + 1$. As L is a linear operator we verify that y is a solution, $Ly = L(y_c + y_p) = Ly_c + Ly_p = 0 + (2x + 1)$. Let us see why we obtain the general solution.

Let y_p and $\tilde{y}_p$ be two different particular solutions to (14.9.1). Write the difference as $w = y_p - \tilde{y}_p$. Then plug w into the left hand side of the equation to get

$$w'' + 5w' + 6w = (y_p'' + 5y_p' + 6y_p) - (\tilde{y}_p'' + 5\tilde{y}_p' + 6\tilde{y}_p) = (2x + 1) - (2x + 1) = 0$$

Using the operator notation the calculation becomes simpler. As L is a linear operator we write

$$Lw = L(y_p - \tilde{y}_p) = Ly_p - L\tilde{y}_p = (2x + 1) - (2x + 1) = 0$$

So $w = y_p - \tilde{y}_p$ is a solution to (14.9.2), that is $Lw = 0$. Any two solutions of (14.9.1) differ by a solution to the homogeneous equation (14.9.2). The solution $y = y_c + y_p$ includes all solutions to (14.9.1), since y_c is the general solution to the associated homogeneous equation.

Theorem 14.9.1

Let $Ly = f(x)$ be a linear ODE (not necessarily constant coefficient). Let y_c be the complementary solution (the general solution to the associated homogeneous equation $Ly = 0$) andlet y_p be any particular solution to $Ly = f(x)$. Then the general solution to $Ly = f(x)$ is

$$y = y_c + y_p.$$

The moral of the story is that we can find the particular solution in any old way. If we find a different particular solution (by a different method, or simply by guessing), then we still get the same general solution. The formula may look different, and the constants we will have to choose to satisfy the initial conditions may be different, but it is the same solution.

14.9.1.1: Undetermined Coefficients

The trick is to somehow, in a smart way, guess one particular solution to (14.9.1). Note that $2x + 1$ is a polynomial, and the left hand side of the equation will be a polynomial if we let y be a polynomial of the same degree. Let us try

$$y_p = Ax + B$$

We plug in to obtain

$$\begin{aligned} y_p'' + 5y_p' + 6y_p &= (Ax + B)'' + 5(Ax + B)' + 6(Ax + B) \\ &= 0 + 5A + 6Ax + 6B = 6Ax + (5A + 6B) \end{aligned} \tag{14.9.3}$$

So $6Ax + (5A + 6B) = 2x + 1$. Therefore, $A = \frac{1}{3}$ and $B = -\frac{1}{9}$. That means $y_p = \frac{1}{3}x - \frac{1}{9} = \frac{3x - 1}{9}$. Solving the complementary problem (exercise!) we get

$$y_c = C_1 e^{-2x} + C_2 e^{-3x}$$

Hence the general solution to (14.9.1) is

$$y = C_1 e^{-2x} + C_2 e^{-3x} + \frac{3x-1}{9}$$

Now suppose we are further given some initial conditions. For example, $y(0) = 0$ and $y'(0) = \frac{1}{3}$. First find $y' = -2C_1 e^{-2x} - 3C_2 e^{-3x} + \frac{1}{3}$. Then

$$0 = y(0) = C_1 + C_2 - \frac{1}{9}, \frac{1}{3} = y'(0) = -2C_1 - 3C_2 + \frac{1}{3}$$

We solve to get $C_1 = \frac{1}{3}$ and $C_2 = -\frac{2}{9}$. The particular solution we want is

$$y(x) = \frac{1}{3}e^{-2x} - \frac{2}{9}e^{-3x} + \frac{3x-1}{9} = \frac{3e^{-2x} - 2e^{-3x} + 3x - 1}{9}$$

? Exercise 14.9.1

Check that y really solves the equation (14.9.1) and the given initial conditions.

Note

A common mistake is to solve for constants using the initial conditions with y_c and only add the particular solution y_p after that. That will not work. You need to first compute $y = y_c + y_p$ and only then solve for the constants using the initial conditions.

A right hand side consisting of exponentials, sines, and cosines can be handled similarly. For example,

$$y'' + 2y' + 2y = \cos(2x)$$

Let us find some y_p. We start by guessing the solution includes some multiple of $\cos(2x)$. We may have to also add a multiple of $\sin(2x)$ to our guess since derivatives of cosine are sines. We try

$$y_p = A\cos(2x) + B\sin(2x)$$

We plug y_p into the equation and we get

$$\underbrace{-4A\cos(2x) - 4B\sin(2x)}_{y_p''} + 2\underbrace{\left(-2A\sin(2x) + 2B\cos(2x)\right)}_{y_p'}$$
$$+2\underbrace{\left(A\cos(2x) + 2B\sin(2x)\right)}_{y_p} = \cos(2x),$$

The left hand side must equal to right hand side. We group terms and we get that $-4A + 4B + 2A = 1$ and $-4B - 4A + 2B = 0$. So $-2A + 4B = 1$ and $2A + B = 0$ and hence $A = \frac{-1}{10}$ and $B = \frac{1}{5}$. So

$$y_p = A\cos(2x) + B\sin(2x) = \frac{-\cos(2x) + 2\sin(2x)}{10}$$

Similarly, if the right hand side contains exponentials we try exponentials. For example, for

$$Ly = e^{3x}$$

we will try $y = Ae^{3x}$ as our guess and try to solve for A.

When the right hand side is a multiple of sines, cosines, exponentials, and polynomials, we can use the product rule for differentiation to come up with a guess. We need to guess a form for y_p such that Ly_p is of the same form, and has all the terms

needed to for the right hand side. For example,

$$Ly = (1+3x^2)e^{-x}\cos(\pi x)$$

For this equation, we will guess

$$y_p = (A+Bx+Cx^2)e^{-x}\cos(\pi x) + (D+Ex+Fx^2)e^{-x}\sin(\pi x)$$

We will plug in and then hopefully get equations that we can solve for $A, B, C, D, E$and F. As you can see this can make for a very long and tedious calculation very quickly.

There is one hiccup in all this. It could be that our guess actually solves the associated homogeneous equation. That is, suppose we have

$$y'' - 9y = e^{3x}$$

We would love to guess $y = Ae^{3x}$, but if we plug this into the left hand side of the equation we get

$$y'' - 9y = 9Ae^{3x} - 9Ae^{3x} = 0 \neq e^{3x}$$

There is no way we can choose A to make the left hand side be e^{3x}. The trick in this case is to multiply our guess by x to get rid of duplication with the complementary solution. That is first we compute y_c (solution to $Ly = 0$)

$$y_c = C_1e^{-3x} + C_2e^{3x}$$

and we note that the e^{3x} term is a duplicate with our desired guess. We modify our guess to $y = Axe^{3x}$ and notice there is no duplication anymore. Let us try. Note that $y' = Ae^{3x} + 3Axe^{3x}$ and $y'' = 6Ae^{3x} + 9Axe^{3x}$. So

$$y'' - 9y = 6Ae^{3x} + 9Axe^{3x} - 9Axe^{3x} = 6Ae^{3x}$$

Thus $6Ae^{3x}$ is supposed to equal e^{3x}. Hence, $6A = 1$ and so $A = \dfrac{1}{6}$. We can now write the general solution as

$$y = y_c + y_p = C_1e^{-3x} + C_2e^{3x} + \frac{1}{6}xe^{3x}$$

It is possible that multiplying by x does not get rid of all duplication. For example,

$$y'' - 6y' + 9y = e^{3x}$$

The complementary solution is $y_c = C_1e^{3x} + C_2xe^{3x}$. Guessing $y = Axe^{3x}$ would not get us anywhere. In this case we want to guess $y_p = Ax^2e^{3x}$. Basically, we want to multiply our guess by x until all duplication is gone. But no more! Multiplying too many times will not work.

Finally, what if the right hand side has several terms, such as

$$Ly = e^{2x} + \cos x$$

In this case we find u that solves $Lu = e^{2x}$ and v that solves $Lv = \cos x$ (that is, do each term separately). Then note that if $y = u + v$, then $Ly = e^{2x} + \cos x$. This is because L is linear; we have $Ly = L(u+v) = Lu + Lv = e^{2x} + \cos x$.

14.9.1.2: Variation of Parameters

The method of undetermined coefficients will work for many basic problems that crop up. But it does not work all the time. It only works when the right hand side of the equation $Ly = f(x)$ has only finitely many linearly independent derivatives, so that we can write a guess that consists of them all. Some equations are a bit tougher. Consider

$$y'' + y = \tan x$$

Note that each new derivative of $\tan x$ looks completely different and cannot be written as a linear combination of the previous derivatives. If we start differentiating $\tan x$, we get

$$\sec^2 x, \quad 2\sec^2 x\,\tan x, \quad 4\sec^2 x\,\tan^2 x + 2\sec^4 x,$$
$$8\sec^2 x\,\tan^3 x + 16\sec^4 x\,\tan x, \quad 16\sec^2 x\,\tan^4 x + 88\sec^4 x\tan^2 x + 16\sec^6 x, \quad \ldots$$

This equation calls for a different method. We present the method of variation of parameters, which will handle any equation of the form $Ly = f(x)$, provided we can solve certain integrals. For simplicity, we restrict ourselves to second order constant coefficient equations, but the method works for higher order equations just as well (the computations become more tedious). The method also works for equations with nonconstant coefficients, provided we can solve the associated homogeneous equation.

Perhaps it is best to explain this method by example. Let us try to solve the equation

$$Ly = y'' + y = \tan x$$

First we find the complementary solution (solution to $Ly_c = 0$). We get $y_c = C_1 y_1 + C_2 y_2$, where $y_1 = \cos x$ and $y_2 = \sin x$. To find a particular solution to the nonhomogeneous equation we try

$$y_p = y = u_1 y_1 + u_2 y_2$$

where u_1 and u_2 are functions and not constants. We are trying to satisfy $Ly = \tan x$. That gives us one condition on the functions u_1 and u_2. Compute (note the product rule!)

$$y' = (u_1' y_1 + u_2' y_2) + (u_1 y_1' + u_2 y_2')$$

We can still impose one more condition at our discretion to simplify computations (we have two unknown functions, so we should be allowed two conditions). We require that $(u_1' y_1 + u_2' y_2) = 0$. This makes computing the second derivative easier.

$$\begin{aligned} y' &= u_1 y_1' + u_2 y_2' \\ y'' &= (u_1' y_1' + u_2' y_2') + (u_1 y_1'' + u_2 y_2'') \end{aligned} \tag{14.9.4}$$

Since y_1 and y_2 are solutions to $y'' + y = 0$, we know that $y_1'' = -y_1$ and $y_2'' = -y_2$. (Note: If the equation was instead $y'' + p(x)y' + q(x)y = 0$ we would have $y_i'' = -p(x)y_i' - q(x)y_i$.) So

$$y'' = (u_1' y_1' + u_2' y_2') - (u_1 y_1 + u_2 y_2)$$

We have $(u_1 y_1 + u_2 y_2) = y$ and so

$$y'' = (u_1' y_1' + u_2' y_2') - y$$

and hence

$$y'' + y = Ly = u_1' y_1' + u_2' y_2'$$

For y to satisfy $Ly = f(x)$ we must have $f(x) = u_1' y_1' + u_2' y_2'$.

So what we need to solve are the two equations (conditions) we imposed on u_1 and u_2

$$\begin{aligned} u_1' y_1 + u_2' y_2 &= 0 \\ u_1' y_1' + u_2' y_2' &= f(x) \end{aligned} \tag{14.9.5}$$

We can now solve for u_1' and u_2' in terms of $f(x), y_1$ and y_2. We will always get these formulas for any $Ly = f(x)$, where $Ly = y'' + p(x)y' + q(x)y$. There is a general formula for the solution we can just plug into, but it is better to just repeat what we do below. In our case the two equations become

$$\begin{aligned} u_1' \cos(x) + u_2' \sin(x) &= 0 \\ -u_1' \sin(x) + u_2' \cos(x) &= \tan(x) \end{aligned} \tag{14.9.6}$$

Hence

$$\begin{aligned} u_1' \cos(x)\sin(x) + u_2' \sin^2(x) &= 0 \\ -u_1' \sin(x)\cos(x) + u_2' \cos^2(x) &= \tan(x)\cos(x) = \sin(x) \end{aligned} \tag{14.9.7}$$

And thus

$$\begin{aligned} u_2'(\sin^2(x) + \cos^2(x)) &= \sin(x) \\ u_2' &= \sin(x) \\ u_1' &= \frac{-\sin^2(x)}{\cos(x)} = -\tan(x)\sin(x) \end{aligned} \tag{14.9.8}$$

Now we need to integrate u_1' and u_2' to get u_1 and u_2.

$$\begin{aligned} u_1 &= \int u_1' dx = \int -\tan(x)\sin(x) dx = \frac{1}{2}\ln\left|\frac{\sin(x)-1}{\sin(x)+1}\right| + \sin(x) \\ u_2 &= \int u_2' dx = \int \sin(x) dx = -\cos(x) \end{aligned} \tag{14.9.9}$$

So our particular solution is

$$\begin{aligned} y_p &= u_1 y_1 + u_2 y_2 = \frac{1}{2}\cos(x)\ln\left|\frac{\sin(x)-1}{\sin(x)+1}\right| + \cos(x)\sin(x) - \cos(x)\sin(x) \\ &= \frac{1}{2}\cos(x)\ln\left|\frac{\sin(x)-1}{\sin(x)+1}\right| \end{aligned} \tag{14.9.10}$$

The general solution to $y'' + y = \tan x$ is, therefore,

$$y = C_1\cos(x) + C_2\sin(x) + \frac{1}{2}\cos(x)\ln\left|\frac{\sin(x)-1}{\sin(x)+1}\right|$$

- Jiří Lebl (Oklahoma State University).These pages were supported by NSF grants DMS-0900885 and DMS-1362337.

This page titled 14.9: Nonhomogeneous Equations is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **2.5: Nonhomogeneous Equations** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

14.10: Forced Oscillations and Resonance

Let us consider to the example of a mass on a spring. We now examine the case of forced oscillations, which we did not yet handle. That is, we consider the equation

$$mx'' + cx' + kx = F(t)$$

for some nonzero $F(t)$. The setup is again: m is mass, c is friction, k is the spring constant, and $F(t)$ is an external force acting on the mass.

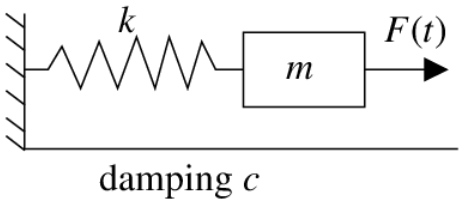


Figure 14.10.1

What we are interested in is periodic forcing, such as noncentered rotating parts, or perhaps loud sounds, or other sources of periodic force. Once we learn about Fourier series in Chapter 4, we will see that we cover all periodic functions by simply considering $F(t) = F_0 \cos(\omega t)$ (or sine instead of cosine, the calculations are essentially the same).

Below is a video on solving a forced oscillations problem.



14.10.1: Undamped Forced Motion and Resonance

First let us consider undamped $c = 0$ motion for simplicity. We have the equation

$$mx'' + kx = F_0 \cos(\omega t)$$

This equation has the complementary solution (solution to the associated homogeneous equation)

$$x_c = C_1 \cos(\omega_0 t) + C_2 \sin(\omega_0 t)$$

where $\omega_0 = \sqrt{\frac{k}{m}}$ is the *natural frequency* (angular), which is the frequency at which the system "wants to oscillate" without external interference.

Let us suppose that $\omega_0 \neq \omega$. We try the solution $x_p = A\cos(\omega t)$ and solve for A. Note that we need not have sine in our trial solution as on the left hand side we will only get cosines anyway. If you include a sine it is fine; you will find that its coefficient will be zero.

We solve using the method of undetermined coefficients. We find that

$$x_p = \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos(\omega t)$$

We leave it as an exercise to do the algebra required.

The general solution is

$$x = C_1 \cos(\omega_0 t) + C_2 \sin(\omega_0 t) + \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos(\omega t)$$

or written another way

$$x = C \cos(\omega_0 t - y) + \frac{F_0}{m(\omega_0^2 - \omega^2)} \cos(\omega t)$$

Hence it is a superposition of two cosine waves at different frequencies.

Example 14.10.1

Take

$$0.5x'' + 8x = 10\cos(\pi t), \quad x(0) = 0, \quad x'(0) = 0$$

Let us compute. First we read off the parameters: $\omega = \pi, \omega_0 = \sqrt{\frac{8}{0.5}} = 4, F_0 = 10, m = 0.5$. The general solution is

$$x = C_1 \cos(4t) + C_2 \sin(4t) + \frac{20}{16 - \pi^2} \cos(\pi t)$$

Solve for C_1 and C_2 using the initial conditions. It is easy to see that $C_1 = \frac{-20}{16-\pi^2}$ and $C_2 = 0$. Hence

$$x = \frac{20}{16 - \pi^2} (\cos(\pi t) - \cos(4t))$$

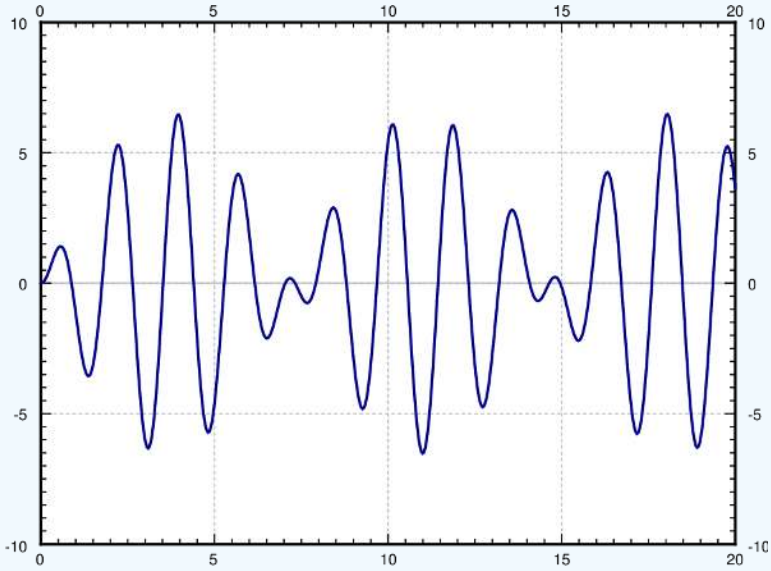


Figure 14.10.2: Graph of $\frac{20}{16-\pi^2}(\cos(\pi t) - \cos(4t))$.

Notice the “beating” behavior in Figure 14.10.2 First use the trigonometric identity

$$2\sin(\frac{A-B}{2})\sin(\frac{A+B}{2}) = \cos B - \cos A$$

to get that

$$x = \frac{20}{16 - \pi^2} (2\sin(\frac{4-\pi}{2}t)\sin(\frac{4+\pi}{2}t))$$

Notice that x is a high frequency wave modulated by a low frequency wave.

Now suppose that $\omega_0 = \omega$. Obviously, we cannot try the solution $A\cos(\omega t)$ and then use the method of undetermined coefficients. We notice that $\cos(\omega t)$ solves the associated homogeneous equation. Therefore, we need to try $x_p = At\cos(\omega t) + Bt\sin(\omega t)$. This time we do need the sine term since the second derivative of $t\cos(\omega t)$ does contain sines. We write the equation

$$x'' + \omega^2 x = \frac{F_0}{m}\cos(\omega t)$$

Plugging x_p into the left hand side we get

$$2B\omega\cos(\omega t) - 2A\omega\sin(\omega t) = \frac{F_0}{m}\cos(\omega t)$$

Hence $A = 0$ and $B = \frac{F_0}{2m\omega}$. Our particular solution is $\frac{F_0}{2m\omega}t\sin(\omega t)$ and our general solution is

$$x = C_1\cos(\omega t) + C_2\sin(\omega t) + \frac{F_0}{2m\omega}t\sin(\omega t)$$

The important term is the last one (the particular solution we found). We can see that this term grows without bound as $t \to \infty$. In fact it oscillates between $\frac{F_0 t}{2m\omega}$ and $\frac{-F_0 t}{2m\omega}$. The first two terms only oscillate between $\pm\sqrt{C_1^2 + C_2^2}$, which becomes smaller and smaller in proportion to the oscillations of the last term as t gets larger. In Figure 14.10.3 we see the graph with $C_1 = C_2 = 0, F_0 = 2, m = 1, \omega = \pi$.

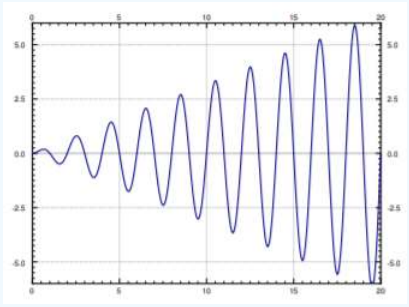


Figure 14.10.3: Graph of $\frac{1}{\pi}t\sin(\pi t)$.

By forcing the system in just the right frequency we produce very wild oscillations. This kind of behavior is called resonance or perhaps pure resonance. Sometimes resonance is desired. For example, remember when as a kid you could start swinging by just moving back and forth on the swing seat in the “correct frequency”? You were trying to achieve resonance. The force of each one of your moves was small, but after a while it produced large swings.

On the other hand resonance can be destructive. In an earthquake some buildings collapse while others may be relatively undamaged. This is due to different buildings having different resonance frequencies. So figuring out the resonance frequency can be very important.

A common (but wrong) example of destructive force of resonance is the Tacoma Narrows bridge failure. It turns out there was a different phenomenon at play.[1]

Below is a video on solving a differential equation that comes from a vibrating system with resonance.



14.10.2: Damped Forced Motion and Practical Resonance

In real life things are not as simple as they were above. There is, of course, some damping. Our equation becomes

$$mx'' + cx' + kx = F_0 \cos(\omega t), \tag{14.10.1}$$

for some $c > 0$. We have solved the homogeneous problem before. We let

$$p = \frac{c}{2m} \quad \omega_0 = \sqrt{\frac{k}{m}}$$

We replace equation (14.10.1)with

$$x'' + 2px' + \omega_0^2 x = \frac{F_0}{m}\cos(\omega t)$$

The roots of the characteristic equation of the associated homogeneous problem are $r_1, r_2 = -p \pm \sqrt{p^2 - \omega_0^2}$. The form of the general solution of the associated homogeneous equation depends on the sign of $p^2 - \omega_0^2$, or equivalently on the sign of $c^2 - 4km$, as we have seen before. That is,

$$x_c = \begin{cases} C_1 e^{r_1 t} + C_2 e^{r_2 t}, & \text{if } c^2 > 4km, \\ C_1 e^{pt} + C_2 t e^{-pt}, & \text{if } c^2 = 4km, \\ e^{-pt}(C_1 \cos(\omega_1 t) + C_2 \sin(\omega_1 t)), & \text{if } c^2 < 4km, \end{cases}$$

where $\omega_1 = \sqrt{\omega_0^2 - p^2}$. In any case, we can see that $x_c(t) \to 0$ as $t \to \infty$. Furthermore, there can be no conflicts when trying to solve for the undetermined coefficients by trying $x_p = A\cos(\omega t) + B\sin(\omega t)$. Let us plug in and solve for A and B. We get (the tedious details are left to reader)

$$((\omega_0^2 - \omega^2)B - 2\omega p A)\sin(\omega t) + ((\omega_0^2 - \omega^2)A + 2\omega p B)\cos(\omega t) = \frac{F_0}{m}\cos(\omega t)$$

We get that

$$A = \frac{(\omega_0^2 - \omega^2)F_0}{m(2\omega p)^2 + m\left(\omega_0^2 - \omega^2\right)^2}$$

$$B = \frac{2\omega p F_0}{m(2\omega p)^2 + m\left(\omega_0^2 - \omega^2\right)^2}$$

We also compute $C = \sqrt{A^2 + B^2}$ to be

$$C = \frac{F_0}{m\sqrt{(2\omega p)^2 + \left(\omega_0^2 - \omega^2\right)^2}}$$

Thus our particular solution is

$$x_P = \frac{(\omega_0^2 - \omega^2)F_0}{m(2\omega p)^2 + m(\omega_0^2 - \omega^2)^2}\cos(\omega t) + \frac{2\omega p F_0}{m(2\omega p)^2 + m(\omega_0^2 - \omega^2)^2}\sin(\omega t)$$

Or in the alternative notation we have amplitude C and phase shift γ where (if $\omega \neq \omega_0$)

$$\tan\gamma = \frac{B}{A} = \frac{2\omega p}{\omega_0^2 - \omega^2}$$

Hence we have

$$x_p = \frac{F_0}{m\sqrt{(2\omega p)^2 + (\omega_0^2 - \omega^2)^2}}\cos(\omega t - \gamma)$$

If $\omega = \omega_0$ we see that $A = 0, B = C = \frac{F_0}{2m\omega p}$, and $\gamma = \frac{\pi}{2}$.

The exact formula is not as important as the idea. Do not memorize the above formula, you should instead remember the ideas involved. For different forcing function F, you will get a different formula for x_p. So there is no point in memorizing this specific formula. You can always recompute it later or look it up if you really need it.

For reasons we will explain in a moment, we call $x_c$the transient solution and denote it by x_{tr}. We call the x_p we found above the steady periodic solution and denote it by x_{sp}. The general solution to our problem is

$$x = x_c + x_p = x_{tr} + x_{sp}$$

We note that $x_c = x_{tr}$ goes to zero as $t \to \infty$, as all the terms involve an exponential with a negative exponent. Hence for large t, the effect of x_{tr} is negligible and we will essentially only see x_{sp}. Hence the name transient. Notice that x_{sp} involves no arbitrary constants, and the initial conditions will only affect x_{tr}. This means that the effect of the initial conditions will be negligible after some period of time. Because of this behavior, we might as well focus on the steady periodic solution and ignore the transient solution. See Figure 14.10.4for a graph of different initial conditions.

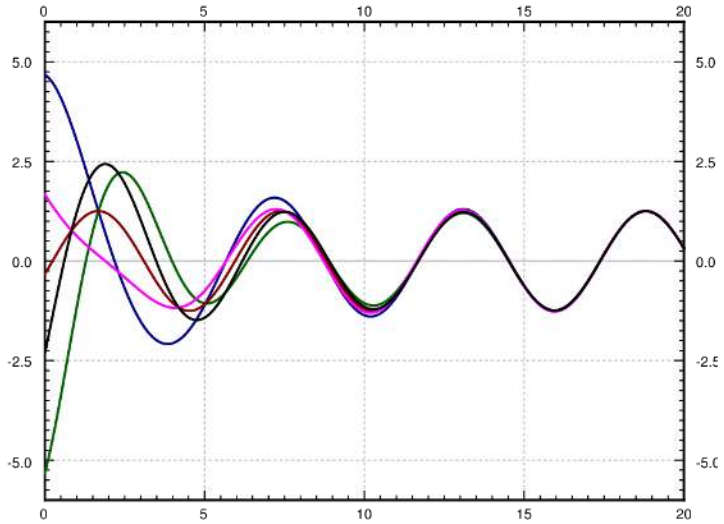


Figure 14.10.4: Solutions with different initial conditions for parameters $k = 1, m = 1, F_0 = 1, c = 0.7,$ and $\omega = 1.1$.

Notice that the speed at which x_{tr} goes to zero depends on P (and hence c). The bigger P is (the bigger c is), the "faster" x_{tr} becomes negligible. So the smaller the damping, the longer the "transient region." This agrees with the observation that when $c = 0$, the initial conditions affect the behavior for all time (i.e. an infinite "transient region").

Let us describe what we mean by resonance when damping is present. Since there were no conflicts when solving with undetermined coefficient, there is no term that goes to infinity. What we will look at however is the maximum value of the amplitude of the steady periodic solution. Let C be the amplitude of x_{sp}. If we plot C as a function of ω (with all other parameters fixed) we can find its maximum. We call the ω that achieves this maximum the practical resonance frequency. We call the maximal amplitude $C(\omega)$ the practical resonance amplitude. Thus when damping is present we talk of practical resonance rather than pure

resonance. A sample plot for three different values of c is given in Figure 14.10.5 As you can see the practical resonance amplitude grows as damping gets smaller, and any practical resonance can disappear when damping is large.

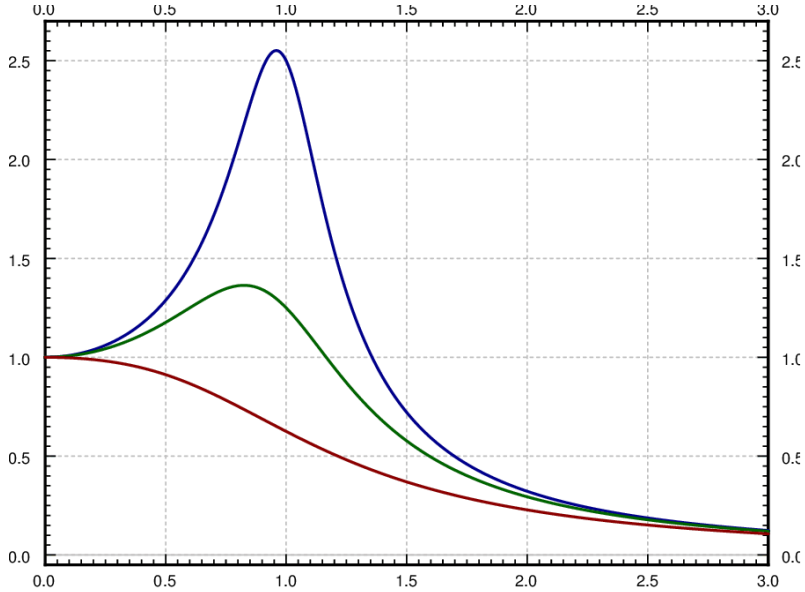


Figure 14.10.5: Graph of $C(\omega)$ showing practical resonance with parameters $k = 1, m = 1, F_0 = 1$. The top line is with $c = 0.4$, the middle line with $c = 0.8$, and the bottom line with $c = 1.6$.

To find the maximum we need to find the derivative $C'(\omega)$. Computation shows

$$C'(\omega) = \frac{-4\omega(2p^2 + \omega^2 - \omega_0^2)F_0}{m\left((2\omega p)^2 + (\omega_0^2 - \omega^2)\right)^{3/2}}$$

This is zero either when $\omega = 0$ or when $2p^2 + \omega^2 - \omega_0^2 = 0$. In other words, $C'(\omega) = 0$ when

$$\omega = \sqrt{\omega_0^2 - 2p^2} \text{ or } \omega = 0$$

It can be shown that if $\omega_0^2 - 2p^2$ is positive, then $\sqrt{\omega_0^2 - 2p^2}$ is the practical resonance frequency (that is the point where $C(\omega)$ is maximal, note that in this case $C'(\omega) > 0$ for small ω). If $\omega = 0$ is the maximum, then essentially there is no practical resonance since we assume that $\omega > 0$ in our system. In this case the amplitude gets larger as the forcing frequency gets smaller.

If practical resonance occurs, the frequency is smaller than ω_0. As the damping c (and hence P) becomes smaller, the practical resonance frequency goes to ω_0. So when damping is very small, ω_0 is a good estimate of the resonance frequency. This behavior agrees with the observation that when $c = 0$, then ω_0 is the resonance frequency.

Another interesting observation to make is that when $\omega \to \infty$, then $\omega \to 0$. This means that if the forcing frequency gets too high it does not manage to get the mass moving in the mass-spring system. This is quite reasonable intuitively. If we wiggle back and forth really fast while sitting on a swing, we will not get it moving at all, no matter how forceful. Fast vibrations just cancel each other out before the mass has any chance of responding by moving one way or the other.

The behavior is more complicated if the forcing function is not an exact cosine wave, but for example a square wave. A general periodic function will be the sum (superposition) of many cosine waves of different frequencies. The reader is encouraged to come back to this section once we have learned about the Fourier series.

14.10.3: Footnotes

[1]K. Billah and R. Scanlan, Resonance, Tacoma Narrows Bridge Failure, and Undergraduate Physics Textbooks, American Journal of Physics, 59(2), 1991, 118–124, http://www.ketchum.org/billah/Billah-Scanlan.pdf

This page titled 14.10: Forced Oscillations and Resonance is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **2.6: Forced Oscillations and Resonance** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

14.E: Higher order linear ODEs (Exercises)

These are homework exercises to accompany Libl's "Differential Equations for Engineering" Textmap. This is a textbook targeted for a one semester first course on differential equations, aimed at engineering students. Prerequisite for the course is the basic calculus sequence.

14.E.1: 2.1: Second order linear ODEs

Exercise 14.E.2.1.1

Show that $y = e^x$ and $y = e^{2x}$ are linearly independent.

Exercise 14.E.2.1.2

Take $y'' + 5y = 10x + 5$. Find (guess!) a solution.

Exercise 14.E.2.1.3

Prove the superposition principle for nonhomogeneous equations. Suppose that y_1 is a solution to $Ly_1 = f(x)$ and y_2 is a solution to $Ly_2 = g(x)$ (same linear operator L). Show that $y = y_1 + y_2$ solves $Ly = f(x) + g(x)$.

Exercise 14.E.2.1.4

For the equation $x^2y'' - xy' = 0$, find two solutions, show that they are linearly independent and find the general solution. Hint: Try $y = x^r$.

Equations of the form $ax^2y'' + bxy' + cy = 0$ are called *Euler's equations* or *Cauchy-Euler equations*. They are solved by trying $y = x^r$ and solving for r (we can assume that $x \geqslant 0$ for simplicity).

Exercise 14.E.2.1.5

Suppose that $(b-a)^2 - 4ac > 0$.

a. Find a formula for the general solution of $ax^2y'' + bxy' + cy = 0$. Hint: Try $y = x^r$ and find a formula for r.
b. What happens when $(b-a)^2 - 4ac = 0$ or $(b-a)^2 - 4ac < 0$?

We will revisit the case when $(b-a)^2 - 4ac < 0$ later.

Exercise 14.E.2.1.6

Same equation as in Exercise 14.E.2.1.5. Suppose $(b-a)^2 - 4ac = 0$. Find a formula for the general solution of $ax^2y'' + bxy' + cy = 0$. Hint: Try $y = x^r \ln x$ for the second solution.

Exercise 14.E.2.1.7: reduction of order

Suppose y_1 is a solution to $y'' + p(x)y' + q(x)y = 0$. Show that

$$y_2(x) = y_1(x) \int \frac{e^{-\int p(x)dx}}{(y_1(x))^2} dx \tag{14.E.1}$$

is also a solution.

Note: If you wish to come up with the formula for reduction of order yourself, start by trying $y_2(x) = y_1(x)v(x)$. Then plug y_2 into the equation, use the fact that y_1 is a solution, substitute $w = v'$, and you have a first order linear equation in w. Solve for w and then for v. When solving for w, make sure to include a constant of integration. Let us solve some famous equations using the method.

Exercise 14.E.2.1.8: Chebyshev's equation of order 1

Take $(1-x^2)y''-xy'+y=0$.

a. Show that $y=x$ is a solution.
b. Use reduction of order to find a second linearly independent solution.
c. Write down the general solution.

Exercise 14.E.2.1.9: Hermite's equation of order 2

Take $y''-2xy'+4y=0$.

a. Show that $y=1-2x^2$ is a solution.
b. Use reduction of order to find a second linearly independent solution.
c. Write down the general solution.

Exercise 14.E.2.1.10

Are $\sin(x)$ and e^x linearly independent? Justify.

Answer

Yes. To justify try to find a constant A such that $\sin(x)=Ae^x$ for all x.

Exercise 14.E.2.1.11

Are e^x and e^{x+2} linearly independent? Justify.

Answer

No. $e^{x+2}=e^2e^x$.

Exercise 14.E.2.1.12

Guess a solution to $y''+y'+y=5$.

Answer

$y=5$

Exercise 14.E.2.1.13

Find the general solution to $xy''+y'=0$. Hint: Notice that it is a first order ODE in y'.

Answer

$y=C_1\ln(x)+C_2$

Exercise 14.E.2.1.14

Write down an equation (guess) for which we have the solutions e^x and e^{2x}. Hint: Try an equation of the form $y''+Ay'+By=0$ for constants A and B, plug in both e^x and e^{2x} and solve for A and B.

Answer

$y''-3y'+2y=0$

https://math.libretexts.org/@go/page/128087

14.E.2: 2.2: Constant coefficient second order linear ODEs

Exercise 14.E.2.2.1

Find the general solution of $2y'' + 2y' - 4y = 0$.

Exercise 14.E.2.2.2

Find the general solution of $y'' + 9y' - 10y = 0$.

Exercise 14.E.2.2.3

Solve $y'' - 8y' + 16y = 0$ for $y(0) = 2, y'(0) = 0$.

Exercise 14.E.2.2.4

Solve $y'' + 9y' = 0$ for $y(0) = 1, y'(0) = 1$.

Exercise 14.E.2.2.5

Find the general solution of $2y'' + 50y = 0$.

Exercise 14.E.2.2.6

Find the general solution of $y'' + 6y' + 13y = 0$.

Exercise 14.E.2.2.7

Find the general solution of $y'' = 0$ using the methods of this section.

Exercise 14.E.2.2.8

The method of this section applies to equations of other orders than two. We will see higher orders later. Try to solve the first order equation $2y' + 3y = 0$ using the methods of this section.

Exercise 14.E.2.2.9

Let us revisit Euler's equations of Exercise $14.E.1$. Suppose now that $(b-a)^2 - 4ac < 0$. Find a formula for the general solution of $ax^2y'' + bxy' + cy = 0$. Hint: Note that $x^r = e^{r \ln x}$.

Exercise 14.E.2.2.10

Find the solution to $y'' - (2\alpha)y' + \alpha^2 y = 0$, $y(0) = a$, $y'(0) = b$, where α, a, and b are real numbers.

Exercise 14.E.2.2.11

Construct an equation such that $y = C_1 e^{-2x}\cos(3x) + C_2 e^{-2x}\sin(3x)$ is the general solution.

Exercise 14.E.2.2.12

Find the general solution to $y'' + 4y' + 2y = 0$.

Answer

$y = C_1 e^{(-2+\sqrt{2})x} + C_2 e^{(-2-\sqrt{2})x}$

Exercise 14.E.2.2.13

Find the general solution to $y''-6y'+9y=0$.

Answer

$y=C_1e^{3x}+C_2xe^{3x}$

Exercise 14.E.2.2.14

Find the solution to $2y''+y'+y=0, y(0)=1, y'(0)=-2$.

Answer

$y=e^{-x/4}\cos\left((\frac{\sqrt{7}}{4})x\right)-\sqrt{7}e^{-x/4}\sin\left((\frac{\sqrt{7}}{4})x\right)$

Exercise 14.E.2.2.15

Find the solution to $2y''+y'-3y=0, y(0)=a, y'(0)=b$.

Answer

$y=\frac{2(a-b)}{5}e^{-3x/2}+\frac{3a+2b}{5}e^{x}$

Exercise 14.E.2.2.16

Find the solution to $z''(t)=-2z'(t)-2z(t), z(0)=2, z'(0)=-2$.

Answer

$z(t)=2e^{-t}\cos(t)$

Exercise 14.E.2.2.17

Find the solution to $y''-(\alpha+\beta)y'+\alpha\beta y=0$, $y(0)=a$, $y'(0)=b$, where α, (\beta\), a, and b are real numbers, and $\alpha\neq\beta$.

Answer

$y=\frac{\alpha\beta-b}{\beta-\alpha}e^{\alpha x}+\frac{b-a\alpha}{\beta-\alpha}e^{\beta x}$

Exercise 14.E.2.2.18

Construct an equation such that $y=C_1e^{3x}+C_2e^{-2x}$ is the general solution.

Answer

$y''-y'-6y=0$

14.E.3: 2.3: Higher order linear ODEs

Exercise 14.E.2.3.1

Find the general solution for $y''' - y'' + y' - y = 0$.

Exercise 14.E.2.3.2

Find the general solution for $y^{(4)} - 5y''' + 6y'' = 0$.

Exercise 14.E.2.3.3

Find the general solution for $y''' + 2y'' + 2y' = 0$.

Exercise 14.E.2.3.4

Suppose the characteristic equation for a differential equation is $(r-1)^2(r-2)^2 = 0$.

a. Find such a differential equation.
b. Find its general solution.

Exercise 14.E.2.3.5

Suppose that a fourth order equation has a solution $y = 2e^{4x}x\cos x$.

a. Find such an equation.
b. Find the initial conditions that the given solution satisfies.

Exercise 14.E.2.3.6

Find the general solution for the equation of Exercise 14.E.2.3.5

Exercise 14.E.2.3.7

Let $f(x) = e^x - \cos x, g(x) = e^x + \cos x$ and $h(x) = \cos x$. Are $f(x), g(x)$, and $h(x)$ and linearly independent? If so, show it, if not, find a linear combination that works.

Exercise 14.E.2.3.8

Let $f(x) = 0, g(x) = \cos x$, and $h(x) = \sin x$. Are $f(x), g(x)$, and $h(x)$ and linearly independent? If so, show it, if not, find a linear combination that works.

Exercise 14.E.2.3.9

Are $x, x^2, and x^4$ linearly independent? If so, show it, if not, find a linear combination that works.

Exercise 14.E.2.3.10

Are $e^x, xe^x, and x^2e^x$ linearly independent? If so, show it, if not, find a linear combination that works.

Exercise 14.E.2.3.11

Find an equation such that $y = xe^{-2x}\sin(3x)$ is a solution.

Exercise 14.E.2.3.12

Find the general solution of $y^{(5)} - y^{(4)} = 0$.

Answer

$y = C_1 e^x + C_2 x^3 + C_3 x^2 + C_4 x + C_5$

Exercise 14.E.2.3.13

Suppose that the characteristic equation of a third order differential equation has roots $3 \pm 2i$.

a. What is the characteristic equation?
b. Find the corresponding differential equation.
c. Find the general solution.

Answer

a. $r^3 - 3r^2 + 4r - 12 = 0$
b. $y''' - 3y'' + 4y' - 12y = 0$
c. $y = C_1 e^{3x} + C_2 \sin(2x) + C_3 \cos(2x)$

Exercise 14.E.2.3.14

Solve $1001y''' + 3.2y'' + \pi y' - \sqrt{4}y = 0, y(0) = 0, y'(0) = 0, y''(0) = 0$.

Answer

$y = 0$

Exercise 14.E.2.3.15

Are $e^x, e^{x+1}, e^{2x}, \sin(x)$ linearly independent? If so, show it, if not find a linear combination that works.

Answer

No. $e^1 e^x - e^{x+1} = 0$.

Exercise 14.E.2.3.16

Are $\sin(x), x, x\sin(x)$ linearly independent? If so, show it, if not find a linear combination that works.

Answer

Yes. (Hint: First note that $\sin(x)$ is bounded. Then note that x and $x\sin(x)$ cannot be multiples of each other.)

Exercise 14.E.2.3.17

Find an equation such that $y = \cos(x)$, $y = \sin(x)$, $y = e^x$ are solutions.

Answer

$y''' - y'' + y' - y = 0$

14.E.4: 2.4: Mechanical Vibrations

Exercise 14.E.2.4.1

Consider a mass and spring system with a mass $m = 2$, spring constant $k = 3$, and damping constant $c = 1$.

a. Set up and find the general solution of the system.
b. Is the system underdamped, overdamped or critically damped?
c. If the system is not critically damped, find a c that makes the system critically damped.

Exercise 14.E.2.4.2

Do Exercise 14.E.2.4.1 for $m = 3, k = 12, and c = 12$.

Exercise 14.E.2.4.3

Using the mks units (meters-kilograms-seconds), suppose you have a spring with spring constant $4\frac{N}{m}$. You want to use it to weigh items. Assume no friction. You place the mass on the spring and put it in motion.

a. You count and find that the frequency is 0.8 Hz (cycles per second). What is the mass?
b. Find a formula for the mass m given the frequency w in Hz.

Exercise 14.E.2.4.4

Suppose we add possible friction to Exercise 14.E.2.4.3 Further, suppose you do not know the spring constant, but you have two reference weights 1 kg and 2 kg to calibrate your setup. You put each in motion on your spring and measure the frequency. For the 1 kg weight you measured 1.1 Hz, for the 2 kg weight you measured 0.8 Hz.

a. Find k (spring constant) and c (damping constant).
b. Find a formula for the mass in terms of the frequency in Hz. Note that there may be more than one possible mass for a given frequency.
c. For an unknown object you measured \(0.2\text{ Hz}\), what is the mass of the object? Suppose that you know that the mass of the unknown object is more than a kilogram.

Exercise 14.E.2.4.5

Suppose you wish to measure the friction a mass of 0.1 kg experiences as it slides along a floor (you wish to find c). You have a spring with spring constant $k = 5\frac{N}{m}$. You take the spring, you attach it to the mass and fix it to a wall. Then you pull on the spring and let the mass go. You find that the mass oscillates with frequency 1 Hz. What is the friction?

Exercise 14.E.2.4.6

A mass of 2 kilograms is on a spring with spring constant k newtons per meter with no damping. Suppose the system is at rest and at time $t = 0$ the mass is kicked and starts traveling at 2 meters per second. How large does k have to be to so that the mass does not go further than 3 meters from the rest position?

Answer

$k = \frac{8}{9}$ (and larger)

Exercise 14.E.2.4.7

Suppose we have an RLC circuit with a resistor of 100 miliohms (0.1 ohms), inductor of inductance of 50 millihenries (0.05 henries), and a capacitor of 5 farads, with constant voltage.

a. Set up the ODE equation for the current I.
b. Find the general solution.

c. Solve for $I(0)=10$ and $I'(0)=0$.

Answer

a. $0.05I''+0.1I'+\left(\frac{1}{5}\right)I=0$
b. $I=Ce^{-t}\cos(\sqrt{3}t-\gamma)$
c. $I=10e^{-t}\cos(\sqrt{3}t)+\frac{10}{\sqrt{3}}e^{-t}\sin(\sqrt{3}t)$

Exercise 14.E.2.4.8

A 5000 kg railcar hits a bumper (a spring) at $1\frac{m}{s}$, and the spring compresses by 0.1 m. Assume no damping.

a. Find k.
b. Find out how far does the spring compress when a 10000 kg railcar hits the spring at the same speed.
c. If the spring would break if it compresses further than 0.3 m, what is the maximum mass of a railcar that can hit it at $1\frac{m}{s}$?
d. What is the maximum mass of a railcar that can hit the spring without breaking at $2\frac{m}{s}$?

Answer

a. $k=500000$
b. $\frac{1}{5\sqrt{2}}\approx 0.141$
c. 45000 kg
d. 11250 kg

Exercise 14.E.2.4.9

A mass of m kg is on a spring with $k=3\frac{\text{N}}{\text{m}}$ and $c=2\frac{\text{Ns}}{\text{m}}$. Find the mass m_0 for which there is critical damping. If $m<m_0$, does the system oscillate or not, that is, is it underdamped or overdamped?

Answer

$m_0=\frac{1}{2}$. If $m<m_0$, then the system is overdamped and will not oscillate.

14.E.5: 2.5: Nonhomogeneous Equations

Exercise 14.E.2.5.1

Find a particular solution of $y''-y'-6y=e^{2x}$.

Exercise 14.E.2.5.2

Find a particular solution of $y''-4y'+4y=e^{2x}$.

Exercise 14.E.2.5.3

Solve the initial value problem $y''+9y=\cos(3x)+\sin(3x)$ for $y(0)=2, y'(0)=1$.

Exercise 14.E.2.5.4

Setup the form of the particular solution but do not solve for the coefficients for $y^{(4)}-2y'''+y''=e^x$.

Exercise 14.E.2.5.5

Setup the form of the particular solution but do not solve for the coefficients for $y^{(4)} - 2y''' + y'' = e^x + x + \sin x$.

Exercise 14.E.2.5.6

a. Using variation of parameters find a particular solution of $y'' - 2y' + y = e^x$.
b. Find a particular solution using undetermined coefficients.
c. Are the two solutions you found the same? What is going on? See also Exercise 14.E.2.5.9

Exercise 14.E.2.5.7

Find a particular solution of $y'' - 2y' + y = \sin(x^2)$. It is OK to leave the answer as a definite integral.

Exercise 14.E.2.5.8

For an arbitrary constant c find a particular solution to $y'' - y = e^{cx}$. Hint: Make sure to handle every possible real c.

Exercise 14.E.2.5.9

a. Using variation of parameters find a particular solution of $y'' - y = e^x$
b. Find a particular solution using undetermined coefficients.
c. Are the two solutions you found the same? What is going on?

Exercise 14.E.2.5.10

Find a polynomial $P(x)$, so that $y = 2x^2 + 3x + 4$ solves $y'' + 5y' + y = P(x)$.

Exercise 14.E.2.5.11

Find a particular solution to $y'' - y' + y = 2\sin(3x)$

Answer

$y = \frac{-16\sin(3x) + 6\cos(3x)}{73}$

Exercise 14.E.2.5.12

a. Find a particular solution to $y'' + 2y = e^x + x^3$.
b. Find the general solution.

Answer

a. $y = \frac{2e^x + 3x^3 - 9x}{6}$
b. $y = C_1\cos(\sqrt{2}x) + C_2\sin(\sqrt{2}x) + \frac{2e^x + 3x^3 - 9x}{6}$

Exercise 14.E.2.5.13

Solve $y'' + 2y' + y = x^2, y(0) = 1, y'(0) = 2$.

Answer

$y(x) = x^2 - 4x + 6 + e^{-x}(x - 5)$

Exercise 14.E.2.5.14

Use variation of parameters to find a particular solution of $y'' - y = \dfrac{1}{e^x + e^{-x}}$.

Answer

$y = \frac{2xe^x - (e^x + e^{-x})\log(e^{2x}+1)}{4}$

Exercise 14.E.2.5.15

For an arbitrary constant c find the general solution to $y'' - 2y = \sin(x + c)$.

Answer

$y = \frac{-\sin(x+c)}{3} + C_1 e^{\sqrt{2}x} + C_2 e^{-\sqrt{2}x}$

Exercise 14.E.2.5.16

Undetermined coefficients can sometimes be used to guess a particular solution to other equations than constant coefficients. Find a polynomial $y(x)$ that solves $y' + xy = x^3 + 2x^2 + 5x + 2$.

Note: Not every right hand side will allow a polynomial solution, for example, $y' + xy = 1$ does not, but a technique based on undetermined coefficients does work, see Chapter 7.

Answer

$y = x^2 + 2x + 3$

14.E.6: 2.6: Forced Oscillations and Resonance

Exercise 14.E.2.6.1

Derive a formula for x_{sp} if the equation is $mx'' + cx' + kx = F_0 \sin(\omega t)$. Assume $c > 0$.

Exercise 14.E.2.6.2

Derive a formula for x_{sp} if the equation is $mx'' + cx' + kx = F_0 \cos(\omega t) + F_1 \cos(3\omega t)$. Assume $c > 0$.

Exercise 14.E.2.6.3

Take $mx'' + cx' + kx = F_0 \cos(\omega t)$. Fix $m > 0$ and $k > 0$. Now think of the function $C(\omega)$. For what values of c (solve in terms of m, k, and F_0) will there be no practical resonance (that is, for what values of c is there no maximum of $C(\omega)$ for $\omega > 0$)?

Exercise 14.E.2.6.4

Take $mx'' + cx' + kx = F_0 \cos(\omega t)$. Fix $c > 0$ and $k > 0$. Now think of the function $C(\omega)$. For what values of m (solve in terms of c, k, and F_0) will there be no practical resonance (that is, for what values of m is there no maximum of $C(\omega)$ for $\omega > 0$)?

Exercise 14.E.2.6.5

Suppose a water tower in an earthquake acts as a mass-spring system. Assume that the container on top is full and the water does not move around. The container then acts as a mass and the support acts as the spring, where the induced vibrations are

horizontal. Suppose that the container with water has a mass of $m = 10,000kg$. It takes a force of 1000 newtons to displace the container 1 meter. For simplicity assume no friction. When the earthquake hits the water tower is at rest (it is not moving).

Suppose that an earthquake induces an external force $F(t) = mA\omega^2 \cos(\omega t)$.

a. What is the natural frequency of the water tower?
b. If ω is not the natural frequency, find a formula for the maximal amplitude of the resulting oscillations of the water container (the maximal deviation from the rest position). The motion will be a high frequency wave modulated by a low frequency wave, so simply find the constant in front of the sines.
c. Suppose $A = 1$ and an earthquake with frequency 0.5 cycles per second comes. What is the amplitude of the oscillations? Suppose that if the water tower moves more than 1.5 meter, the tower collapses. Will the tower collapse?

? Exercise 14.E. 2.6.6

A mass of 4 kg on a spring with $k = 4$ and a damping constant $c = 1$. Suppose that $F_0 = 2$. Using forcing of $F_0 \cos(\omega t)$. Find the ω that causes practical resonance and find the amplitude.

Answer

$\omega = \frac{\sqrt{31}}{4\sqrt{2}} \approx 0.984 \quad C(\omega) = \frac{16}{3\sqrt{7}} \approx 2.016$

? Exercise 14.E. 2.6.7

Derive a formula for x_{sp} for $mx'' + cx' + kx = F_0 \cos(\omega t) + A$ where A is some constant. Assume $c > 0$.

Answer

$x_{sp} = \frac{(\omega_0^2 - \omega^2)F_0}{m(2\omega p)^2 + m(\omega_0^2 - \omega^2)^2}\cos(\omega t) + \frac{2\omega p F_0}{m(2\omega p)^2 + m(\omega_0^2 - \omega^2)^2}\sin(\omega t) + \frac{A}{k}$, where $p = \frac{c}{2m}$ and $\omega_0 = \sqrt{\frac{k}{m}}$

? Exercise 14.E. 2.6.8

Suppose there is no damping in a mass and spring system with $m = 5, k = 20,$ and $F_0 = 5$. Suppose that ω is chosen to be precisely the resonance frequency.

a. Find ω.
b. Find the amplitude of the oscillations at time $t = 100$.

Answer

a. $\omega = 2$
b. 25

This page titled 14.E: Higher order linear ODEs (Exercises) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **2.E: Higher order linear ODEs (Exercises)** has no license indicated.

CHAPTER OVERVIEW

15: Fourier series and PDEs

15.1: Boundary value problems
15.2: The Trigonometric Series
15.3: More on the Fourier Series
15.4: Sine and Cosine Series
15.5: Applications of Fourier Series
15.6: PDEs, Separation of Variables, and The Heat Equation
15.7: One Dimensional Wave Equation
15.8: D'Alembert Solution of The Wave Equation
15.9: Steady State Temperature and the Laplacian
15.10: Dirichlet Problem in the Circle and the Poisson Kernel
15.E: Fourier Series and PDEs (Exercises)

Contributors and Attributions

- Jiří Lebl (Oklahoma State University).

These pages were supported by NSF grants DMS-0900885 and DMS-1362337.

This page titled 15: Fourier series and PDEs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

15.1: Boundary value problems

Before we tackle the Fourier series, we need to study the so-called **boundary value problems** (or endpoint problems). For example, suppose we have

$$x'' + \lambda x = 0, \quad x(a) = 0, \quad x(b) = 0,$$

for some constant λ, where $x(t)$ is defined for t in the interval $[a, b]$. Unlike before, when we specified the value of the solution and its derivative at a single point, we now specify the value of the solution at two different points. Note that $x = 0$ is a solution to this equation, so existence of solutions is not an issue here. Uniqueness of solutions is another issue. The general solution to $x'' + \lambda x = 0$ has two arbitrary constants present.[1] It is, therefore, natural (but wrong) to believe that requiring two conditions guarantees a unique solution.

Example 15.1.1

Take $\lambda = 1, a = 0, b = \pi$. That is,

$$x'' + x = 0, \quad x(0) = 0, \quad x(\pi) = 0.$$

Then $x = \sin t$ is another solution (besides $x = 0$) satisfying both boundary conditions. There are more. Write down the general solution of the differential equation, which is $x = A\cos t + B\sin t$. The condition $x(0) = 0$ forces $A = 0$. Letting $x(\pi) = 0$ does not give us any more information as $x = B\sin t$ already satisfies both boundary conditions. Hence, there are infinitely many solutions of the form $x = B\sin t$, where B is an arbitrary constant.

Example 15.1.2

On the other hand, change to $\lambda = 2$.

$$x'' + 2x = 0, \quad x(0) = 0, \quad x(\pi) = 0.$$

Then the general solution is $x = A\cos(\sqrt{2}t) + B\sin(\sqrt{2}t)$. Letting $x(0) = 0$ still forces $A = 0$. We apply the second condition to find $0 = x(\pi) = B\sin(\sqrt{2}t)$. As $\sin(\sqrt{2}t) \neq 0$ we obtain $B = 0$. Therefore $x = 0$ is the unique solution to this problem.

What is going on? We will be interested in finding which constants λ allow a nonzero solution, and we will be interested in finding those solutions. This problem is an analogue of finding eigenvalues and eigenvectors of matrices.

15.1.1: Eigenvalue Problems

For basic Fourier series theory we will need the following three eigenvalue problems. We will consider more general equations, but we will postpone this until Chapter 5.

$$x'' + \lambda x = 0, \qquad x(a) = 0, \quad x(b) = 0, \tag{15.1.1}$$
$$x'' + \lambda x = 0, \qquad x'(a) = 0, \quad x'(b) = 0, \tag{15.1.2}$$

and

$$x'' + \lambda x = 0, \quad x(a) = x(b), \quad x'(a) = x'(b), \tag{15.1.3}$$

A number λ is called an eigenvalue of (15.1.1) (resp. (15.1.2) or (15.1.3)) if and only if there exists a nonzero (not identically zero) solution to (15.1.1) (resp. (15.1.2) or (15.1.3)) given that specific λ. The nonzero solution we found is called the corresponding *eigenfunction*.

Note the similarity to eigenvalues and eigenvectors of matrices. The similarity is not just coincidental. If we think of the equations as differential operators, then we are doing the same exact thing. Think of a function $x(t)$ as a vector with infinitely many components (one for each t). Let $L = -\frac{d^2}{dt^2}$ be the linear operator. Then the eigenvalue/eigenfunction pair should be λ and nonzero x such that $Lx = \lambda x$. In other words, we are looking for nonzero functions x satisfying certain endpoint conditions that solve $(L - \lambda)x = 0$. A lot of the formalism from linear algebra still applies here, though we will not pursue this line of reasoning too far.

Example 15.1.3

Let us find the eigenvalues and eigenfunctions of

$$x'' + \lambda x = 0, \quad x(0) = 0, \quad x(\pi) = 0.$$

For reasons that will be clear from the computations, we will have to handle the cases $\lambda > 0, \lambda = 0, \lambda < 0$ separately. First suppose that $\lambda > 0$, then the general solution to $x'' + \lambda x = 0$ is

$$x = A\cos(\sqrt{\lambda}t) + B\sin(\sqrt{\lambda}t).$$

The condition $x(0) = 0$ implies immediately A. Next

$$0 = x(\pi) = B\sin(\sqrt{\lambda}\pi).$$

If B is zero, then x is not a nonzero solution. So to get a nonzero solution we must have that $\sin(\sqrt{\lambda}\pi) = 0$. Hence, $\sqrt{\lambda}\pi$ must be an integer multiple of π. In other words, $\sqrt{\lambda} = k$ for a positive integer k. Hence the positive eigenvalues are k^2 for all integers $k \geq 1$. The corresponding eigenfunctions can be taken as $x = \sin(kt)$. Just like for eigenvectors, we get all the multiples of an eigenfunction, so we only need to pick one.

Now suppose that $\lambda = 0$. In this case the equation is $x'' = 0$ and the general solution is $x = At + B$. The condition $x(0) = 0$ implies that $B = 0$, and $x(\pi) = 0$ implies that $A = 0$. This means that $\lambda = 0$ is not an eigenvalue.

Finally, suppose that $\lambda < 0$. In this case we have the general solution[2]

$$x = A\cosh(\sqrt{-\lambda}t) + B\sinh(\sqrt{-\lambda}t).$$

Letting $x(0) = 0$ implies that $A = 0$ (recall $\cosh 0 = 1$ and $\sinh 0 = 0$). So our solution must be $x = B\sinh(\sqrt{-\lambda}t)$ and satisfy $x(\pi) = 0$. This is only possible if B is zero. Why? Because $\sinh\xi$ is only zero when $\xi = 0$. You should plot sinh to see this fact. We can also see this from the definition of sinh. We get $0 = \sinh t = \frac{e^t - e^{-t}}{2}$. Hence $e^t = e^{-t}$, which implies $t = -t$ and that is only true if $t = 0$. So there are no negative eigenvalues.

In summary, the eigenvalues and corresponding eigenfunctions are

$$\lambda_k = k^2 \quad \text{with an eigenfucntion} \quad x_k = \sin(kt) \quad \text{for all integers } k \geq 1.$$

Example 15.1.4

Let us compute the eigenvalues and eigenfunctions of

$$x'' + \lambda x = 0, \quad x'(0) = 0, \quad x'(\pi) = 0.$$

Again we will have to handle the cases $\lambda > 0, \lambda = 0, \lambda < 0$ separately. First suppose that $\lambda > 0$. The general solution to $x'' + \lambda x = 0$ is $x = A\cos(\sqrt{\lambda}t) + B\sin(\sqrt{\lambda}t)$. So

$$x' = -A\sqrt{\lambda}\sin(\sqrt{\lambda}t) + B\sqrt{\lambda}\cos(\sqrt{\lambda}t).$$

The condition $x'(0) = 0$ implies immediately $B = 0$. Next

$$0 = x'(\pi) = -A\sqrt{\lambda}\sin(\sqrt{\lambda}\pi).$$

Again A cannot be zero if λ is to be an eigenvalue, and $\sin(\sqrt{\lambda}\pi)$ is only zero if $\sqrt{\lambda} = k$ for a positive integer k. Hence the positive eigenvalues are again k^2 for all integers $k \geq 1$. And the corresponding eigenfunctions can be taken as $x = \cos(kt)$.

Now suppose that $\lambda = 0$. In this case the equation is $x'' = 0$ and the general solution is $x = At + B$ so $x' = A$. The condition $x'(0) = 0$ implies that $A = 0$. Now $x'(\pi) = 0$ also simply implies $A = 0$. This means that B could be anything (let us take it to be 1). So $\lambda = 0$ is an eigenvalue and $x = 1$ is a corresponding eigenfunction.

Finally, let $\lambda < 0$. In this case we have the general solution $x = A\cosh(\sqrt{-\lambda}t) + B\sinh(\sqrt{-\lambda}t)$ and hence

$$x' = A\sqrt{-\lambda}\sinh(\sqrt{-\lambda}t) + B\sqrt{-\lambda}\cosh(\sqrt{-\lambda}t).$$

We have already seen (with roles of A and B switched) that for this to be zero at $t = 0$ and $t = \pi$ it implies that $A = B = 0$. Hence there are no negative eigenvalues.

In summary, the eigenvalues and corresponding eigenfunctions are

$$\lambda_k = k^2 \quad \text{with an eigenfunction} \quad x_k = \cos(kt) \quad \text{for all integers } k \geq 1,$$

and there is another eigenvalue

$$\lambda_0 = 0 \quad \text{with an eigenfunction} \quad x_0 = 1.$$

The following problem is the one that leads to the general Fourier series.

Example 15.1.5

Let us compute the eigenvalues and eigenfunctions of

$$x'' + \lambda x = 0, \quad x(-\pi) = x(\pi), \quad x'(-\pi) = x'(\pi).$$

Notice that we have not specified the values or the derivatives at the endpoints, but rather that they are the same at the beginning and at the end of the interval.

Let us skip $\lambda < 0$. The computations are the same as before, and again we find that there are no negative eigenvalues.

For $\lambda = 0$, the general solution is $x = At + B$. The condition $x(-\pi) = x(\pi)$ implies that $A = 0$ ($A\pi + B = -A\pi + B$ implies $A = 0$). The second condition $x'(-\pi) = x'(\pi)$ says nothing about B and hence $\lambda = 0$ is an eigenvalue with a corresponding eigenfunction $x = 1$.

For $\lambda > 0$ we get that $x = A\cos(\sqrt{\lambda}t) + B\sin(\sqrt{\lambda}t)$. Now

$$\underbrace{A\cos(-\sqrt{\lambda}\,\pi) + B\sin(-\sqrt{\lambda}\,\pi)}_{x(-\pi)} = \underbrace{A\cos(\sqrt{\lambda}\,\pi) + B\sin(\sqrt{\lambda}\,\pi)}_{x(\pi)}.$$

We remember that $\cos(-\theta) = \cos(\theta)$ and $\sin(-\theta) = -\sin(\theta)$. Therefore,

$$A\cos(\sqrt{\lambda}\pi) - B\sin(\sqrt{\lambda}\pi) = A\cos(\sqrt{\lambda}\pi) + B\sin(\sqrt{\lambda}\pi).$$

Hence either $B = 0$ or $\sin(\sqrt{\lambda}\pi) = 0$. Similarly (exercise) if we differentiate x and plug in the second condition we find that $A = 0$ or $\sin(\sqrt{\lambda}\pi) = 0$. Therefore, unless we want A and B to both be zero (which we do not) we must have $\sin(\sqrt{\lambda}\pi) = 0$. Hence, $\sqrt{\lambda}$ is an integer and the eigenvalues are yet again $\lambda = k^2$ for an integer $k \geq 1$. In this case, however, $x = A\cos(kt) + B\sin(kt)$ is an eigenfunction for any A and any B. So we have two linearly independent eigenfunctions $\sin(kt)$ and $\cos(kt)$. Remember that for a matrix we could also have had two eigenvectors corresponding to a single eigenvalue if the eigenvalue was repeated.

In summary, the eigenvalues and corresponding eigenfunctions are

$$\begin{array}{lll} \lambda_k = k^2 & \text{with eigenfunctions} & \cos(kt) \quad \text{and} \quad \sin(kt) \quad \text{for all integers } k \geq 1, \\ \lambda_0 = 0 & \text{with an eigenfunction} & x_0 = 1. \end{array} \tag{15.1.4}$$

15.1.2: Orthogonality of Eigenfunctions

Something that will be very useful in the next section is the *orthogonality* property of the eigenfunctions. This is an analogue of the following fact about eigenvectors of a matrix. A matrix is called *symmetric* if $A = A^T$. *Eigenvectors for two distinct eigenvalues of a symmetric matrix are orthogonal*. That symmetry is required. We will not prove this fact here. The differential operators we are dealing with act much like a symmetric matrix. We, therefore, get the following theorem.

Theorem 15.1.1: Orthogonal

Suppose that $x_1(t)$ and $x_2(t)$ are two eigenfunctions of the problem (15.1.1), (15.1.2) or (15.1.3) for two different eigenvalues λ_1 and λ_2. Then they are **orthogonal** in the sense that

$$\int_a^b x_1(t)x_2(t)dt = 0.$$

Note that the terminology comes from the fact that the integral is a type of inner product. We will expand on this in the next section. The theorem has a very short, elegant, and illuminating proof so let us give it here. First note that we have the following two equations.

$$x_1'' + \lambda_1 x_1 = 0 \quad \text{and} \quad x_2'' + \lambda_2 x_2 = 0.$$

Multiply the first by x_2 and the second by x_1 and subtract to get

$$(\lambda_1 - \lambda_2)x_1 x_2 = x_2'' x_1 - x_2 x_1''.$$

Now integrate both sides of the equation.

$$\begin{aligned}(\lambda_1 - \lambda_2)\int_a^b x_1 x_2 dt &= \int_a^b x_2'' x_1 - x_2 x_1'' dt \\ &= \int_a^b \frac{d}{dt}(x_2' x_1 - x_2 x_1')dt \\ &= [x_2' x_1 - x_2 x_1']_{t=a}^b = 0.\end{aligned} \tag{15.1.5}$$

The last equality holds because of the boundary conditions. For example, if we consider (15.1.1) we have $x_1(a) = x_1(b) = x_2(a) = x_2(b) = 0$ and so $x_2' x_1 - x_2 x_1'$ is zero at both a and b. As $\lambda_1 \neq \lambda_2$, the theorem follows.

? Exercise 15.1.1: (easy)

Finish the theorem (check the last equality in the proof) for the cases (15.1.2) and (15.1.3).

We have seen previously that $\sin(nt)$ was an eigenfunction for the problem $x'' + \lambda x = 0,\ x(0) = 0,\ x(\pi) = 0$. Hence we have the integral

$$\int_0^\pi \sin(mt)\sin(nt)\,dt = 0, \quad \text{when } m \neq n.$$

Similarly,

$$\int_0^\pi \cos(mt)\cos(nt)\,dt = 0, \quad \text{when } m \neq n, \qquad \text{and} \qquad \int_0^\pi \cos(nt)\,dt = 0.$$

And finally we also get

$$\int_{-\pi}^\pi \sin(mt)\sin(nt)\,dt = 0, \quad \text{when } m \neq n, \qquad \text{and} \qquad \int_{-\pi}^\pi \sin(nt)\,dt = 0,$$

$$\int_{-\pi}^\pi \cos(mt)\cos(nt)\,dt = 0, \quad \text{when } m \neq n, \qquad \text{and} \qquad \int_{-\pi}^\pi \cos(nt)\,dt = 0,$$

and

$$\int_{-\pi}^\pi \cos(mt)\sin(nt)\,dt = 0 \qquad (\text{even if } m = n).$$

15.1.3: Fredholm Alternative

We now touch on a very useful theorem in the theory of differential equations. The theorem holds in a more general setting than we are going to state it, but for our purposes the following statement is sufficient. We will give a slightly more general version in Chapter 5.

Theorem 15.1.2

Fredholm alternative[3]

Exactly one of the following statements holds. Either

$$x'' + \lambda x = 0, \quad x(a) = 0, \quad x(b) = 0$$

has a nonzero solution, or

$$x'' + \lambda x = f(t), \quad x(a) = 0, \quad x(b) = 0 \tag{15.1.6}$$

has a unique solution for every function f continuous on $[a, b]$.

The theorem is also true for the other types of boundary conditions we considered. The theorem means that if λ is not an eigenvalue, the nonhomogeneous equation (15.1.6) has a unique solution for every right hand side. On the other hand if λ is an eigenvalue, then (15.1.6) need not have a solution for every f, and furthermore, even if it happens to have a solution, the solution is not unique.

We also want to reinforce the idea here that linear differential operators have much in common with matrices. So it is no surprise that there is a finite dimensional version of Fredholm alternative for matrices as well. Let A be an $n \times n$ matrix. The Fredholm alternative then states that either $(A - \lambda I)\vec{x} = \vec{0}$ has a nontrivial solution, or $(A - \lambda I)\vec{x} = \vec{b}$ has a solution for every $\vec{b}$.

A lot of intuition from linear algebra can be applied to linear differential operators, but one must be careful of course. For example, one difference we have already seen is that in general a differential operator will have infinitely many eigenvalues, while a matrix has only finitely many.

15.1.3.1: Application

Let us consider a physical application of an endpoint problem. Suppose we have a tightly stretched quickly spinning elastic string or rope of uniform linear density ρ. Let us put this problem into the xy–plane. The x axis represents the position on the string. The string rotates at angular velocity ω, so we will assume that the whole xy–plane rotates at angular velocity ω. We will assume that the string stays in this xy–plane and y will measure its deflection from the equilibrium position, $y = 0$, on the x axis. Hence, we will find a graph giving the shape of the string. We will idealize the string to have no volume to just be a mathematical curve. If we take a small segment and we look at the tension at the endpoints, we see that this force is tangential and we will assume that the magnitude is the same at both end points. Hence the magnitude is constant everywhere and we will call its magnitude T. If we assume that the deflection is small, then we can use Newton's second law to get an equation

$$Ty'' + \rho\omega^2 y = 0.$$

Let L be the length of the string and the string is fixed at the beginning and end points. Hence, $y(0) = 0$ and $y(L) = 0$. See Figure 15.1.1.

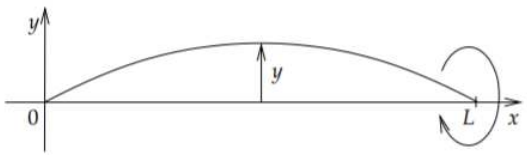


Figure 15.1.1: Whirling string.

We rewrite the equation as $y'' + \frac{\rho\omega^2}{T} y = 0$. The setup is similar to Example 15.1.3, except for the interval length being L instead of π. We are looking for eigenvalues of $y'' + \lambda y = 0, y(0) = 0, y(L) = 0$ where $\lambda = \frac{\rho\omega^2}{T}$. As before there are no nonpositive eigenvalues. With $\lambda > 0$, the general solution to the equation is $y = A\cos(\sqrt{\lambda}x) + B\sin(\sqrt{\lambda}x)$. The condition $y(0) = 0$ implies that $A = 0$ as before. The condition $y(L) = 0$ implies that $\sin(\sqrt{\lambda}L) = 0$ and hence $\sqrt{\lambda}L = k\pi$ for some integer $k > 0$, so

$$\frac{\rho\omega^2}{T} = \lambda = \frac{k^2\pi^2}{L^2}.$$

What does this say about the shape of the string? It says that for all parameters ρ, ω, T not satisfying the equation above, the string is in the equilibrium position, $y = 0$. When $\frac{\rho\omega^2}{T} = \frac{k^2\pi^2}{L^2}$, then the string will some distance B. We cannot compute B with the

information we have.

Let us assume that ρ and T are fixed and we are changing ω. For most values of ω the string is in the equilibrium state. When the angular velocity ω hits a value $\omega = \frac{k\pi\sqrt{T}}{L\sqrt{\rho}}$, then the string pops out and has the shape of a sin wave crossing the x-axis $k-1$ times between the end points. For example, at $k=1$, the string does not cross the x-axis and the shape looks like in Figure 15.1.1. On the other hand, when $k=3$ the string crosses the x-axis 2 times, see Figure 15.1.2 When ω changes again, the string returns to the equilibrium position. The higher the angular velocity, the more times it crosses the x-axis when it is popped out.

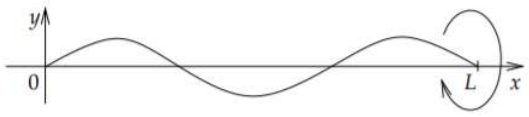


Figure 15.1.2: Whirling string at the third eigenvalue $(k=3)$.

For another example, if you have a spinning jump rope (then $k=1$ as it is completely) and you pull on the ends to increase the tension, then the velocity also increases for the rope to stay "popped out".

15.1.4: Footnotes

[1] See Example 2.2.1 and Example 2.2.3.

[2] Recall that $\cosh s = \frac{1}{2}(e^s + e^{-s})$ and $\sinh s = \frac{1}{2}(e^s - e^{-s})$. As an exercise try the computation with the general solution written as $x = Ae^{\sqrt{-\lambda}t} + Be^{-\sqrt{-\lambda}t}$ (for different A and B of course).

[3] Named after the Swedish mathematician Erik Ivar Fredholm (1866–1927).

This page titled 15.1: Boundary value problems is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.1: Boundary value problems** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.2: The Trigonometric Series

15.2.1: Periodic Functions and Motivation

As motivation for studying Fourier series, suppose we have the problem

$$x'' + \omega_0^2 x = f(t), \tag{15.2.1}$$

for some periodic function $f(t)$. We have already solved

$$x'' + \omega_0^2 x = F_0 \cos(\omega t). \tag{15.2.2}$$

One way to solve (15.2.1) is to decompose $f(t)$ as a sum of cosines (and sines) and then solve many problems of the form (15.2.2). We then use the principle of superposition, to sum up all the solutions we got to get a solution to (15.2.1).

Before we proceed, let us talk a little bit more in detail about periodic functions. A function is said to be *periodic* with period P if $f(t)$ for all t. For brevity we will say $f(t)$ is $P-$periodic. Note that a $P-$periodic function is also $2P-$periodic, $3P-$periodic and so on. For example, $\cos(t)$ and $\sin(t)$ are $2\pi-$periodic. So are $\cos(kt)$ and $\sin(kt)$ for all integers k. The constant functions are an extreme example. They are periodic for any period (exercise).

Normally we will start with a function $f(t)$ defined on some interval $[-L, L]$ and we will want to *extend* $f(t)$ *periodically* to make it a $2L-$periodic function. We do this extension by defining a new function $F(t)$ such that for t in$[-L, L]$, $F(t) = f(t)$. For t in $[L, 3L]$, we define $F(t) = f(t - 2L)$, for t in $[-3L, -L]$, $F(t) = f(t + 2L)$, and so on. We assumed that $f(-L) = f(L)$. We could have also started with f defined only on the half-open interval $(-L, L]$ and then define $f(-L) = f(L)$.

Example 15.2.1

Define $f(t) = 1 - t^2$ on $[-1, 1]$. Now extend $f(t)$ periodically to a 2-periodic function. See Figure 15.2.1on the facing page.

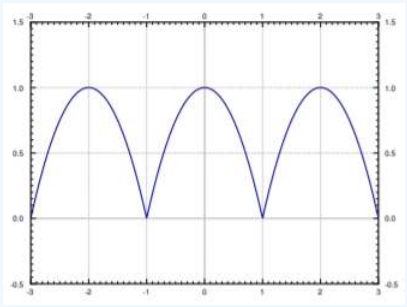


Figure 15.2.1: Periodic extension of the function $1 - t^2$.

You should be careful to distinguish between $f(t)$ and its extension. A common mistake is to assume that a formula for $f(t)$ holds for its extension. It can be confusing when the formula for $f(t)$ is periodic, but with perhaps a different period.

Exercise 15.2.1

Define $f(t) = \cos t$ on $[\frac{-\pi}{2}, \frac{\pi}{2}]$. Take the $\pi-$periodic extension and sketch its graph. How does it compare to the graph of $\cos t$?

15.2.2: Inner Product and Eigenvector Decomposition

Suppose we have a *symmetric matrix*, that is $A^T = A$. We have said before that the eigenvectors of A are then orthogonal. Here the word *orthogonal* means that if $\vec{v}$ and $\vec{w}$ are two distinct (and not multiples of each other) eigenvectors of A, then $\langle \vec{v}, \vec{w} \rangle = 0$. In this case the inner product $\langle \vec{v}, \vec{w} \rangle$ is the dot product, which can be computed as $\vec{v}^T \vec{w}$.

To decompose a vector $\vec{v}$ in terms of mutually orthogonal vectors $\vec{w}_1$ and $\vec{w}_2$ we write

$$\vec{v} = a_1 \vec{w}_1 + a_2 \vec{w}_2.$$

Let us find the formula for a_1 and a_2. First let us compute

$$\langle \vec{v}, \vec{w_1} \rangle = \langle a_1\vec{w}_1 + a_2\vec{w}_2, \vec{w_1} \rangle = a_1\langle \vec{w}_1, \vec{w_1} \rangle + a_2 \underbrace{\langle \vec{w}_2, \vec{w_1} \rangle}_{=0} = a_1\langle \vec{w}_1, \vec{w_1} \rangle.$$

Therefore,

$$a_1 = \frac{\langle \vec{v}, \vec{w}_1 \rangle}{\langle \vec{w}_1, \vec{w}_1 \rangle}.$$

Similarly

$$a_2 = \frac{\langle \vec{v}, \vec{w}_2 \rangle}{\langle \vec{w}_2, \vec{w}_2 \rangle}.$$

You probably remember this formula from vector calculus.

Example 15.2.2

Write $\vec{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ as a linear combination of $\vec{w}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\vec{w}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

First note that $\vec{w}_1$ and $\vec{w}_2$ are orthogonal as $\langle \vec{w}_1, \vec{w}_2 \rangle = 1(1) + (-1)1 = 0$. Then

$$\begin{aligned} a_1 &= \frac{\langle \vec{v}, \vec{w}_1 \rangle}{\langle \vec{w}_1, \vec{w}_1 \rangle} = \frac{2(1)+3(-1)}{1(1)+(-1)(-1)} = \frac{-1}{2}, \\ a_2 &= \frac{\langle \vec{v}, \vec{w}_2 \rangle}{\langle \vec{w}_2, \vec{w}_2 \rangle} = \frac{2+3}{1+1} = \frac{5}{2}. \end{aligned} \tag{15.2.3}$$

Hence

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \frac{-1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} + \frac{5}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

15.2.3: Trigonometric Series

Instead of decomposing a vector in terms of eigenvectors of a matrix, we will decompose a function in terms of eigenfunctions of a certain eigenvalue problem. The eigenvalue problem we will use for the Fourier series is

$$x'' + \lambda x = 0, \quad x(-\pi) = x(\pi) \quad x'(-\pi) = x'(\pi).$$

We have previously computed that the eigenfunctions are $1, \cos(kt), \sin(kt)$. That is, we will want to find a representation of a $2\pi-$periodic function $f(t)$ as

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + b_n \sin(nt).$$

This series is called the *Fourier series*[1] or the *trigonometric series* for $f(t)$. We write the coefficient of the eigenfunction 1 as $\frac{a_0}{2}$ for convenience. We could also think of $1 = \cos(0t)$, so that we only need to look at $\cos(kt)$ and $\sin(kt)$.

As for matrices we want to find a *projection* of $f(t)$ onto the subspace generated by the eigenfunctions. So we will want to define an *inner product of functions*. For example, to find a_n we want to compute $\langle f(t), \cos(nt) \rangle$. We define the inner product as

$$\langle f(t), g(t) \rangle = \int_{-\pi}^{\pi} f(t)g(t)dt.$$

With this definition of the inner product, we have seen in the previous section that the eigenfunctions $\cos(kt)$(including the constant eigenfunction), and $\sin(kt)$ are *orthogonal* in the sense that

$$\begin{aligned} &\langle \cos(mt), \cos(nt) \rangle = 0 && \text{for } m \neq n, \\ &\langle \sin(mt), \sin(nt) \rangle = 0 && \text{for } m \neq n, \\ &\langle \sin(mt), \cos(nt) \rangle = 0 && \text{for all } m \text{ and } n. \end{aligned} \tag{15.2.4}$$

By elementary calculus for $n = 1, 2, 3, \ldots$. we have $\langle \cos(nt), \cos(nt) \rangle = \pi$ and $\langle \sin(nt), \sin(nt) \rangle = \pi$. For the constant we get

$$\langle 1, 1 \rangle = \int_{\pi}^{\pi} 1 \cdot 1 \, dt = 2\pi.$$

The coefficients are given by

$$\begin{aligned} a_n &= \frac{\langle f(t), \cos(nt) \rangle}{\langle \cos(nt), \cos(nt) \rangle} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(nt) dt, \\ b_n &= \frac{\langle f(t), \sin(nt) \rangle}{\langle \sin(nt), \sin(nt) \rangle} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(nt) dt. \end{aligned} \tag{15.2.5}$$

Compare these expressions with the finite-dimensional example. For a_0 we get a similar formula

$$a_0 = 2 \frac{\langle f(t), 1 \rangle}{\langle 1, 1 \rangle} \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) dt.$$

Let us check the formulas using the orthogonality properties. Suppose for a moment that

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + b_n \sin(nt).$$

Then for $m \geq 1$ we have

$$\begin{aligned} \langle f(t), \cos(mt) \rangle &= \left\langle \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + b_n \sin(nt), \cos(mt) \right\rangle \\ &= \frac{a_0}{2} \langle 1, \cos(mt) \rangle + \sum_{n=1}^{\infty} a_n \langle \cos(nt), \cos(mt) \rangle + b_n \langle \sin(nt), \sin(mt) \rangle \\ &= a_m \langle \cos(mt), \cos(mt) \rangle . \end{aligned} \tag{15.2.6}$$

And hence $a_m = \frac{\langle f(t), \cos(mt) \rangle}{\langle \cos(mt), \cos(mt) \rangle}$.

? Exercise 15.2.2

Carry out the calculation for a_0 and b_m.

✓ Example 15.2.3

Take the function

$$f(t) = t$$

for t in $(-\pi, \pi]$. Extend $f(t)$ periodically and write it as a Fourier series. This function is called the *sawtooth*.

The plot of the extended periodic function is given in Figure 15.2.2 Let us compute the coefficients.

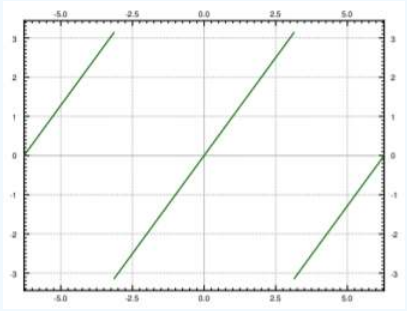


Figure 15.2.2: The graph of the sawtooth function.

Solution

We start with a_0,

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} t dt = 0.$$

We will often use the result from calculus that says that the integral of an odd function over a symmetric interval is zero. Recall that an odd function is a function $\varphi(t)$ such that $\varphi(-t) = -\varphi(t)$. For example the functions $t, \sin t$, or (importantly for us) $t\cos(nt)$ are all odd functions. Thus

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} t\cos(nt) dt = 0.$$

Let us move to b_n. Another useful fact from calculus is that the integral of an even function over a symmetric interval is twice the integral of the same function over half the interval. Recall an *even function* is a function $\varphi(t)$ such that $\varphi(-t) = \varphi(t)$. For example $t\sin(nt)$ is even.

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} t\sin(nt) dt \\ &= \frac{2}{\pi}\int_{0}^{\pi} t\sin(nt) dt \\ &= \frac{2}{\pi}\left(\left[\frac{-t\cos(nt)}{n} \right]_{t=0}^{\pi} + \frac{1}{n}\int_0^{\pi} \cos(nt) dt \right) \\ &= \frac{2}{\pi}\left(\frac{-\pi\cos(n\pi)}{n} + 0 \right) \\ &= \frac{-2\cos(n\pi)}{n} = \frac{2(-1)^{n+1}}{n}. \end{aligned} \tag{15.2.7}$$

We have used the fact that

$$\cos(n\pi) = (-1)^n = \begin{cases} 1 & \text{if } n \text{ even}, \\ -1 & \text{if } n \text{ odd}. \end{cases}$$

The series, therefore, is

$$\sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n}\sin(nt).$$

Let us write out the first 3 harmonics of the series for $f(t)$.

$$2\sin(t) - \sin(2t) + \frac{2}{3}\sin(3t) + \cdots$$

The plot of these first three terms of the series, along with a plot of the first 20 terms is given in Figure 15.2.3

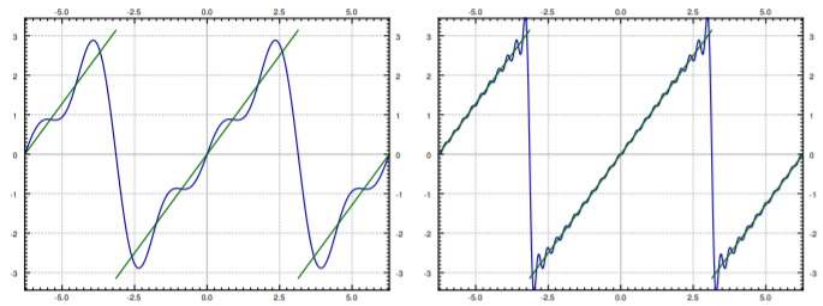


Figure 15.2.3: First 3 (left graph) and 20 (right graph) harmonics of the sawtooth function.

Example 15.2.4

Take the function

$$f(t) = \begin{cases} 0 & \text{if } -\pi < t \le 0, \\ \pi & \text{if } 0 < t \le \pi. \end{cases}$$

Extend $f(t)$ periodically and write it as a Fourier series. This function or its variants appear often in applications and the function is called the square wave.

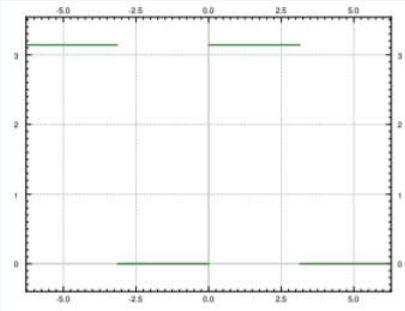


Figure 15.2.4: The graph of the square wave function.

The plot of the extended periodic function is given in Figure 15.2.4 Now we compute the coefficients. Let us start with a_0

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)dt = \frac{1}{\pi}\int_{0}^{\pi} \pi dt = \pi.$$

Next,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos(nt)dt = \frac{1}{\pi}\int_{0}^{\pi} \pi\cos(nt)dt = 0.$$

And finally

$$\begin{aligned} b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(nt)dt \\ &= \frac{1}{\pi}\int_{0}^{\pi} \pi\sin(nt)dt \\ &= \left[\frac{-\cos(nt)}{n}\right]_{t=0}^{\pi} \\ &= \frac{1-\cos(\pi n)}{n} = \frac{1-(-1)^n}{n} = \begin{cases} \frac{2}{n} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases} \end{aligned} \tag{15.2.8}$$

The Fourier series is

$$\frac{\pi}{2} + \sum_{n=1\ n\text{ odd}}^{\infty} \frac{2}{n}\sin(nt) + \sum_{k=1}^{\infty} \frac{2}{2k-1}\sin((2k-1)t).$$

Let us write out the first 3 harmonics of the series for $f(t)$.

$$\frac{\pi}{2} + 2\sin(t) + \frac{2}{3}\sin(3t) + \cdots$$

The plot of these first three and also of the first 20 terms of the series is given in Figure 15.2.5

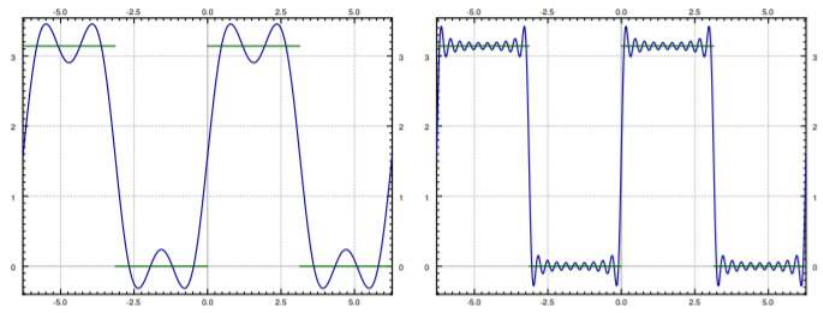


Figure 15.2.5: First 3 (left graph) and 20 (right graph) harmonics of the square wave function.

We have so far skirted the issue of convergence. For example, if $f(t)$ is the square wave function, the equation

$$f(t) = \frac{\pi}{2} + \sum_{k=1}^{\infty} \frac{2}{2k-1} \sin((2k-1)t).$$

is only an equality for such t where $f(t)$ is continuous. That is, we do not get an equality for $t = -\pi, 0, \pi$ and all the other discontinuities of $f(t)$. It is not hard to see that when t is an integer multiple of π (which includes all the discontinuities), then

$$\frac{\pi}{2} + \sum_{k=1}^{\infty} \frac{2}{2k-1} \sin((2k-1)t) = \frac{\pi}{2}.$$

We redefine $f(t)$ on $[-\pi, \pi]$ as

$$f(t) = \begin{cases} 0 & \text{if } -\pi < t < 0, \\ \pi & \text{if } 0 < t < \pi, \\ \pi/2 & \text{if } t = -\pi, t = 0, \text{ or } t = \pi, \end{cases}$$

and extend periodically. The series equals this extended $f(t)$ everywhere, including the discontinuities. We will generally not worry about changing the function values at several (finitely many) points.

We will say more about convergence in the next section. Let us however mention briefly an effect of the discontinuity. Let us zoom in near the discontinuity in the square wave. Further, let us plot the first 100 harmonics, see Figure 15.2.6. You will notice that while the series is a very good approximation away from the discontinuities, the error (the overshoot) near the discontinuity at $t = \pi$ does not seem to be getting any smaller. This behavior is known as the *Gibbs phenomenon*. The region where the error is large does get smaller, however, the more terms in the series we take.

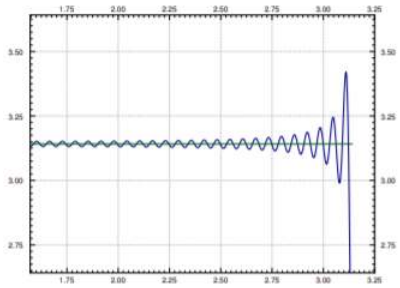


Figure 15.2.6: Gibbs phenomenon in action.

We can think of a periodic function as a "signal" being a superposition of many signals of pure frequency. For example, we could think of the square wave as a tone of certain base frequency. This base frequency is called the *fundamental frequency*. The square wave will be a superposition of many different pure tones of frequencies that are multiples of the fundamental frequency. In music, the higher frequencies are called the *overtones*. All the frequencies that appear are called the *spectrum* of the signal. On the other hand a simple sine wave is only the pure tone (no overtones). The simplest way to make sound using a computer is the square wave, and the sound is very different from a pure tone. If you ever played video games from the 1980s or so, then you heard what square waves sound like.

15.2.4: Footnotes

[1] Named after the French mathematician Jean Baptiste Joseph Fourier (1768 – 1830).

This page titled 15.2: The Trigonometric Series is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.2: The Trigonometric Series** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.3: More on the Fourier Series

Before reading the lecture, it may be good to first try Project IV (Fourier series) from the IODE website: https://conf.math.illinois.edu/iode/fsgui.html. After reading the lecture it may be good to continue with Project V (Fourier series again).

15.3.1: 2L-Periodic Functions

We have computed the Fourier series for a 2π-periodic function, but what about functions of different periods. Well, fear not, the computation is a simple case of change of variables. We can just rescale the independent axis. Suppose that we have a $2L$-periodic function $f(t)$ (L is called the *half period*). Let $S = \frac{\pi}{L}t$. Then the function

$$g(s) = f\left(\frac{L}{\pi}s\right)$$

is 2π-periodic. We want to also rescale all our sines and cosines. We want to write

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}t\right) + b_n \sin\left(\frac{n\pi}{L}t\right).$$

If we change variables to s we see that

$$g(s) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(ns) + b_n \sin(ns).$$

We compute a_n and b_n as before. After we write down the integrals we change variables from s back to t.

$$\begin{aligned}
a_0 &= \frac{1}{\pi}\int_{-\pi}^{\pi} g(s)ds = \frac{1}{L}\int_{-L}^{L} f(t)dt, \\
a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} g(s)\cos(ns)ds = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{n\pi}{L}t\right)dt, \\
b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} g(s)\sin(ns)ds = \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{n\pi}{L}t\right)dt.
\end{aligned}$$

The two most common half periods that show up in examples are π and 1 because of the simplicity. We should stress that we have done no new mathematics, we have only changed variables. If you understand the Fourier series for 2π-periodic functions, you understand it for $2L$-periodic functions. All that we are doing is moving some constants around, but all the mathematics is the same.

Example 15.3.1

Let

$$f(t) = |t| \quad \text{for } -1 < t \le 1,$$

extended periodically. The plot of the periodic extension is given in Figure 15.3.1. Compute the Fourier series of $f(t)$.

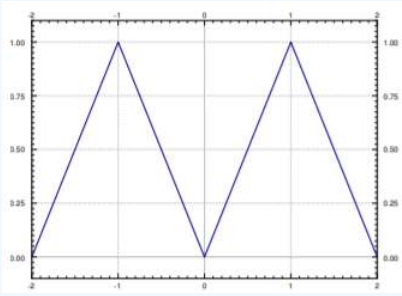


Figure 15.3.1: Periodic extension of the function $f(t)$.

Solution

We want to write $f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi t) + b_n \sin(n\pi t)$. For $n \geq 1$ we note that $|t| \cos(n\pi t)$ is even and hence

$$\begin{aligned} a_n &= \int_{-1}^{1} f(t)\cos(n\pi t)dt \\ &= 2\int_{0}^{1} t\cos(n\pi t)dt \\ &= 2\left[\frac{t}{n\pi}\sin(n\pi t)\right]_{t=0}^{1} - 2\int_{0}^{1}\frac{1}{n\pi}\sin(n\pi t)dt \\ &= 0 + \frac{1}{n^2\pi^2}[\cos(n\pi t)]_{t=0}^{1} = \frac{2((-1)^n - 1)}{n^2\pi^2} = \begin{cases} 0 & \text{if } n \text{ is even,} \\ \dfrac{-4}{n^2\pi^2} & \text{if } n \text{ is odd.} \end{cases} \end{aligned} \tag{15.3.1}$$

Next we find a_0

$$a_0 = \int_{-1}^{1} |t| dt = 1.$$

You should be able to find this integral by thinking about the integral as the area under the graph without doing any computation at all. Finally we can find b_n. Here, we notice that $|t| \sin(n\pi t)$ is odd and, therefore,

$$b_n = \int_{-1}^{1} f(t)\sin(n\pi t)dt = 0.$$

Hence, the series is

$$\frac{1}{2} + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{-4}{n^2\pi^2}\cos(n\pi t)$$

Let us explicitly write down the first few terms of the series up to the 3rd harmonic.

$$\frac{1}{2} - \frac{4}{\pi^2}\cos(\pi t) - \frac{4}{9\pi^2}\cos(3\pi t) - \cdots$$

The plot of these few terms and also a plot up to the 20th harmonic is given in Figure 15.3.2 You should notice how close the graph is to the real function. You should also notice that there is no "Gibbs phenomenon" present as there are no discontinuities.

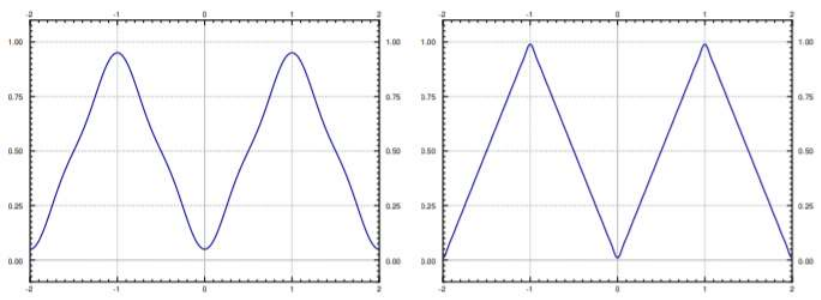


Figure 15.3.2: Fourier series of $f(t)$ up to the 3rd harmonic (left graph) and up to the 20th harmonic (right graph).

15.3.2: Convergence

We will need the one sided limits of functions. We will use the following notation

$$f(c-) = \lim_{t \uparrow c} f(t), \qquad \text{and} \qquad f(c+) = \lim_{t \downarrow c} f(t).$$

If you are unfamiliar with this notation, $\lim_{t\uparrow c} f(t)$ means we are taking a limit of $f(t)$ as t approaches c from below (i.e. $t<c$) and $\lim_{t\downarrow c} f(t)$ means we are taking a limit of $f(t)$ as t approaches c from above (i.e. $t>c$). For example, for the square wave function

$$f(t)=\begin{cases}0 & \text{if} \quad -\pi<t\le 0,\\ \pi & \text{if} \quad 0<t\le \pi,\end{cases} \tag{15.3.2}$$

we have $f(0-)=0$ and $f(0+)=\pi$.

Let $f(t)$ be a function defined on an interval $[a,b]$. Suppose that we find finitely many points $a=t_0,t_1,t_2,\ldots,t_k=b$ in the interval, such that $f(t)$ is continuous on the intervals $(t_0,t_1),(t_1,t_2),\ldots,(t_{k-1},t_k)$. Also suppose that all the one sided limits exist, that is, all of $f(t_0+),f(t_1-),f(t_1+),f(t_2-),f(t_2+),\ldots,f(t_k-)$ exist and are finite. Then we say $f(t)$ is *piecewise continuous*.

If moreover, $f(t)$ is differentiable at all but finitely many points, and $f'(t)$ is piecewise continuous, then $f(t)$ is said to be *piecewise smooth*.

Example 15.3.2

The square wave function (15.3.2) is piecewise smooth on $[-\pi,\pi]$ or any other interval. In such a case we simply say that the function is *piecewise smooth*.

Example 15.3.3

The function $f(t)=|t|$ is piecewise smooth.

Example 15.3.4

The function $f(t)=\dfrac{1}{t}$ is not piecewise smooth on $[-1,1]$ (or any other interval containing zero). In fact, it is not even piecewise continuous.

Example 15.3.5

The function $f(t)=\sqrt[3]{t}$ is not piecewise smooth on $[-1,1]$ (or any other interval containing zero). $f(t)$ is continuous, but the derivative of $f(t)$ is unbounded near zero and hence not piecewise continuous.

Piecewise smooth functions have an easy answer on the convergence of the Fourier series.

Theorem 15.3.1

Suppose $f(t)$ is a $2L$-periodic piecewise smooth function. Let

$$\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n\cos\Big(\frac{n\pi}{L}t\Big)+b_n\sin\Big(\frac{n\pi}{L}t\Big)$$

be the Fourier series for $f(t)$. Then the series converges for all t. If $f(t)$ is continuous near t, then

$$f(t)=\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n\cos\Big(\frac{n\pi}{L}t\Big)+b_n\sin\Big(\frac{n\pi}{L}t\Big).$$

Otherwise

$$\frac{f(t-)+f(t+)}{2}=\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n\cos\Big(\frac{n\pi}{L}t\Big)+b_n\sin\Big(\frac{n\pi}{L}t\Big).$$

If we happen to have that $f(t) = \dfrac{f(t-) + f(t+)}{2}$ at all the discontinuities, the Fourier series converges to $f(t)$ everywhere. We can always just redefine $f(t)$ by changing the value at each discontinuity appropriately. Then we can write an equals sign between $f(t)$ and the series without any worry. We mentioned this fact briefly at the end last section.

Note that the theorem does not say how fast the series converges. Think back to the discussion of the Gibbs phenomenon in the last section. The closer you get to the discontinuity, the more terms you need to take to get an accurate approximation to the function.

15.3.3: Differentiation and Integration of Fourier Series

Not only does Fourier series converge nicely, but it is easy to differentiate and integrate the series. We can do this just by differentiating or integrating term by term.

Theorem 15.3.2

Suppose

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}t\right) + b_n \sin\left(\frac{n\pi}{L}t\right)$$

is a piecewise smooth continuous function and the derivative $f'(t)$ is piecewise smooth. Then the derivative can be obtained by differentiating term by term,

$$f'(t) = \sum_{n=1}^{\infty} \frac{-a_n n\pi}{L} \sin\left(\frac{n\pi}{L}t\right) + \frac{b_n n\pi}{L} \cos\left(\frac{n\pi}{L}t\right).$$

It is important that the function is continuous. It can have corners, but no jumps. Otherwise the differentiated series will fail to converge. For an exercise, take the series obtained for the square wave and try to differentiate the series. Similarly, we can also integrate a Fourier series.

Theorem 15.3.3

Suppose

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}t\right) + b_n \sin\left(\frac{n\pi}{L}t\right)$$

is a piecewise smooth function. Then the antiderivative is obtained by antidifferentiating term by term and so

$$F(t) = \frac{a_0 t}{2} + C + \sum_{n=1}^{\infty} \frac{a_n L}{n\pi} \sin\left(\frac{n\pi}{L}t\right) + \frac{-b_n L}{n\pi} \cos\left(\frac{n\pi}{L}t\right),$$

where $F'(t) = f(t)$ and C is an arbitrary constant.

Note that the series for $F(t)$ is no longer a Fourier series as it contains the $\dfrac{a_0 t}{2}$ term. The antiderivative of a periodic function need no longer be periodic and so we should not expect a Fourier series.

15.3.4: Rates of Convergence and Smoothness

Let us do an example of a periodic function with one derivative everywhere.

Example 15.3.6

Take the function

$$f(t) = \begin{cases} (t+1)t & \text{if} \quad -1 < t \leq 0, \\ (1-t)t & \text{if} \quad 0 < t \leq 1, \end{cases}$$

and extend to a 2-periodic function. The plot is given in Figure 15.3.3

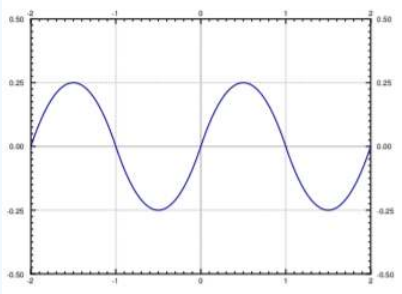


Figure 15.3.3: Smooth 2-periodic function.

Note that this function has one derivative everywhere, but it does not have a second derivative whenever t is an integer.

? Exercise 15.3.1

Compute $f''(0+)$ and $f''(0-)$.

Let us compute the Fourier series coefficients. The actual computation involves several integration by parts and is left to student.

$$\begin{aligned} a_0 &= \int_{-1}^{1} f(t)dt = \int_{-1}^{0} (t+1)tdt + \int_{0}^{1} (1-t)tdt = 0, \\ a_n &= \int_{-1}^{1} f(t)\cos(n\pi t)dt = \int_{-1}^{0} (t+1)t\cos(n\pi t)dt + \int_{0}^{1} (1-t)t\cos(n\pi t)dt = 0 \\ b_n &= \int_{-1}^{1} f(t)\sin(n\pi t)dt = \int_{-1}^{0} (t+1)t\sin(n\pi t)dt + \int_{0}^{1} (1-t)t\sin(n\pi t)dt \\ &= \frac{4(1-(-1)^n)}{\pi^3 n^3} = \begin{cases} \frac{8}{\pi^3 n^3} & \text{if } n \text{ is odd,} \\ 0 & \text{if } n \text{ is even.} \end{cases} \end{aligned} \tag{15.3.3}$$

That is, the series is

$$\sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{8}{\pi^3 n^3}\sin(n\pi t).$$

This series converges very fast. If you plot up to the third harmonic, that is the function

$$\frac{8}{\pi^3}\sin(\pi t) + \frac{8}{27\pi^3}\sin(3\pi t),$$

it is almost indistinguishable from the plot of $f(t)$ in Figure 15.3.3. In fact, the coefficient $\frac{8}{27\pi^3}$ is already just 0.0096 (approximately). The reason for this behavior is the n^3 term in the denominator. The coefficients b_n in this case go to zero as fast as $\frac{1}{n^3}$ goes to zero.

For functions constructed piecewise from polynomials as above, it is generally true that if you have one derivative, the Fourier coefficients will go to zero approximately like $\frac{1}{n^3}$. If you have only a continuous function, then the Fourier coefficients will go to zero as $\frac{1}{n^2}$. If you have discontinuities, then the Fourier coefficients will go to zero approximately as $\frac{1}{n}$. For more general functions the story is somewhat more complicated but the same idea holds, the more derivatives you have, the faster the coefficients go to zero. Similar reasoning works in reverse. If the coefficients go to zero like $\frac{1}{n^2}$ you always obtain a continuous function. If they go to zero like $\frac{1}{n^3}$ you obtain an everywhere differentiable function.

To justify this behavior, take for example the function defined by the Fourier series

$$f(t) = \sum_{n=1}^{\infty} \frac{1}{n^3} \sin(nt).$$

When we differentiate term by term we notice

$$f'(t) = \sum_{n=1}^{\infty} \frac{1}{n^2} \cos(nt).$$

Therefore, the coefficients now go down like $\frac{1}{n^2}$, which means that we have a continuous function. The derivative of $f'(t)$ is defined at most points, but there are points where $f'(t)$ is not differentiable. It has corners, but no jumps. If we differentiate again (where we can) we find that the function $f''(t)$, now fails to be continuous (has jumps)

$$f''(t) = \sum_{n=1}^{\infty} \frac{-1}{n} \sin(nt).$$

This function is similar to the sawtooth. If we tried to differentiate the series again we would obtain

$$\sum_{n=1}^{\infty} -\cos(nt),$$

which does not converge!

? Exercise 15.3.2

Use a computer to plot the series we obtained for $f(t)$, $f'(t)$ and $f''(t)$. That is, plot say the first 5 harmonics of the functions. At what points does $f''(t)$ have the discontinuities?

This page titled 15.3: More on the Fourier Series is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.3: More on the Fourier Series** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.4: Sine and Cosine Series

15.4.1: 4.4.1:Even Periodic Functions

You may have noticed by now that an odd function has no cosine terms in the Fourier series and an even function has no sine terms in the Fourier series. This observation is not a coincidence. Let us look at even and odd periodic function in more detail.

Recall that a function $f(t)$ is odd if $f(-t) = -f(t)$. A function $f(t)$ is even if $f(-t) = f(t)$. For example, $\cos(nt)$ is even and $\sin(nt)$ is odd. Similarly the function t^k is even if k is even and odd when k is odd.

? Exercise 15.4.1

Take two functions $f(t)$ and $g(t)$ and define their product $h(t) = f(t)\,g(t)$.

a. Suppose both $f(t)$ and $g(t)$ are odd, is $h(t)$ odd or even?
b. Suppose one is even and one is odd, is $h(t)$ odd or even?
c. Suppose both are even, is $h(t)$ odd or even?

If $f(t)$ and $g(t)$ are both odd, then $f(t)+g(t)$ is odd. Similarly for even functions. On the other hand, if $f(t)$ is odd and $g(t)$ even, then we cannot say anything about the sum $f(t)+g(t)$. In fact, the Fourier series of any function is a sum of an odd (the sine terms) and an even (the cosine terms) function.

In this section we consider odd and even periodic functions. We have previously defined the $2L$-periodic extension of a function defined on the interval $[-L, L]$. Sometimes we are only interested in the function on the range $[0, L]$ and it would be convenient to have an odd (resp. even) function. If the function is odd (resp. even), all the cosine (resp. sine) terms will disappear. What we will do is take the odd (resp. even) extension of the function to $[-L, L]$ and then extend periodically to a $2L$-periodic function.

Take a function $f(t)$ defined on $[0, L]$. On $(-L, L]$ define the functions

$$\begin{aligned} F_{\text{odd}}(t) &\stackrel{\text{def}}{=} \begin{cases} f(t) & \text{if} \quad 0 \le t \le L, \\ -f(-t) & \text{if} \quad -L < t < 0, \end{cases} \\ F_{\text{even}}(t) &\stackrel{\text{def}}{=} \begin{cases} f(t) & \text{if} \quad 0 \le t \le L, \\ f(-t) & \text{if} \quad -L < t < 0. \end{cases} \end{aligned} \tag{15.4.1}$$

Extend $F_{odd}(t)$ and $F_{even}(t)$ to be $2L$-periodic. Then $F_{odd}(t)$ is called the *odd periodic extension* of $f(t)$, and $F_{even}(t)$ is called the *even periodic extension* of $f(t)$. For the odd extension we generally assume that $f(0) = f(L) = 0$.

? Exercise 15.4.2

Check that $F_{odd}(t)$ is odd and that $F_{even}(t)$ is even. For F_{odd}, assume $f(0) = f(L) = 0$.

✓ Example 15.4.1

Take the function $f(t) = t(1-t)$ defined on $[0, 1]$. Figure 15.4.1 shows the plots of the odd and even extensions of $f(t)$.

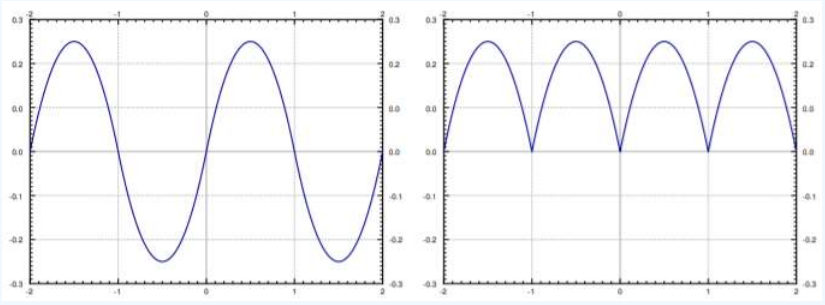


Figure 15.4.1: Odd and even 2-periodic extension of $f(t) = t(1-t)$, $0 \le t \le 1$.

15.4.2: Sine and Cosine Series

Let $f(t)$ be an odd $2L$-periodic function. We write the Fourier series for $f(t)$. First, we compute the coefficients a_n (including $n=0$) and get

$$a_n = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{n\pi}{L}t\right)dt = 0.$$

That is, there are no cosine terms in the Fourier series of an odd function. The integral is zero because $f(t)\cos\left(\frac{n\pi}{L}t\right)$ is an odd function (product of an odd and an even function is odd) and the integral of an odd function over a symmetric interval is always zero. The integral of an even function over a symmetric interval $[-L,L]$ is twice the integral of the function over the interval $[0,L]$. The function $f(t)\sin\left(\frac{n\pi}{L}t\right)$ is the product of two odd functions and hence is even.

$$b_n = \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{n\pi}{L}t\right)dt = \frac{2}{L}\int_{0}^{L} f(t)\sin\left(\frac{n\pi}{L}t\right)dt.$$

We now write the Fourier series of $f(t)$ as

$$\sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}t\right).$$

Similarly, if $f(t)$ is an even $2L$-periodic function. For the same exact reasons as above, we find that $b_n = 0$ and

$$a_n = \frac{2}{L}\int_{0}^{L} f(t)\cos\left(\frac{n\pi}{L}t\right)dt.$$

The formula still works for $n=0$, in which case it becomes

$$a_0 = \frac{2}{L}\int_{0}^{L} f(t)\,dt.$$

The Fourier series is then

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}t\right).$$

An interesting consequence is that the coefficients of the Fourier series of an odd (or even) function can be computed by just integrating over the half interval $[0,L]$. Therefore, we can compute the Fourier series of the odd (or even) extension of a function by computing certain integrals over the interval where the original function is defined.

Theorem 15.4.1

Let $f(t)$ be a piecewise smooth function defined on $[0,L]$. Then the odd periodic extension of $f(t)$ has the Fourier series

$$F_{odd}(t) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}t\right),$$

where

$$b_n = \frac{2}{L}\int_{0}^{L} f(t)\sin\left(\frac{n\pi}{L}t\right)dt.$$

The even periodic extension of $f(t)$ has the Fourier series

$$F_{even}(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}t\right),$$

where

$$a_n = \frac{2}{L}\int_{0}^{L} f(t)\cos\left(\frac{n\pi}{L}t\right)dt.$$

The series $\sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}t\right)$ is called the *sine series* of $f(t)$ and the series $\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}t\right)$ is called the *cosine series* of $f(t)$. We often do not actually care what happens outside of $[0, L]$. In this case, we pick whichever series fits our problem better.

It is not necessary to start with the full Fourier series to obtain the sine and cosine series. The sine series is really the eigenfunction expansion of $f(t)$ using eigenfunctions of the eigenvalue problem $x''+\lambda x=0$ $x(0)=0$, $x(L)=L$. The cosine series is the eigenfunction expansion of $f(t)$ using eigenfunctions of the eigenvalue problem $x''+\lambda x=0$, $x'(0)=0$, $x'(L)=L$. We could have, therefore, gotten the same formulas by defining the inner produ

$$\langle f(t), g(y)\rangle = \int_0^L f(t)g(t)\,dt,$$

and following the procedure of Section 4.2. This point of view is useful, as we commonly use a specific series that arose because our underlying question led to a certain eigenvalue problem. If the eigenvalue problem is not one of the three we covered so far, you can still do an eigenfunction expansion, generalizing the results of this chapter. We will deal with such a generalization in Chapter 5.

Example 15.4.2

Find the Fourier series of the even periodic extension of the function $f(t)=t^2$ for $0\le t\le \pi$.

Solution

We want to write

$$f(t)=\frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos(nt),$$

where

$$a_0=\frac{2}{\pi}\int_0^{\pi} t^2 dt=\frac{2\pi^2}{3},$$

and

$$\begin{aligned} a_n &= \frac{2}{\pi}\int_0^{\pi} t^2\cos(nt)dt = \frac{2}{\pi}\left[t^2\frac{1}{n}\sin(nt)\right]_0^{\pi} - \frac{4}{n\pi}\int_0^{\pi} t\sin(nt)dt \\ &= \frac{4}{n^2\pi}[t\cos(nt)]_0^{\pi} + \frac{4}{n^2\pi}\int_0^{\pi}\cos(nt)dt = \frac{4(-1)^n}{n^2}. \end{aligned} \tag{15.4.2}$$

Note that we have "detected" the continuity of the extension since the coefficients decay as $\frac{1}{n^2}$. That is, the even extension of t^2 has no jump discontinuities. It does have corners, since the derivative, which is an odd function and a sine series, has jumps; it has a Fourier series whose coefficients decay only as $\frac{1}{n}$.

Explicitly, the first few terms of the series are

$$\frac{\pi^2}{3}-4\cos(t)+\cos(2t)-\frac{4}{9}\cos(3t)+\cdots$$

Exercise 15.4.3

a. Compute the derivative of the even extension of $f(t)$ above and verify it has jump discontinuities. Use the actual definition of $f(t)$, not its cosine series!
b. Why is it that the derivative of the even extension of $f(t)$ is the odd extension of $f'(t)$?

15.4.3: Application

Fourier series ties in to the boundary value problems we studied earlier. Let us see this connection in more detail.

Suppose we have the boundary value problem for $0<t<L$.

$$x''(t)+\lambda x(t)=f(t),$$

for the *Dirichlet boundary conditions* $x(0)=0, x(L)=0$. By using the Fredholm alternative (Theorem 4.1.2) we note that as long as λ is *not* an eigenvalue of the underlying homogeneous problem, there exists a unique solution. Note that the eigenfunctions of this eigenvalue problem are the functions $\sin\left(\frac{n\pi}{L}t\right)$. Therefore, to find the solution, we first find the Fourier sine series for $f(t)$. We write x also as a sine series, but with unknown coefficients. We substitute the series for x into the equation and solve for the unknown coefficients. If we have the *Neumann boundary conditions* $x'(0)=0$ and $x'(L)=0$, we do the same procedure using the cosine series.

Let us see how this method works on examples.

Example 15.4.3

Take the boundary value problem for $0<t<1$,

$$x''(t)+2x(t)=f(t),$$

where $f(t)=t$ on $0<t<1$, and satisfying the **Dirichlet boundary** conditions $x(0)=0$ and $x(1)=0$. We write $f(t)$ as a sine series

$$f(t)=\sum_{n=1}^{\infty} c_n \sin(n\pi t).$$

Compute

$$c_n = 2\int_0^1 t\,\sin(n\pi t)\,dt = \frac{2(-1)^{n+1}}{n\pi}.$$

We write $x(t)$ as

$$x(t)=\sum_{n=1}^{\infty} b_n \sin(n\pi t).$$

We plug in to obtain

$$\begin{aligned} x''(t)+2x(t) &= \underbrace{\sum_{n=1}^{\infty} -b_n n^2\pi^2 \sin(n\pi t)}_{x''} + 2\underbrace{\sum_{n=1}^{\infty} b_n \sin(n\pi t)}_{x} \\ &= \sum_{n=1}^{\infty} b_n(2-n^2\pi^2)\sin(n\pi t) \\ &= f(t)=\sum_{n=1}^{\infty}\frac{2\,(-1)^{n+1}}{n\pi}\sin(n\pi t). \end{aligned} \tag{15.4.3}$$

Therefore,

$$b_n(2-n^2\pi^2)=\frac{2(-1)^{n+1}}{n\pi}.$$

or

$$b_n=\frac{2(-1)^{n+1}}{n\pi(2-n^2\pi^2)}.$$

That $2-n^2\pi^2$ is not zero for any n, and that we can solve for b_n, is precisely because 2 is not an eigenvale of the problem. We have thus obtained a Fourier series for the solution

$$x(t)=\sum_{n=1}^{\infty}\frac{2(-1)^{n+1}}{n\pi(2-n^2\pi^2)}\sin(n\pi t).$$

See Figure 15.4.2 for a graph of the solution. Notice that because the eigenfunctions satisfy the boundary conditions, and x is written in terms of the boundary conditions, then x satisfies the boundary conditions.

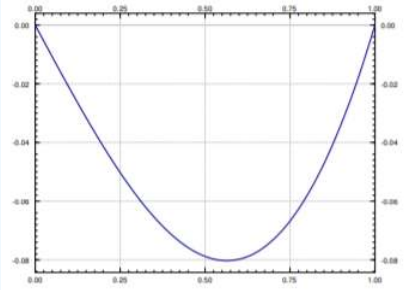


Figure 15.4.2: Plot of the solution of $x'' + 2x = t$, $x(0) = 0$, $x(1) = 0$.

Example 15.4.4

Similarly we handle the Neumann conditions. Take the boundary value problem for $0 < t < 1$,

$$x''(t) + 2x(t) = f(t),$$

where again $f(t) = t$ on $0 < t < 1$, but now satisfying the Neumann boundary conditions $x'(0) = 0$ and $x'(1) = 0$. We write $f(t)$ as a cosine series

$$f(t) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n \cos(n\pi t),$$

where

$$c_0 = 2\int_0^1 t\, dt = 1$$

and

$$c_n = 2\int_0^1 t\cos(n\pi t)dt = \frac{2((-1)^n - 1)}{\pi^2 n^2} = \begin{cases} \frac{-4}{\pi^2 n^2} & \text{if } n \text{ odd}, \\ 0 & \text{if } n \text{ even}. \end{cases}$$

We also write $x(t)$ as a cosine series

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi t).$$

We plug in to obtain

$$\begin{aligned} x''(t) + 2x(t) &= \sum_{n=1}^{\infty} \left[-a_n n^2\pi^2 \cos(n\pi t)\right] + a_0 + 2\sum_{n=1}^{\infty} \left[a_n \cos(n\pi t)\right] \\ &= a_0 + \sum_{n=1}^{\infty} a_n(2 - n^2\pi^2)\cos(n\pi t) \\ &= f(t) = \frac{1}{2} + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{-4}{\pi^2 n^2}\cos(n\pi t). \end{aligned} \tag{15.4.4}$$

Therefore, $a = \frac{1}{2}$ and $a_n = 0$ for n even ($n \geq 2$) and for n odd we have

$$a_n(2 - n^2\pi^2) = \frac{-4}{\pi^2 n^2},$$

or

$$a_n = \frac{-4}{n^2\,\pi^2(2 - n^2\pi^2)}.$$

The Fourier series for the solution $x(t)$ is

$$x(t) = \frac{1}{4} + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{-4}{n^2\pi^2(2 - n^2\pi^2)} \cos(n\pi t).$$

This page titled 15.4: Sine and Cosine Series is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.4: Sine and Cosine Series** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.5: Applications of Fourier Series

15.5.1: Periodically Forced Oscillation

Let us return to the forced oscillations. Consider a mass-spring system as before, where we have a mass m on a spring with spring constant k, with damping c, and a force $F(t)$ applied to the mass. Suppose the forcing function $F(t)$ is $2L$-periodic for some $L > 0$. We have already seen this problem in chapter 2 with a simple $F(t)$.

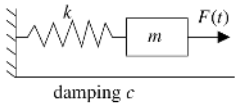


Figure 15.5.1

The equation that governs this particular setup is

$$mx''(t) + cx'(t) + kx(t) = F(t). \tag{15.5.1}$$

The general solution consists of (15.5.1) consists of the complementary solution x_c, which solves the associated homogeneous equation $mx'' + cx' + kx = 0$, and a particular solution of Equation (15.5.1) we call x_p. For $c > 0$, the complementary solution x_c will decay as time goes by. Therefore, we are mostly interested in a particular solution x_p that does not decay and is periodic with the same period as $F(t)$. We call this particular solution the steady periodic solution and we write it as x_{sp} as before. What will be new in this section is that we consider an arbitrary forcing function $F(t)$ instead of a simple cosine.

For simplicity, let us suppose that $c = 0$. The problem with $c > 0$ is very similar. The equation

$$mx'' + kx = 0$$

has the general solution

$$x(t) = A\cos(\omega_0 t) + B\sin(\omega_0 t),$$

where $\omega_0 = \sqrt{\dfrac{k}{m}}$. Any solution to $mx''(t) + kx(t) = F(t)$ is of the form $A\cos(\omega_0 t) + B\sin(\omega_0 t) + x_{sp}$. The steady periodic solution x_{sp} has the same period as $F(t)$.

In the spirit of the last section and the idea of undetermined coefficients we first write

$$F(t) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n \cos\left(\frac{n\pi}{L}t\right) + d_n \sin\left(\frac{n\pi}{L}t\right).$$

Then we write a proposed steady periodic solution x as

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi}{L}t\right) + b_n \sin\left(\frac{n\pi}{L}t\right),$$

where a_n and b_n are unknowns. We plug x into the differential equation and solve for a_n and b_n in terms of c_n and d_n. This process is perhaps best understood by example.

Example 15.5.1

Suppose that $k = 2$, and $m = 1$. The units are again the mks units (meters-kilograms-seconds). There is a jetpack strapped to the mass, which fires with a force of 1 newton for 1 second and then is off for 1 second, and so on. We want to find the steady periodic solution.

Solution

The equation is, therefore,

$$x'' + 2x = F(t),$$

where $F(t)$ is the step function

$$F(t) = \begin{cases} 0 & \text{if} \quad -1 < t < 0, \\ 1 & \text{if} \quad 0 < t < 1, \end{cases}$$

extended periodically. We write

$$F(t) = \frac{c_0}{2} + \sum_{n=1}^{\infty} c_n \cos(n\pi t) + d_n \sin(n\pi t).$$

We compute

$$\begin{aligned} c_n &= \int_{-1}^{1} F(t)\cos(n\pi t)dt = \int_{0}^{1} \cos(n\pi t)dt = 0 \quad \text{for } n \geq 1, \\ c_0 &= \int_{-1}^{1} F(t)dt = \int_{0}^{1} dt = 1, \\ d_n &= \int_{-1}^{1} F(t)\sin(n\pi t)dt \\ &= \int_{0}^{1} \sin(n\pi t)dt \\ &= \left[\frac{-\cos(n\pi t)}{n\pi}\right]_{t=0}^{1} \\ &= \frac{1-(-1)^n}{\pi n} = \begin{cases} \frac{2}{\pi n} & \text{if } n \text{ odd}, \\ 0 & \text{if } n \text{ even}. \end{cases} \end{aligned} \tag{15.5.2}$$

So

$$F(t) = \frac{1}{2} + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{2}{\pi n}\sin(n\pi t).$$

We want to try

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\pi t) + b_n \sin(n\pi t).$$

Once we plug . into the differential equation $x'' + 2x = F(t)$, it is clear that $a_n = 0$ for $n \geq 1$ as there are no corresponding terms in the series for $F(t)$. Similarly $b_n = 0$ for n even. Hence we try

$$x(t) = \frac{a_0}{2} + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} b_n \sin(n\pi t).$$

We plug into the differential equation and obtain

$$\begin{aligned} x'' + 2x &= \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \left[-b_n n^2\pi^2 \sin(n\pi t)\right] + a_0 + 2\sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \left[b_n \sin(n\pi t)\right] \\ &= a_0 + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} b_n(2 - n^2\pi^2)\sin(n\pi t) \\ &= F(t) = \frac{1}{2} + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{2}{\pi n}\sin(n\pi t). \end{aligned} \tag{15.5.3}$$

So $a_0 = \frac{1}{2}$, $b_n = 0$ for even n, and for odd n we get

$$b_n = \frac{2}{\pi n(2 - n^2\pi^2)}.$$

The steady periodic solution has the Fourier series

$$x_{sp}(t) = \frac{1}{4} + \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{2}{\pi n(2-n^2\pi^2)} \sin(n\pi t).$$

We know this is the steady periodic solution as it contains no terms of the complementary solution and it is periodic with the same period as $F(t)$ itself. See Figure 15.5.1 for the plot of this solution.

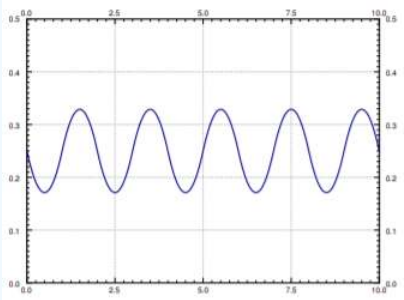


Figure 15.5.1: Plot of the steady periodic solution x_{sp} of Example 15.5.1.

15.5.1.1: Resonance

Just like when the forcing function was a simple cosine, resonance could still happen. Let us assume $c = 0$ and we will discuss only pure resonance. Again, take the equation

$$mx''(t) + kx(t) = F(t).$$

When we expand $F(t)$ and find that some of its terms coincide with the complementary solution to $mx'' + kx = 0$, we cannot use those terms in the guess. Just like before, they will disappear when we plug into the left hand side and we will get a contradictory equation (such as $0 = 1$). That is, suppose

$$x_c = A\cos(\omega_0 t) + B\sin(\omega_0 t),$$

where $\omega_0 = \dfrac{N\pi}{L}$ for some positive integer N. In this case we have to modify our guess and try

$$x(t) = \frac{a_0}{2} + t\left(a_N \cos\left(\frac{N\pi}{L}t\right) + b_N \sin\left(\frac{N\pi}{L}t\right)\right) + \sum_{\substack{n=1 \\ n\neq N}}^{\infty} a_n \cos\left(\frac{n\pi}{L}t\right) + b_n \sin\left(\frac{n\pi}{L}t\right).$$

In other words, we multiply the offending term by t. From then on, we proceed as before.

Of course, the solution will not be a Fourier series (it will not even be periodic) since it contains these terms multiplied by t. Further, the terms $t\left(a_N \cos\left(\dfrac{N\pi}{L}t\right) + b_N \sin\left(\dfrac{N\pi}{L}t\right)\right)$ will eventually dominate and lead to wild oscillations. As before, this behavior is called *pure resonance* or just *resonance*.

Note that there now may be infinitely many resonance frequencies to hit. That is, as we change the frequency of F (we change L), different terms from the Fourier series of F may interfere with the complementary solution and will cause resonance. However, we should note that since everything is an approximation and in particular c is never actually zero but something very close to zero, only the first few resonance frequencies will matter.

Example 15.5.2

Find the steady periodic solution to the equation

$$2x'' + 18\pi^2 x = F(t), \tag{15.5.4}$$

where

$$F(t) = \begin{cases} -1 & \text{if} \quad -1 < t < 0, \\ 1 & \text{if} \quad 0 < t < 1, \end{cases}$$

extended periodically. We note that

$$F(t) = \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{4}{\pi n} \sin(n\pi t).$$

? Exercise 15.5.1

Compute the Fourier series of F to verify the above equation.

As $\sqrt{\frac{k}{m}} = \sqrt{\frac{18\pi^2}{2}} = 3\pi$, the solution to (15.5.4) is

$$x(t) = c_1 \cos(3\pi t) + c_2 \sin(3\pi t) + x_p(t)$$

for some particular solution x_p.

If we just try an x_p given as a Fourier series with $\sin(n\pi t)$ as usual, the complementary equation, $2x'' + 18\pi^2 x = 0$, eats our 3rd harmonic. That is, the term with $\sin(3\pi t)$ is already in in our complementary solution. Therefore, we pull that term out and multiply it by t. We also add a cosine term to get everything right. That is, we try

$$x_p(t) = a_3 t \cos(3\pi t) + b_3 t \sin(3\pi t) + \sum_{\substack{n=1 \\ n \text{ odd} \\ n \neq 3}}^{\infty} b_n \sin(n\pi t).$$

Let us compute the second derivative.

$$x_p''(t) = -6a_3\pi \sin(3\pi t) - 9\pi^2 a_3 t \cos(3\pi t) + 6b_3\pi \cos(3\pi t) - 9\pi^2 b_3 t \sin(3\pi t) + \sum_{\substack{n=1 \\ n \text{ odd} \\ n \neq 3}}^{\infty} (-n^2\pi^2 b_n) \sin(n\pi t).$$

We now plug into the left hand side of the differential equation.

$$\begin{aligned} 2x_p'' + 18\pi^2 x_p = \; & -12a_3\pi \sin(3\pi t) - 18\pi^2 a_3 t \cos(3\pi t) + 12b_3\pi \cos(3\pi t) - 18\pi^2 b_3 t \sin(3\pi t) \\ & + 18\pi^2 a_3 t \cos(3\pi t) \qquad\qquad + 18\pi^2 b_3 t \sin(3\pi t) \\ & + \sum_{\substack{n=1 \\ n \text{ odd} \\ n \neq 3}}^{\infty} \left(-2n^2\pi^2 b_n + 18\pi^2 b_n\right) \sin(n\pi t). \end{aligned} \tag{15.5.5}$$

If we simplify we obtain

$$2x_p'' + 18\pi^2 x = -12a_3\pi \sin(3\pi t) + 12b_3\pi \cos(3\pi t) + \sum_{\substack{n=1 \\ n \text{ odd} \\ n \neq 3}}^{\infty} (-2n^2\pi^2 b_n + 18\pi^2 b_n) \sin(n\pi t. \,)$$

This series has to equal to the series for $F(t)$. We equate the coefficients and solve for a_3 and b_n.

$$\begin{aligned} a_3 &= \frac{4/(3\pi)}{-12\pi} = \frac{-1}{9\pi^2}, \\ b_3 &= 0, \\ b_n &= \frac{4}{n\pi(18\pi^2 - 2n^2\pi^2)} = \frac{2}{\pi^3 n(9 - n^2)} \qquad \text{for } n \text{ odd and } n \neq 3. \end{aligned} \tag{15.5.6}$$

That is,

$$x_p(t) = \frac{-1}{9\pi^2} t \cos(3\pi t) + \sum_{\substack{n=1 \\ n \text{ odd} \\ n \neq 3}}^{\infty} \frac{2}{\pi^3 n(9 - n^2)} \sin(n\pi t. \,)$$

When $c > 0$, you will not have to worry about pure resonance. That is, there will never be any conflicts and you do not need to multiply any terms by t. There is a corresponding concept of practical resonance and it is very similar to the ideas we already explored in Chapter 2. Basically what happens in practical resonance is that one of the coefficients in the series for x_{sp} can get very big. We will not go into details here.

- Jiří Lebl (Oklahoma State University).These pages were supported by NSF grants DMS-0900885 and DMS-1362337.

This page titled 15.5: Applications of Fourier Series is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.5: Applications of Fourier Series** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.6: PDEs, Separation of Variables, and The Heat Equation

Let us recall that a *partial differential equation* or PDE is an equation containing the partial derivatives with respect to several independent variables. Solving PDEs will be our main application of Fourier series.

A PDE is said to be *linear* if the dependent variable and its derivatives appear at most to the first power and in no functions. We will only talk about linear PDEs. Together with a PDE, we usually have specified some *boundary conditions*, where the value of the solution or its derivatives is specified along the boundary of a region, and/or some *initial conditions* where the value of the solution or its derivatives is specified for some initial time. Sometimes such conditions are mixed together and we will refer to them simply as *side conditions*.

We will study three specific partial differential equations, each one representing a more general class of equations. First, we will study the *heat equation*, which is an example of a *parabolic* PDE. Next, we will study the *wave equation*, which is an example of a *hyperbolic PDE*. Finally, we will study the *Laplace equation*, which is an example of an *elliptic PDE*. Each of our examples will illustrate behavior that is typical for the whole class.

15.6.1: Heat on an Insulated Wire

Let us first study the heat equation. Suppose that we have a wire (or a thin metal rod) of length L that is insulated except at the endpoints. Let x denote the position along the wire and let t denote time. See Figure 15.6.1.

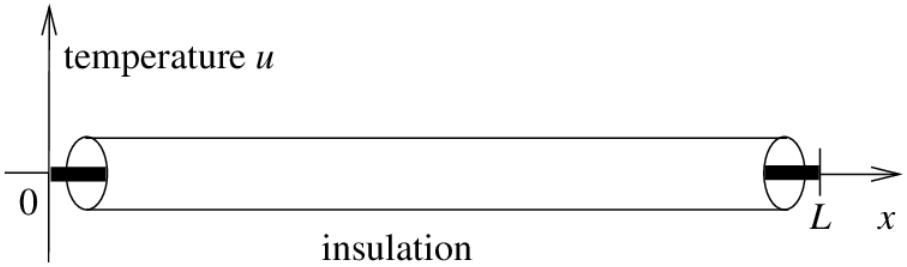


Figure 15.6.1: Insulated wire.

Let $u(x,t)$ denote the temperature at point x at time t. The equation governing this setup is the so-called *one-dimensional heat equation*:

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2},$$

where $k > 0$ is a constant (the *thermal conductivity* of the material). That is, the change in heat at a specific point is proportional to the second derivative of the heat along the wire. This makes sense; if at a fixed t the graph of the heat distribution has a maximum (the graph is concave down), then heat flows away from the maximum. And vice-versa.

We will generally use a more convenient notation for partial derivatives. We will write u_t instead of $\frac{\partial u}{\partial t}$, and we will write u_{xx} instead of $\frac{\partial^2 u}{\partial x^2}$. With this notation the heat equation becomes

$$u_t = ku_{xx}\,.$$

For the heat equation, we must also have some boundary conditions. We assume that the ends of the wire are either exposed and touching some body of constant heat, or the ends are insulated. For example, if the ends of the wire are kept at temperature 0, then we must have the conditions

$$u(0,t) = 0 \quad \text{and} \quad u(L,t) = 0.$$

If, on the other hand, the ends are also insulated we get the conditions

$$u_x(0,t) = 0 \quad \text{and} \quad u_x(L,t) = 0.$$

Let us see why that is so. If u_x is positive at some point x_0, then at a particular time, u is smaller to the left of x_0, and higher to the right of x_0. Heat is flowing from high heat to low heat, that is to the left. On the other hand if u_x is negative then heat is again flowing from high heat to low heat, that is to the right. So when u_x is zero, that is a point through which heat is not flowing. In other words, $u_x(0,t) = 0$ means no heat is flowing in or out of the wire at the point $x = 0$.

We have two conditions along the x-axis as there are two derivatives in the x direction. These side conditions are said to be *homogeneous* (i.e., u or a derivative of u is set to zero).

We also need an initial condition—the temperature distribution at time $t = 0$. That is,

$$u(x, 0) = f(x),$$

for some known function $f(x)$. This initial condition is not a homogeneous side condition.

15.6.2: Separation of Variables

The heat equation is linear as u and its derivatives do not appear to any powers or in any functions. Thus the principle of superposition still applies for the heat equation (without side conditions). If u_1 and u_2 are solutions and c_1, c_2 are constants, then $u = c_1u_1 + c_2u_2$ is also a solution.

? Exercise 15.6.1

Verify the principle of superposition for the heat equation.

Superposition also preserves some of the side conditions. In particular, if u_1 and u_2 are solutions that satisfy $u(0, t) = 0$ and $u_(L, t) = 0$, and c_1, c_2 are constants, then $u = c_1u_1 + c_2u_2$ is still a solution that satisfies $u(0, t) = 0$ and$u_(L, t) = 0$. Similarly for the side conditions $u_x(0, t) = 0$ and $u_x(L, t) = 0$. In general, superposition preserves all homogeneous side conditions.

The method of *separation of variables* is to try to find solutions that are sums or products of functions of one variable. For example, for the heat equation, we try to find solutions of the form

$$u(x, t) = X(x)T(t).$$

That the desired solution we are looking for is of this form is too much to hope for. What is perfectly reasonable to ask, however, is to find enough "building-block" solutions of the form $u(x, t) = X(x)T(t)$ using this procedure so that the desired solution to the PDE is somehow constructed from these building blocks by the use of superposition.

Let us try to solve the heat equation

$$u_t = ku_{xx} \quad \text{with} \quad u(0, t) = 0, \quad u(L, t) = 0, \quad \text{and} \quad u(x, 0) = f(x).$$

Let us guess $u(x, t) = X(x)T(t)$. We will try to make this guess satisfy the differential equation, $u_t = ku_{xx}$, and the homogeneous side conditions, $u(0, t) = 0$ and $u(L, t) = 0$. Then, as superposition preserves the differential equation and the homogeneous side conditions, we will try to build up a solution from these building blocks to solve the nonhomogeneous initial condition $u(x, 0) = f(x)$.

First we plug $u(x, t) = X(x)T(t)$ into the heat equation to obtain

$$X(x)T'(t) = kX''(x)T(t).$$

We rewrite as

$$\frac{T'(t)}{kT(t)} = \frac{X''(x)}{X(x)}.$$

This equation must hold for all x and all t. But the left hand side does not depend on x and the right hand side does not depend on t. Hence, each side must be a constant. Let us call this constant $-\lambda$ (the minus sign is for convenience later). We obtain the two equations

$$\frac{T'(t)}{kT(t)} = -\lambda = \frac{X''(x)}{X(x)}.$$

In other words

$$\begin{aligned} X''(x) + \lambda X(x) &= 0, \\ T'(t) + \lambda kT(t) &= 0. \end{aligned} \tag{15.6.1}$$

The boundary condition $u(0,t)=0$ implies $X(0)T(t)=0$. We are looking for a nontrivial solution and so we can assume that $T(t)$ is not identically zero. Hence $X(0)=0$. Similarly, $u(L,t)=0$ implies $X(L)=0$. We are looking for nontrivial solutions X of the eigenvalue problem $X''+\lambda X=0, X(0)=0, X(L)=0$. We have previously found that the only eigenvalues are $\lambda_n=\frac{n^2\pi^2}{L^2}$, for integers $n\geq 1$, where eigenfunctions are $\sin\left(\frac{n\pi}{L}x\right)$. Hence, let us pick the solutions

$$X_n(x)=\sin\left(\frac{n\pi}{L}x\right).$$

The corresponding T_n must satisfy the equation

$$T_n'(t)+\frac{n^2\pi^2}{L^2}kT_n(t)=0.$$

By the method of integrating factor, the solution of this problem is

$$T_n(t)=e^{\frac{-n^2\pi^2}{L^2}kt}.$$

It will be useful to note that $T_n(0)=1$. Our building-block solutions are

$$u_n(x,t)=X_n(x)T_n(t)=\sin\left(\frac{n\pi}{L}x\right)e^{\frac{-n^2\pi^2}{L^2}kt}.$$

We note that $u_n(x,0)=\sin\left(\frac{n\pi}{L}x\right)$. Let us write $f(x)$ as the sine series

$$f(x)=\sum_{n=1}^{\infty}b_n\sin\left(\frac{n\pi}{L}x\right).$$

That is, we find the Fourier series of the odd periodic extension of $f(x)$. We used the sine series as it corresponds to the eigenvalue problem for $X(x)$ above. Finally, we use superposition to write the solution as

$$u(x,t)=\sum_{n=1}^{\infty}b_nu_n(x,t)=\sum_{n=1}^{\infty}b_n\sin\left(\frac{n\pi}{L}x\right)e^{\frac{-n^2\pi^2}{L^2}kt}.$$

Why does this solution work? First note that it is a solution to the heat equation by superposition. It satisfies $u(0,t)=0$ and $u(L,t)=0$, because $x=0$ or $x=L$ makes all the sines vanish. Finally, plugging in $t=0$, we notice that $T_n(0)=1$ and so

$$u(x,0)=\sum_{n=1}^{\infty}b_nu_n(x,0)=\sum_{n=1}^{\infty}b_n\sin\left(\frac{n\pi}{L}x\right)=f(x).$$

Example 15.6.1

Suppose that we have an insulated wire of length 1, such that the ends of the wire are embedded in ice (temperature 0). Let $k=0.003$. Then suppose that initial heat distribution is $u(x,0)=50x(1-x)$. See Figure 15.6.2

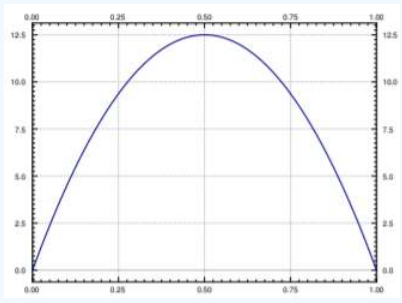


Figure 15.6.2: Initial distribution of temperature in the wire.

We want to find the temperature function $u(x,t)$. Let us suppose we also want to find when (at what t) does the maximum temperature in the wire drop to one half of the initial maximum of 12.5.

We are solving the following PDE problem:

$$\begin{aligned} u_t &= 0.003u_{xx}, \\ u(0,t) &= u(1,t) = 0, \\ u(x,0) &= 50x(1-x) \quad \text{for } 0 < x < 1. \end{aligned} \tag{15.6.2}$$

We write $f(x) = 50x(1-x)$ for $0 < x < 1$ as a sine series. That is, $f(x) = \sum_{n=1}^{\infty} b_n \sin(n\pi x)$, where

$$b_n = 2\int_0^1 50x(1-x)\sin(n\pi x)dx = \frac{200}{\pi^3 n^3} - \frac{200(-1)^n}{\pi^3 n^3} = \begin{cases} 0 & \text{if } n \text{ even}, \\ \frac{400}{\pi^3 n^3} & \text{if } n \text{ odd}. \end{cases}$$

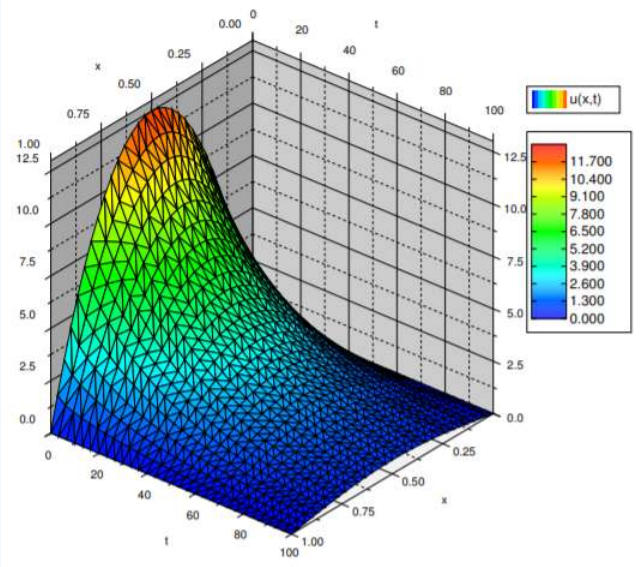


Figure 15.6.3: Plot of the temperature of the wire at position x at time t.

The solution $u(x,t)$, plotted in Figure 15.6.3 for \(0 \leq t \leq 100\), is given by the series:

$$u(x,t) = \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{400}{\pi^3 n^3} \sin(n\pi x) e^{-n^2\pi^2 0.003t}.$$

Finally, let us answer the question about the maximum temperature. It is relatively easy to see that the maximum temperature will always be at $x = 0.5$, in the middle of the wire. The plot of $u(x,t)$ confirms this intuition.

If we plug in $x = 0.5$ we get

$$u(0.5,t) = \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{400}{\pi^3 n^3} \sin(n\pi 0.5) e^{-n^2\pi^2 0.003t}.$$

For $n = 3$ and higher (remember n is only odd), the terms of the series are insignificant compared to the first term. The first term in the series is already a very good approximation of the function. Hence

$$u(0.5,t) \approx \frac{400}{\pi^3} e^{-\pi^2 0.003t}.$$

The approximation gets better and better as t gets larger as the other terms decay much faster. Let us plot the function $0.5, t$, the temperature at the midpoint of the wire at time t, in Figure 15.6.4 The figure also plots the approximation by the first term.

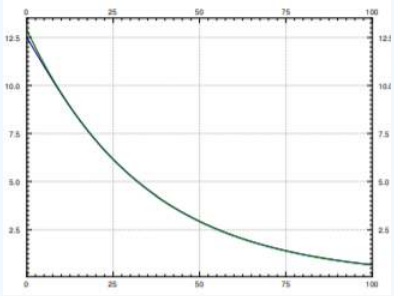


Figure 15.6.4: Temperature at the midpoint of the wire (the bottom curve), and the approximation of this temperature by using only the first term in the series (top curve).

After $t = 5$ or so it would be hard to tell the difference between the first term of the series for $u(x,t)$ and the real solution $u(x,t)$. This behavior is a general feature of solving the heat equation. If you are interested in behavior for large enough t, only the first one or two terms may be necessary.

Let us get back to the question of when is the maximum temperature one half of the initial maximum temperature. That is, when is the temperature at the midpoint $12.5/2 = 6.25$. We notice on the graph that if we use the approximation by the first term we will be close enough. We solve

$$6.25 = \frac{400}{\pi^3} e^{-\pi^2 0.003 t}.$$

That is,

$$t = \frac{\ln \frac{6.25\pi^3}{400}}{-\pi^2 0.003} \approx 24.5.$$

So the maximum temperature drops to half at about $t = 24.5$.

We mention an interesting behavior of the solution to the heat equation. The heat equation "smoothes" out the function $f(x)$ as t grows. For a fixed t, the solution is a Fourier series with coefficients $b_n e^{\frac{-n^2\pi^2}{L^2} kt}$. If $t > 0$, then these coefficients go to zero faster than any $\frac{1}{n^P}$ for any power p. In other words, the Fourier series has infinitely many derivatives everywhere. Thus even if the function $f(x)$ has jumps and corners, then for a fixed $t > 0$, the solution $u(x,t)$ as a function of x is as smooth as we want it to be.

Example 15.6.2

When the initial condition is already a sine series, then there is no need to compute anything, you just need to plug in. Consider

$$u_t = 0.3\, u_{xx}, \qquad u(0,t) = u(1,t) = 0, \qquad u(x,0) = 0.1\sin(\pi t) + \sin(2\pi t).$$

The solution is then

$$u(x,t) = 0.1\sin(\pi t) e^{-0.3\pi^2 t} + \sin(2\pi t) e^{-1.2\pi^2 t}.$$

15.6.3: Insulated Ends

Now suppose the ends of the wire are insulated. In this case, we are solving the equation

$$u_t = k u_{xx} \quad \text{with} \quad u_x(0,t) = 0, \quad u_x(L,t) = 0, \quad \text{and} \quad u(x,0) = f(x).$$

Yet again we try a solution of the form $u(x,t) = X(x)T(t)$. By the same procedure as before we plug into the heat equation and arrive at the following two equations

$$\begin{aligned} X''(x) + \lambda X(x) &= 0, \\ T'(t) + \lambda k T(t) &= 0. \end{aligned} \tag{15.6.3}$$

At this point the story changes slightly. The boundary condition $u_x(0,t)=0$ implies $X'(0)T(t)=0$. Hence $X'(0)=0$. Similarly, $u_x(L,t)=0$ implies $X'(L)=0$. We are looking for nontrivial solutions X of the eigenvalue problem $X''+\lambda X=0$, $X'(0)=0$, $X'(L)=0,$. We have previously found that the only eigenvalues are $\lambda_n=\frac{n^2\pi^2}{L^2}$, for integers $n\geq 0$, where eigenfunctions are $\cos(\frac{n\pi}{L})X$ (we include the constant eigenfunction). Hence, let us pick solutions

$$X_n(x)=\cos(\frac{n\pi}{L}x) \quad \text{and} \quad X_0(x)=1.$$

The corresponding T_n must satisfy the equation

$$T_n'(t)+\frac{n^2\pi^2}{L^2}kT_n(t)=0.$$

For $n\geq 1$, as before,

$$T_n(t)=e^{\frac{-n^2\pi^2}{L^2}kt}.$$

For $n=0$, we have $T_0'(t)=0$ and hence $T_0(t)=1$. Our building-block solutions will be

$$u_n(x,t)=X_n(x)T_n(t)=\cos\Big(\frac{n\pi}{L}x\Big)e^{\frac{-n^2\pi^2}{L^2}kt},$$

and

$$u_0(x,t)=1.$$

We note that $u_n(x,0)=\cos\big(\frac{n\pi}{L}x\big)$. Let us write f using the cosine series

$$f(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}a_n\cos\Big(\frac{n\pi}{L}x\Big).$$

That is, we find the Fourier series of the even periodic extension of $f(x)$.

We use superposition to write the solution as

$$u(x,t)=\frac{a_0}{2}+\sum_{n=1}^{\infty}a_nu_n(x,t)=\frac{a_0}{2}+\sum_{n=1}^{\infty}a_n\cos\Big(\frac{n\pi}{L}x\Big)e^{\frac{-n^2\pi^2}{L^2}kt}.$$

Example 15.6.3

Let us try the same equation as before, but for insulated ends. We are solving the following PDE problem

$$\begin{aligned} u_t &= 0.003u_{xx}, \\ u_x(0,t) &= u_x(1,t)=0, \\ u(x,0) &= 50x(1-x) \quad \text{for } 0<x<1. \end{aligned} \tag{15.6.4}$$

For this problem, we must find the cosine series of $u(x,0)$. For $0<x<1$ we have

$$50x(1-x)=\frac{25}{3}+\sum_{\substack{n=2\\ n \text{ even}}}^{\infty}\left(\frac{-200}{\pi^2n^2}\right)\cos(n\pi x).$$

The calculation is left to the reader. Hence, the solution to the PDE problem, plotted in Figure 15.6.5, is given by the series

$$u(x,t)=\frac{25}{3}+\sum_{\substack{n=2\\ n \text{ even}}}^{\infty}\left(\frac{-200}{\pi^2n^2}\right)\cos(n\pi x)e^{-n^2\pi^2 0.003t}.$$

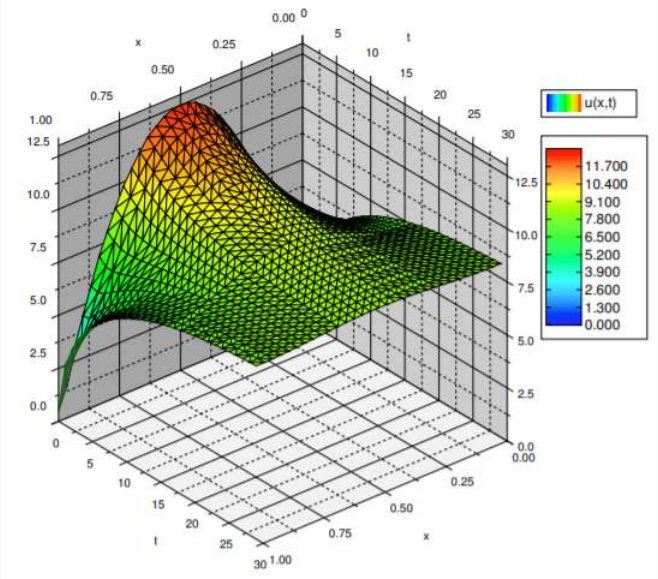


Figure 15.6.5: Plot of the temperature of the insulated wire at position x at time t.

Note in the graph that the temperature evens out across the wire. Eventually, all the terms except the constant die out, and you will be left with a uniform temperature of \(\frac{25}{3} \approx{8.33}\) along the entire length of the wire.

Let us expand on the last point. The constant term in the series is

$$\frac{a_0}{2} = \frac{1}{L} \int_0^L f(x)\,dx.$$

In other words, $\frac{a_0}{2}$ is the average value of $f(x)$, that is, the average of the initial temperature. As the wire is insulated everywhere, no heat can get out, no heat can get in. So the temperature tries to distribute evenly over time, and the average temperature must always be the same, in particular it is always $\frac{a_0}{2}$. As time goes to infinity, the temperature goes to the constant $\frac{a_0}{2}$ everywhere.

This page titled 15.6: PDEs, Separation of Variables, and The Heat Equation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.6: PDEs, Separation of Variables, and The Heat Equation** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.7: One Dimensional Wave Equation

Imagine we have a tensioned guitar string of length L. Suppose we only consider vibrations in one direction. That is, let x denote the position along the string, let t denote time, and let y denote the displacement of the string from the rest position. See Figure 15.7.1

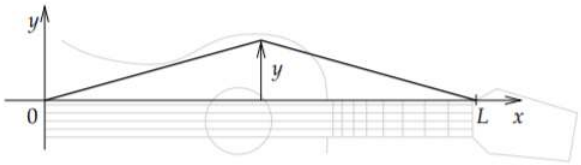


Figure 15.7.1: Vibrating string of length L, x is position, y is displacement.

The equation that governs this setup is the so-called *one-dimensional wave equation*:

$$y_{tt} = a^2 y_{xx},$$

for some constant $a > 0$. The intuition is similar to the heat equation, replacing velocity with acceleration: the acceleration at a specific point is proportional to the second derivative of the shape of the string. In other words when the string is concave down then u_{xx} is negative and the string wants to accelerate downwards, so u_{tt} should be negative. And vice versa. The wave equation is an example of a hyperbolic PDE.

Assume that the ends of the string are fixed in place:

$$y(0,t) = 0 \quad \text{and} \quad y(L,t) = 0.$$

Note that we have two conditions along the x axis as there are two derivatives in the x direction.

There are also two derivatives along the t direction and hence we need two further conditions here. We need to know the initial position and the initial velocity of the string. That is, for some known functions $f(x)$ and $g(x)$, we impose

$$y(x,0) = f(x) \quad \text{and} \quad y_t(x,0) = g(x).$$

As the equation is again linear, superposition works just as it did for the heat equation. And again we will use separation of variables to find enough building-block solutions to get the overall solution. There is one change however. It will be easier to solve two separate problems and add their solutions.

The two problems we will solve are

$$\begin{aligned} &w_{tt} = a^2 w_{xx}, \\ &w(0,t) = w(L,t) = 0, \\ &w(x,0) = 0 && \text{for } 0 < x < L, \\ &w_t(x,0) = g(x) && \text{for } 0 < x < L, \end{aligned} \tag{15.7.1}$$

and

$$\begin{aligned} &z_{tt} = a^2 z_{xx}, \\ &z(0,t) = z(L,t) = 0, \\ &z(x,0) = f(x) && \text{for } 0 < x < L, \\ &z_t(x,0) = 0 && \text{for } 0 < x < L. \end{aligned} \tag{15.7.2}$$

The principle of superposition implies that $y = w + z$ solves the wave equation and furthermore $y(x,0) = w(x,0) + z(x,0) = f(x)$ and $y_t(x,0) = w_t(x,0) + z_t(x,0) = g(x)$. Hence, y is a solution to

$$\begin{aligned} &y_{tt} = a^2 y_{xx}, \\ &y(0,t) = y(L,t) = 0, \\ &y(x,0) = f(x) && \text{for } 0 < x < L, \\ &y_t(x,0) = g(x) && \text{for } 0 < x < L. \end{aligned} \tag{15.7.3}$$

The reason for all this complexity is that superposition only works for homogeneous conditions such as $y(0,t) = y(L,t) = 0$, $y(x,0) = 0$, or $y_t(x,0) = 0$. Therefore, we will be able to use the idea of separation of variables to find many building-block solutions solving all the homogeneous conditions. We can then use them to construct a solution solving the remaining nonhomogeneous condition.

Let us start with (15.7.1). We try a solution of the form $w(x,t) = X(x)T(t)$ again. We plug into the wave equation to obtain

$$X(x)T''(t) = a^2 X''(x)T(t).$$

Rewriting we get

$$\frac{T''(t)}{a^2 T(t)} = \frac{X''(x)}{X(x)}.$$

Again, left hand side depends only on t and the right hand side depends only on x. Therefore, both equal a constant, which we will denote by $-\lambda$.

$$\frac{T''(t)}{a^2 T(t)} = -\lambda = \frac{X''(x)}{X(x)}.$$

We solve to get two ordinary differential equations

$$\begin{aligned} X''(x) + \lambda X(x) &= 0, \\ T''(t) + \lambda a^2 T(t) &= 0. \end{aligned} \tag{15.7.4}$$

The conditions $0 = w(0,t) = X(0)T(t)$ implies $X(0) = 0$ and $w(L,t) = 0$ implies that $X(L) = 0$. Therefore, the only nontrivial solutions for the first equation are when $\lambda = \lambda_n = \dfrac{n^2\pi^2}{L^2}$ and they are

$$X_n(x) = \sin\left(\frac{n\pi}{L}x\right).$$

The general solution for T for this particular λ_n is

$$T_n(t) = A\cos\left(\frac{n\pi a}{L}t\right) + B\sin\left(\frac{n\pi a}{L}t\right).$$

We also have the condition that $w(x,0) = 0$ or $X(x)T(0) = 0$. This implies that $T(0) = 0$, which in turn forces $A = 0$. It is convenient to pick $B = \dfrac{L}{n\pi a}$ (you will see why in a moment) and hence

$$T_n(t) = \frac{L}{n\pi a}\sin\left(\frac{n\pi a}{L}t\right).$$

Our building-block solutions are

$$w_n(x,t) = \frac{L}{n\pi a}\sin\left(\frac{n\pi}{L}x\right)\sin\left(\frac{n\pi a}{L}t\right).$$

We differentiate in t, that is

$$\frac{\partial w_n}{\partial t}(x,t) = \sin\left(\frac{n\pi}{L}x\right)\cos\left(\frac{n\pi a}{L}t\right).$$

Hence,

$$\frac{\partial w_n}{\partial t}(x,0) = \sin\left(\frac{n\pi}{L}x\right).$$

We expand $g(x)$ in terms of these sines as

$$g(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right).$$

Using superposition we can just write down the solution to (15.7.1) as a series

$$w(x,t) = \sum_{n=1}^{\infty} b_n w_n(x,t) = \sum_{n=1}^{\infty} b_n \frac{L}{n\pi a} \sin\Big(\frac{n\pi}{L}x\Big) \sin\Big(\frac{n\pi a}{L}t\Big).$$

Exercise 15.7.1

Check that $w(x,0) = 0$ and $w_t(x,0) = g(x)$.

Similarly we proceed to solve (15.7.2). We again try $z(x,y) = X(x)T(t)$. The procedure works exactly the same at first. We obtain

$$\begin{aligned} X''(x) + \lambda X(x) &= 0, \\ T''(t) + \lambda a^2 T(t) &= 0. \end{aligned} \tag{15.7.5}$$

and the conditions $X(0) = 0$, $X(L) = 0$. So again $\lambda = \lambda_n = \dfrac{n^2\pi^2}{L^2}$ and

$$X_n(x) = \sin\Big(\frac{n\pi}{L}x\Big).$$

This time the condition on T is $T'(0) = 0$. Thus we get that $B = 0$ and we take

$$T_n(t) = \cos\Big(\frac{n\pi a}{L}t\Big).$$

Our building-block solution will be

$$z_n(x,t) = \sin\Big(\frac{n\pi}{L}x\Big) \cos\Big(\frac{n\pi a}{L}t\Big).$$

As $z_n(x,0) = \sin\big(\frac{n\pi}{L}x\big)$, we expand $f(x)$ in terms of these sines as

$$f(x) = \sum_{n=1}^{\infty} c_n \sin\Big(\frac{n\pi}{L}x\Big).$$

And we write down the solution to (15.7.2) as a series

$$z(x,t) = \sum_{n=1}^{\infty} c_n z_n(x,t) = \sum_{n=1}^{\infty} c_n \sin\Big(\frac{n\pi}{L}x\Big) \cos\Big(\frac{n\pi a}{L}t\Big).$$

Exercise 15.7.2

Fill in the details in the derivation of the solution of (15.7.2). Check that the solution satisfies all the side conditions.

Putting these two solutions together, let us state the result as a theorem.

Theorem 15.7.1

Take the equation

$$\begin{aligned} &y_{tt} = a^2 y_{xx}, \\ &y(0,t) = y(L,t) = 0, \\ &y(x,0) = f(x) && \text{for } 0 < x < L, \\ &y_t(x,0) = g(x) && \text{for } 0 < x < L, \end{aligned} \tag{15.7.6}$$

where

$$f(x) = \sum_{n=1}^{\infty} c_n \sin\Big(\frac{n\pi}{L}x\Big),$$

and

$$g(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{L}x\right).$$

Then the solution $y(x,t)$ can be written as a sum of the solutions of (4.7.4) and (4.7.5). In other words,

$$\begin{aligned} y(x,t) &= \sum_{n=1}^{\infty} b_n \frac{L}{n\pi a}\sin\left(\frac{n\pi}{L}x\right)\sin\left(\frac{n\pi a}{L}t\right) + c_n \sin\left(\frac{n\pi}{L}x\right)\cos\left(\frac{n\pi a}{L}t\right) \\ &= \sum_{n=1}^{\infty} \sin\left(\frac{n\pi}{L}x\right)\left[b_n \frac{L}{n\pi a}\sin\left(\frac{n\pi a}{L}t\right) + c_n \cos\left(\frac{n\pi a}{L}t\right)\right]. \end{aligned} \tag{15.7.7}$$

Example 15.7.1

Let us try a simple example of a plucked string. Suppose that a string of length 2 is plucked in the middle such that it has the initial shape given in Figure 15.7.2 That is

$$f(x) = \begin{cases} 0.1x & \text{if } 0 \le x \le 1, \\ 0.1(2-x) & \text{if } 1 < x \le 2 \end{cases}.$$

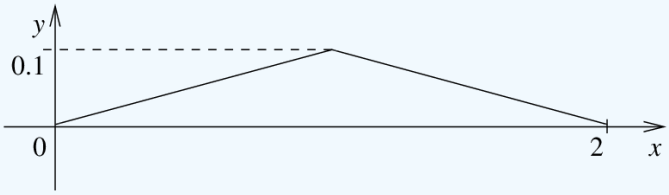


Figure 15.7.2: Initial shape of a plucked string from Example 15.7.1.

The string starts at rest ($g(x) = 0$). Suppose that $a = 1$ in the wave equation for simplicity. In other words, we wish to solve the problem:

$$\begin{aligned} &y_{tt} = y_{xx}, \\ &y(0,t) = y(2,t) = 0, \\ &y(x,0) = f(x) \quad \text{and} \quad y_t(x,0) = 0. \end{aligned} \tag{15.7.8}$$

We leave it to the reader to compute the sine series of $f(x)$. The series will be

$$f(x) = \sum_{n=1}^{\infty} \frac{0.8}{n^2\pi^2}\sin\left(\frac{n\pi}{2}\right)\sin\left(\frac{n\pi}{2}x\right).$$

Note that $\sin\left(\frac{n\pi}{2}\right)$ is the sequence $1, 0, -1, 0, 1, 0, -1, \ldots$ for $n = 1, 2, 3, 4, \ldots$. Therefore,

$$f(x) = \frac{0.8}{\pi^2}\sin\left(\frac{\pi}{2}x\right) - \frac{0.8}{9\pi^2}\sin\left(\frac{3\pi}{2}x\right) + \frac{0.8}{25\pi^2}\sin\left(\frac{5\pi}{2}x\right) - \cdots$$

The solution $y(x,t)$ is given by

$$\begin{aligned} y(x,t) &= \sum_{n=1}^{\infty} \frac{0.8}{n^2\pi^2}\sin\left(\frac{n\pi}{2}\right)\sin\left(\frac{n\pi}{2}x\right)\cos\left(\frac{n\pi}{2}t\right) \\ &= \sum_{m=1}^{\infty} \frac{0.8(-1)^{m+1}}{(2m-1)^2\pi^2}\sin\left(\frac{(2m-1)\pi}{2}x\right)\cos\left(\frac{(2m-1)\pi}{2}t\right) \\ &= \frac{0.8}{\pi^2}\sin\left(\frac{\pi}{2}x\right)\cos\left(\frac{\pi}{2}t\right) - \frac{0.8}{9\pi^2}\sin\left(\frac{3\pi}{2}x\right)\cos\left(\frac{3\pi}{2}t\right) \\ &\qquad + \frac{0.8}{25\pi^2}\sin\left(\frac{5\pi}{2}x\right)\cos\left(\frac{5\pi}{2}t\right) - \cdots \end{aligned} \tag{15.7.9}$$

See Figure 15.7.3 for a plot for $0 < t < 3$. Notice that unlike the heat equation, the solution does not become "smoother," the "sharp edges" remain. We will see the reason for this behavior in the next section where we derive the solution to the wave

equation in a different way.

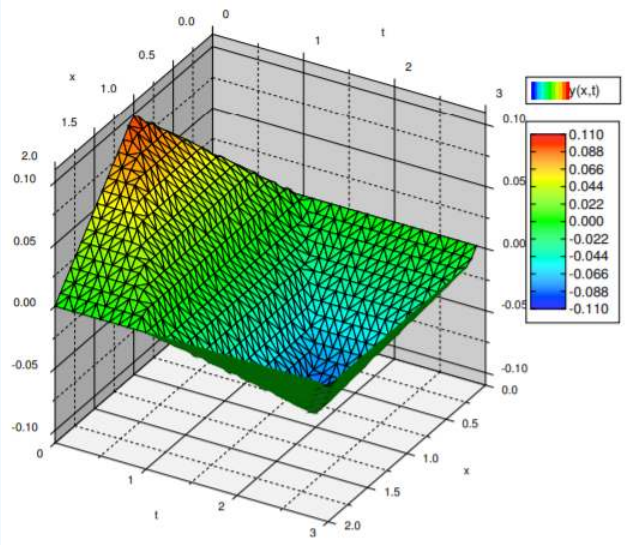


Figure 15.7.3: Shape of the plucked string for $0 < t < 3$.

Make sure you understand what the plot, such as the one in the figure, is telling you. For each fixed t, you can think of the function $y(x,t)$ as just a function of x. This function gives you the shape of the string at time t. See Figure 15.7.4 for plots of at y as a function of x at several different values of t. On this plot you can see the sharp edges remaining much better.

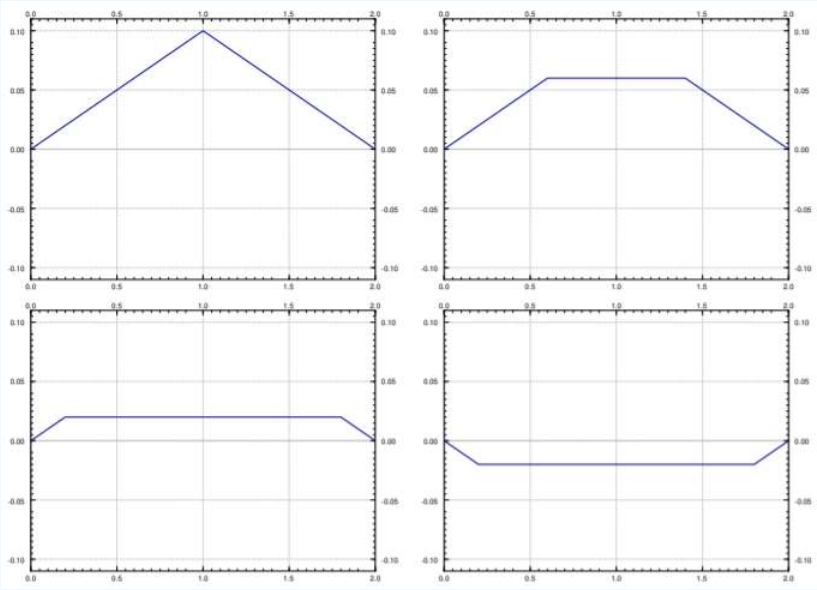


Figure 15.7.4: Plucked string for $t = 0$, $t = 0.4$, $t = 0.8$, and $t = 1.2$.

One thing to take away from all this is how a guitar sounds. Notice that the (angular) frequencies that come up in the solution are $n\frac{\pi a}{L}$. That is, there is a certain base *fundamental frequency* $\frac{\pi a}{L}$, and then we also get all the multiples of this frequency, which in music are called the *overtones*. Which overtones appear and with what amplitude is what musicians call the *timbre* of the note. Mathematicians usually call this the *spectrum*. Because all the frequencies are multiples of one frequency (the fundamental) we get a nice pleasing sound.

The fundamental frequency $\frac{\pi a}{L}$ increases as we decrease length L. That is, if we place a finger on the fingerboard and then pluck a string we get a higher note. The constant a is given by

$$a = \sqrt{\frac{T}{\rho}},$$

where T is tension and ρ is the linear density of the string. Tightening the string (turning the tuning peg on a guitar) increases a and hence produces a higher fundamental frequency (a higher note). On the other hand using a heavier string reduces a and produces a lower fundamental frequency (a lower note). A bass guitar has longer thicker strings, while a ukulele has short strings made of lighter material.

Something rather interesting is the almost symmetry between space and time. In its simplest form we see this symmetry in the solutions

$$\sin\left(\frac{n\pi}{L}x\right)\sin\left(\frac{n\pi a}{L}t\right).$$

Except for the a, time and space are just the same.

In general, the solution for a fixed x is a Fourier series in t, for a fixed t it is a Fourier series in x, and the coefficients are related. If the shape $f(x)$ or the initial velocity have lots of corners, then the sound wave will have lots of corners. That is because the Fourier coefficients of the initial shape decay to zero (as $n \to \infty$) at the same rate as the Fourier coefficients of the wave in time (for some fixed x). So if you use a sharp object to pick the string, you get a sharper sound with lots of high frequency components, while if you use your thumb, you get a softer sound without so many high overtones. Similarly if you pluck close to the bridge, you are getting a pluck that looks more like the sawtooth, and you get an even sharper sound.

In fact, if you look at the formula for the solution, you see that for any fixed x we get an almost arbitrary Fourier series in t, everything except the constant term. You can essentially obtain any sound you want by plucking the string in just the right way. Of course we are considering an ideal string of no stiffness and no air resistance. Those variables clearly impact the sound as well.

This page titled 15.7: One Dimensional Wave Equation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.7: One Dimensional Wave Equation** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.8: D'Alembert Solution of The Wave Equation

We have solved the wave equation by using Fourier series. But it is often more convenient to use the so-called *d'Alembert solution to the wave equation.*[1] While this solution can be derived using Fourier series as well, it is really an awkward use of those concepts. It is easier and more instructive to derive this solution by making a correct change of variables to get an equation that can be solved by simple integration.

Suppose we have the wave equation

$$y_{tt} = a^2 y_{xx}. \tag{15.8.1}$$

We wish to solve the equation (15.8.1) given the conditions

$$\begin{aligned} y(0,t) &= y(L,t) = 0 \qquad && \text{for all } t, \\ y(x,0) &= f(x) && 0 < x < L, \qquad (15.8.2) \\ y_t(x,0) &= g(x) && 0 < x < L. \qquad (15.8.3) \end{aligned}$$

15.8.1: Change of Variables

We will transform the equation into a simpler form where it can be solved by simple integration. We change variables to $\xi = x - at$, $\eta = x + at$. The chain rule says:

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial \xi}{\partial x}\frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial x}\frac{\partial}{\partial \eta} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}, \\ \frac{\partial}{\partial t} &= \frac{\partial \xi}{\partial t}\frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial t}\frac{\partial}{\partial \eta} = -a\frac{\partial}{\partial \xi} + a\frac{\partial}{\partial \eta}. \end{aligned} \tag{15.8.4}$$

We compute

$$\begin{aligned} y_{xx} &= \frac{\partial^2 y}{\partial x^2} = \left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}\right)\left(\frac{\partial y}{\partial \xi} + \frac{\partial y}{\partial \eta}\right) = \frac{\partial^2 y}{\partial \xi^2} + 2\frac{\partial^2 y}{\partial \xi \partial \eta} + \frac{\partial^2 y}{\partial \eta^2}, \\ y_{tt} &= \frac{\partial^2 y}{\partial^2} = \left(-a\frac{\partial}{\partial \xi} + a\frac{\partial}{\partial \eta}\right)\left(-a\frac{\partial y}{\partial \xi} + a\frac{\partial y}{\partial \eta}\right) = a^2\frac{\partial^2 y}{\partial \xi^2} - 2a^2\frac{\partial^2 y}{\partial \xi \partial \eta} + a^2\frac{\partial^2 y}{\partial \eta^2}. \end{aligned} \tag{15.8.5}$$

In the above computations, we used the fact from calculus that $\frac{\partial^2 y}{\partial \xi \partial \eta} = \frac{\partial^2 y}{\partial \eta \partial \xi}$. We plug what we got into the wave equation,

$$0 = a^2 y_{xx} - y_{tt} = 4a^2\frac{\partial^2 y}{\partial \xi \partial \eta} = 4a^2 y_{\xi\eta}.$$

Therefore, the wave equation (15.8.1) transforms into $y_{\xi\eta} = 0$. It is easy to find the general solution to this equation by integrating twice. Keeping ξ constant, we integrate with respect to η first[2] and notice that the constant of integration depends on ξ; for each ξ we might get a different constant of integration. We get $y_\xi = C(\xi)$. Next, we integrate with respect to ξ and notice that the constant of integration must depend on η. Thus, $y = \int C(\xi)d\xi + B(\eta)$. The solution must, therefore, be of the following form for some functions $A(\xi)$ and $B(\eta)$:

$$y = A(\xi) + B(\eta) = A(x - at) + B(x + at).$$

The solution is a superposition of two functions (waves) traveling at speed a in opposite directions. The coordinates ξ and η are called the *characteristic coordinates*, and a similar technique can be applied to more complicated hyperbolic PDE. And in fact, in Section 1.9 it is used to solve first order linear PDE. Basically, to solve the wave equation (or more general hyperbolic equations) we find certain characteristic curves along which the equation is really just an ODE, or a pair of ODEs. In this case these are the curves where ξ and η are constant.

15.8.2: D'Alembert's Formula

We know what any solution must look like, but we need to solve for the given side conditions. We will just give the formula and see that it works. First let $F(x)$ denote the odd extension of $f(x)$, and let $G(x)$ denote the odd extension of $g(x)$. Define

$$A(x) = \frac{1}{2}F(x) - \frac{1}{2a}\int_0^x G(s)ds, \quad B(x) = \frac{1}{2}F(x) + \frac{1}{2a}\int_0^x G(s)ds.$$

We claim this $A(x)$ and $B(x)$ give the solution. Explicitly, the solution is $y(x,t) = A(x-at) + B(x+at)$ or in other words:

$$\begin{aligned} y(x,t) &= \frac{1}{2}F(x-at) - \frac{1}{2a}\int_0^{x-at} G(s)ds + \frac{1}{2}F(x+at) + \frac{1}{2a}\int_0^{x+at} G(s)ds \qquad (15.8.6)\\ &= \frac{F(x-at)+F(x+at)}{2} + \frac{1}{2a}\int_{x-at}^{x+at} G(s)ds. \qquad (15.8.7)\end{aligned}$$

Let us check that the d'Alembert formula really works.

$$y(x,0) = \frac{1}{2}F(x) - \frac{1}{2a}\int_0^{x} G(s)ds + \frac{1}{2}F(x) + \frac{1}{2a}\int_0^{x} G(s)ds = F(x).$$

So far so good. Assume for simplicity F is differentiable. And we use the first form of (15.8.6) as it is easier to differentiate. By the fundamental theorem of calculus we have

$$y_t(x,t) = \frac{-a}{2}F'(x-at) + \frac{1}{2}G(x-at) + \frac{a}{2}F'(x+at) + \frac{1}{2}G(x+at).$$

So

$$y_t(x,0) = \frac{-a}{2}F'(x) + \frac{1}{2}G(x) + \frac{a}{2}F'(x) + \frac{1}{2}G(x) = G(x).$$

Yay! We're smoking now. OK, now the boundary conditions. Note that $F(x)$ and $G(x)$ are odd. Also $\int_0^x G(s)ds$ is an even function of x because $G(x)$ is odd (to see this fact, do the substitution $s = -v$). So

$$\begin{aligned} y(0,t) &= \frac{1}{2}F(-at) - \frac{1}{2a}\int_0^{-at} G(s)ds + \frac{1}{2}F(at) + \frac{1}{2a}\int_0^{at} G(s)ds \\ &= \frac{-1}{2}F(at) - \frac{1}{2a}\int_0^{at} G(s)ds + \frac{1}{2}F(at) + \frac{1}{2a}\int_0^{at} G(s)ds = 0. \end{aligned} \qquad (15.8.8)$$

Note that $F(x)$ and $G(x)$ are $2L$ periodic. We compute

$$\begin{aligned} y(L,t) &= \frac{1}{2}F(L-at) - \frac{1}{2a}\int_0^{L-at} G(s)ds + \frac{1}{2}F(L+at) + \frac{1}{2a}\int_0^{L+at} G(s)ds \\ &= \frac{1}{2}F(-L-at) - \frac{1}{2a}\int_0^{L} G(s)ds - \frac{1}{2a}\int_0^{-at} G(s)ds + \\ &= \frac{1}{2}F(L+at) + \frac{1}{2a}\int_0^{L} G(s)ds + \frac{1}{2a}\int_0^{at} G(s)ds \\ &= \frac{-1}{2}F(L+at) - \frac{1}{2a}\int_0^{at} G(s)ds + \frac{1}{2}F(L+at) + \frac{1}{2a}\int_0^{at} G(s)ds = 0. \end{aligned} \qquad (15.8.9)$$

And voilà, it works.

Example 15.8.1

D'Alembert says that the solution is a superposition of two functions (waves) moving in the opposite direction at "speed" a. To get an idea of how it works, let us work out an example. Consider the simpler setup

$$\begin{aligned} y_{tt} &= y_{xx}, \\ y(0,t) &= y(1,t) = 0, \\ y(x,0) &= f(x), \\ y_t(x,0) &= 0. \end{aligned} \qquad (15.8.10)$$

Here $f(x)$ is an impulse of height 1 centered at $x = 0.5$:

$$f(x) = \begin{cases} 0 & \text{if} \quad 0 \le x < 0.45, \\ 20(x-0.45) & \text{if} \quad 0.45 \le x < 0.5, \\ 20(0.55-x) & \text{if} \quad 0.5 \le x < 0.55 \\ 0 & \text{if} \quad 0.55 \le x \le 1. \end{cases}$$

The graph of this impulse is the top left plot in Figure 15.8.1.

Let $F(x)$ be the odd periodic extension of $f(x)$. Then from (15.8.6) we know that the solution is given as

$$y(x,t) = \frac{F(x-t)+F(x+t)}{2}.$$

It is not hard to compute specific values of $y(x,t)$. For example, to compute $y(0.1, 0.6)$ we notice $x-t=-0.5$ and $x+t=0.7$. Now $F(-0.5) = -f(0.5) = -20(0.55-0.5) = -1$ and $F(0.7) = f(0.7) = 0$. Hence $y(0.1, 0.6) = \frac{-1+0}{2} = -0.5$. As you can see the d'Alembert solution is much easier to actually compute and to plot than the Fourier series solution. See Figure 15.8.1 for plots of the solution y for several different t.

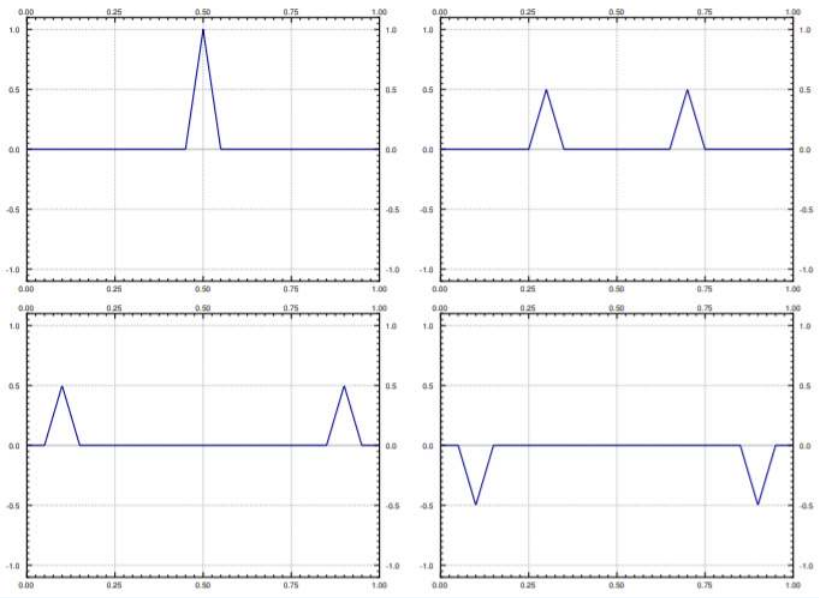


Figure 15.8.1: Plot of the d'Alembert solution for $t=0$, $t=0.2$, $t=0.4$, and $t=0.6$.

15.8.3: Another Way to Solve for the Side Conditions

It is perhaps easier and more useful to memorize the procedure rather than the formula itself. The important thing to remember is that a solution to the wave equation is a superposition of two waves traveling in opposite directions. That is,

$$y(x,t) = A(x-at) + B(x+at).$$

If you think about it, the exact formulas for A and B are not hard to guess once you realize what kind of side conditions $y(x,t)$ is supposed to satisfy. Let us give the formula again, but slightly differently. Best approach is to do this in stages. When $g(x)=0$ (and hence $G(x)=0$) we have the solution

$$\frac{F(x-at)+F(x+at)}{2}.$$

On the other hand, when $f(x)=0$ (and hence $F(x)=0$), we let

$$H(x) = \int_0^x G(s)ds.$$

The solution in this case is

$$\frac{1}{2a}\int_{x-at}^{x+at} G(s)ds = \frac{-H(x-at)+H(x+at)}{2a}.$$

By superposition we get a solution for the general side conditions (15.8.2) (when neither $f(x)$ nor $g(x)$ are identically zero).

$$y(x,t) = \frac{F(x-at)+F(x+at)}{2} + \frac{-H(x-at)+H(x+at)}{2a}. \tag{15.8.11}$$

Do note the minus sign before the H, and the a in the second denominator.

? Exercise 15.8.1

Check that the new formula (15.8.11)satisfies the side conditions (15.8.2).

Warning: Make sure you use the odd extensions $F(x)$ and $G(x)$, when you have formulas for $f(x)$ and $g(x)$. The thing is, those formulas in general hold only for $0 < x < L$, and are not usually equal to $F(x)$ and $G(x)$ for other x.

15.8.4: Remarks

Let us remark that the formula $y(x,t) = A(x-at) + B(x+at)$ is the reason why the solution of the wave equation doesn't get as time goes on, that is, why in the examples where the initial conditions had corners, the solution also has corners at every time t.

The corners bring us to another interesting remark. Nobody ever notices at first that our example solutions are not even differentiable (they have corners): In Example 15.8.1 above, the solution is not differentiable whenever $x = t + 0.5$ or $x = -t + 0.5$ for example. Really to be able to compute u_{xx} or u_{tt}, you need not one, but two derivatives. Fear not, we could think of a shape that is very nearly $F(x)$ but does have two derivatives by rounding the corners a little bit, and then the solution would be very nearly $\frac{F(x-t)+F(x+t)}{2}$ and nobody would notice the switch.

One final remark is what the d'Alembert solution tells us about what part of the initial conditions influence the solution at a certain point. We can figure this out by Let us suppose that the string is very long (perhaps infinite) for simplicity. Since the solution at time t is

$$y(x,t) = \frac{F(x-at)+F(x+at)}{2} + \frac{1}{2a}\int_{x-at}^{x+at} G(s)\,ds,$$

we notice that we have only used the initial conditions in the interval $[x-at, x+at]$. These two endpoints are called the *wavefronts*, as that is where the wave front is given an initial ($t = 0$) disturbance at x. So if $a = 1$, an observer sitting at $x = 0$ at time $t = 1$ has only seen the initial conditions for x in the range $[-1, 1]$ and is blissfully unaware of anything else. This is why for example we do not know that a supernova has occurred in the universe until we see its light, millions of years from the time when it did in fact happen.

15.8.5: Footnotes

[1] Named after the French mathematician Jean le Rond d'Alembert (1717 – 1783).

[2] There is nothing special about η, you can integrate with ξ first, if you wish.

- Jiří Lebl (Oklahoma State University).These pages were supported by NSF grants DMS-0900885 and DMS-1362337.

This page titled 15.8: D'Alembert Solution of The Wave Equation is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.8: D'Alembert Solution of The Wave Equation** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.9: Steady State Temperature and the Laplacian

Suppose we have an insulated wire, a plate, or a 3-dimensional object. We apply certain fixed temperatures on the ends of the wire, the edges of the plate, or on all sides of the 3-dimensional object. We wish to find out what is the steady state temperature distribution. That is, we wish to know what will be the temperature after long enough period of time.

We are really looking for a solution to the heat equation that is not dependent on time. Let us first solve the problem in one space variable. We are looking for a function u that satisfies

$$u_t = ku_{xx},$$

but such that $u_t = 0$ for all x and t. Hence, we are looking for a function of x alone that satisfies $u_{xx} = 0$. It is easy to solve this equation by integration and we see that $u = Ax + B$ for some constants A and B.

Suppose we have an insulated wire, and we apply constant temperature T_1 at one end (say where $x = 0$) and T_2 on the other end (at $x = L$ where L is the length of the wire). Then our steady state solution is

$$u(x) = \frac{T_2 - T_1}{L}x + T_1.$$

This solution agrees with our common sense intuition with how the heat should be distributed in the wire. So in one dimension, the steady state solutions are basically just straight lines.

Things are more complicated in two or more space dimensions. Let us restrict to two space dimensions for simplicity. The heat equation in two space variables is

$$u_t = k(u_{xx} + u_{yy}), \tag{15.9.1}$$

or more commonly written as $u_t = k\Delta u$ or $u_t = k\nabla^2 u$. Here the Δ and ∇^2 symbols mean $\dfrac{\partial^2}{\partial x^2} + \dfrac{\partial^2}{\partial y^2}$. We will use Δ from now on. The reason for using such a notation is that you can define Δ to be the right thing for any number of space dimensions and then the heat equation is always $u_t = k\Delta u$. The operator Δ is called the **Laplacian**.

OK, now that we have notation out of the way, let us see what does an equation for the steady state solution look like. We are looking for a solution to Equation (15.9.1) that does not depend on t, or in other words $u_t = 0$. Hence we are looking for a function $u(x, y)$ such that

$$\Delta u = u_{xx} + u_{yy} = 0.$$

This equation is called the *Laplace equation*[1]. Solutions to the Laplace equation are called **harmonic functions** and have many nice properties and applications far beyond the steady state heat problem.

Harmonic functions in two variables are no longer just linear (plane graphs). For example, you can check that the functions $x^2 - y^2$ and xy are harmonic. However, if you remember your multi-variable calculus we note that if u_{xx} is positive, u is concave up in the x direction, then u_{yy} must be negative and u must be concave down in the y direction. Therefore, a harmonic function can never have any "hilltop" or "valley" on the graph. This observation is consistent with our intuitive idea of steady state heat distribution; the hottest or coldest spot will not be inside.

Commonly the Laplace equation is part of a so-called *Dirichlet problem*[2]. That is, we have a region in the xy-plane and we specify certain values along the boundaries of the region. We then try to find a solution u defined on this region such that u agrees with the values we specified on the boundary.

For simplicity, we consider a rectangular region. Also for simplicity we specify boundary values to be zero at 3 of the four edges and only specify an arbitrary function at one edge. As we still have the principle of superposition, we can use this simpler solution to derive the general solution for arbitrary boundary values by solving 4 different problems, one for each edge, and adding those solutions together. This setup is left as an exercise.

We wish to solve the following problem. Let h and w be the height and width of our rectangle, with one corner at the origin and lying in the first quadrant.

$$\Delta u = 0, \tag{15.9.2}$$
$$u(0,y) = 0 \qquad \text{for } 0 < y < h, \tag{15.9.3}$$
$$u(x,h) = 0 \qquad \text{for } 0 < x < w, \tag{15.9.4}$$
$$u(w,y) = 0 \qquad \text{for } 0 < y < h, \tag{15.9.5}$$
$$u(x,0) = f(x) \qquad \text{for } 0 < x < w. \tag{15.9.6}$$

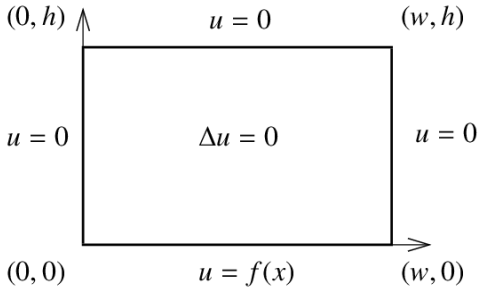


Figure 15.9.1

The method we apply is separation of variables. Again, we will come up with enough building-block solutions satisfying all the homogeneous boundary conditions (all conditions except (15.9.6)). We notice that superposition still works for the equation and all the homogeneous conditions. Therefore, we can use the Fourier series for $f(x)$ to solve the problem as before.

We try $u(x,y) = X(x)Y(y)$. We plug u into the equation to get

$$X''Y + XY'' = 0.$$

We put the Xs on one side and the Ys on the other to get

$$-\frac{X''}{X} = \frac{Y''}{Y}.$$

The left hand side only depends on x and the right hand side only depends on y. Therefore, there is some constant λ such that $\lambda = \dfrac{-X''}{X} = \dfrac{Y''}{Y}$. And we get two equations

$$X'' + \lambda X = 0,$$
$$Y'' - \lambda Y = 0.$$

Furthermore, the homogeneous boundary conditions imply that $X(0) = X(w) = 0$ and $Y(h) = 0$. Taking the equation for X we have already seen that we have a nontrivial solution if and only if $\lambda = \lambda_n = \dfrac{n^2\pi^2}{w^2}$ and the solution is a multiple of

$$X_n(x) = \sin\left(\frac{n\pi}{w}x\right).$$

For these given λ_n, the general solution for Y (one for each n) is

$$Y_n(y) = A_n \cosh\left(\frac{n\pi}{w}y\right) + B_n \sinh\left(\frac{n\pi}{w}y\right). \tag{15.9.7}$$

We only have one condition on Y_n and hence we can pick one of A_n or B_n to be something convenient. It will be useful to have $Y_n(0) = 1$, so we let $A_n = 1$. Setting $Y_n(h) = 0$ and solving for B_n we get that

$$B_n = \frac{-\cosh\left(\dfrac{n\pi h}{w}\right)}{\sinh\left(\dfrac{n\pi h}{w}\right)}.$$

After we plug the A_n and B_n we into (15.9.7) and simplify, we find

$$Y_n(y) = \frac{\sinh\left(\frac{n\pi(h-y)}{w}\right)}{\sinh\left(\frac{n\pi h}{w}\right)}.$$

We define $u_n(x,y) = X_n(x)Y_n(y)$. And note that u_n satisfies (15.9.2) - (15.9.5).

Observe that

$$u_n(x,0) = X_n(x)Y_n(0) = \sin\left(\frac{n\pi}{w}x\right).$$

Suppose

$$f(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{w}\right).$$

Then we get a solution of (15.9.2) - (15.9.6) of the following form.

$$u(x,y) = \sum_{n=1}^{\infty} b_n u_n(x,y) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi}{w}x\right)\left(\frac{\sinh\left(\frac{n\pi(h-y)}{w}\right)}{\sinh\left(\frac{n\pi h}{w}\right)}\right).$$

As u_n satisfies Equation (15.9.2) - (15.9.5) and any linear combination (finite or infinite) of u_n must also satisfy (15.9.2) - (15.9.5), we see that u must satisfy Equations (15.9.2) - (15.9.5). By plugging in $y = 0$ it is easy to see that u satisfies (15.9.6) as well.

Example 15.9.1

Suppose that we take $w = h = \pi$ and we let $f(x) = \pi$. We compute the sine series for the function π(we will get the square wave). We find that for $0 < x < \pi$ we have

$$f(x) = \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{4}{n}\sin(nx).$$

Therefore the solution $u(x,y)$, see Figure 15.9.2, to the corresponding Dirichlet problem is given as

$$u(x,y) = \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{4}{n}\sin(nx)\left(\frac{\sinh(n(\pi-y))}{\sinh(n\pi)}\right).$$

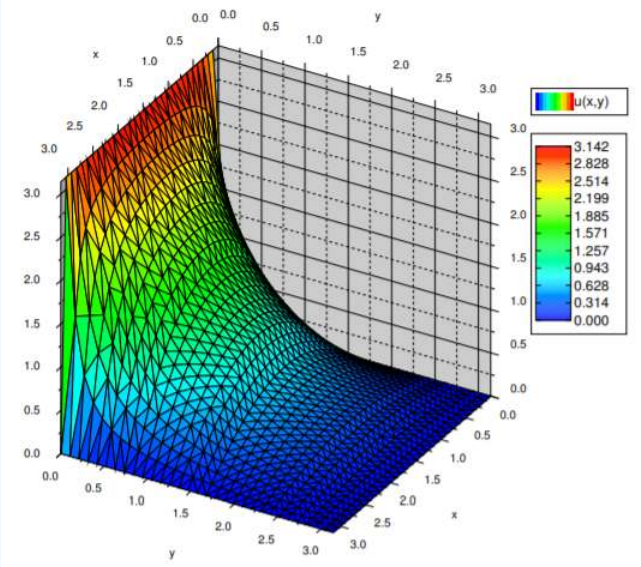


Figure 15.9.2: Steady state temperature of a square plate, three sides held at zero and one side held at π.

This scenario corresponds to the steady state temperature on a square plate of width π with 3 sides held at 0 degrees and one side held at π degrees. If we have arbitrary initial data on all sides, then we solve four problems, each using one piece of nonhomogeneous data. Then we use the principle of superposition to add up all four solutions to have a solution to the original problem.

A different way to visualize solutions of the Laplace equation is to take a wire and bend it so that it corresponds to the graph of the temperature above the boundary of your region. Cut a rubber sheet in the shape of your region—a square in our case—and stretch it fixing the edges of the sheet to the wire. The rubber sheet is a good approximation of the graph of the solution to the Laplace equation with the given boundary data.

15.9.1: Footnotes

[1] Named after the French mathematician Pierre-Simon, marquis de Laplace (1749–1827).

[2] Named after the German mathematician Johann Peter Gustav Lejeune Dirichlet (1805–1859).

This page titled 15.9: Steady State Temperature and the Laplacian is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.9: Steady State Temperature and the Laplacian** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.10: Dirichlet Problem in the Circle and the Poisson Kernel

15.10.1: Laplace in Polar Coordinates

A more natural setting for the Laplace equation $\Delta u = 0$ is the circle rather than the square. On the other hand, what makes the problem somewhat more difficult is that we need polar coordinates.

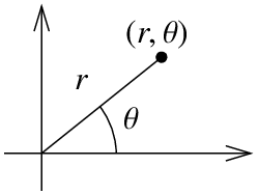


Figure 15.10.1

Recall that the polar coordinates for the (x, y)-plane are (r, θ):

$$x = r\cos\theta, \quad y = r\sin\theta,$$

where $r \geq 0$ and $-\pi < \theta < \pi$. So (x, y) is distance r from the origin at angle θ from the positive x-axis.

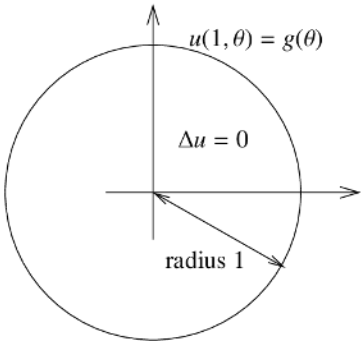


Figure 15.10.2

Now that we know our coordinates, let us give the problem we wish to solve. We have a circular region of radius 1, and we are interested in the Dirichlet problem for the Laplace equation for this region. Let $u(r, \theta)$ denote the temperature at the point (r, θ) in polar coordinates. We have the problem:

$$\Delta u = 0, \qquad \text{for } r < 1, \tag{15.10.1}$$

$$u(1, \theta) = g(\theta), \qquad \text{for } \pi < \theta \leq \pi. \tag{15.10.2}$$

The first issue we face is that we do not know what the Laplacian is in polar coordinates. Normally we would find u_{xx} and u_{yy} in terms of the derivatives in r and θ. We would need to solve for r and θ in terms of x and y. While this is certainly possible, it happens to be more convenient to work in reverse. Let us instead compute derivatives in r and θ in terms of derivatives in x and y and then solve. The computations are easier this way. First

$$\begin{aligned} x_r &= \cos\theta, & x_\theta &= -r\sin\theta, \\ y_r &= \sin\theta, & y_\theta &= r\cos\theta. \end{aligned} \tag{15.10.3}$$

Next by chain rule we obtain

$$\begin{aligned} u_r &= u_x x_r + u_y y_r = \cos(\theta)u_x + \sin(\theta)u_y, \\ u_{rr} &= \cos(\theta)(u_{xx}x_r + u_{xy}y_r) + \sin(\theta)(u_{yx}x_r + u_{yy}y_r) \\ &= \cos^2(\theta)u_{xx} + 2\cos(\theta)\sin(\theta)u_{xy} + \sin^2(\theta)u_{yy}. \end{aligned} \tag{15.10.4}$$

Similarly for the θ derivative. Note that we have to use product rule for the second derivative.

$$\begin{aligned} u_\theta &= u_x x_\theta + u_y y_\theta = -r\sin(\theta)u_x + r\cos(\theta)u_y, \\ u_{\theta\theta} &= -r\cos(\theta)(u_x) - r\sin(\theta)(u_{xx}x_\theta + u_{xy}y_\theta) - r\sin(\theta)(u_y) + r\cos(\theta)(u_{yx}x_\theta + u_{yy}y_\theta) \\ &= -r\cos(\theta)u_x - r\sin(\theta)u_y + r^2\sin^2(\theta)u_{xx} - r^2 2\sin(\theta)\cos(\theta)u_{xy} + r^2\cos^2(\theta)u_{yy}. \end{aligned} \tag{15.10.5}$$

Let us now try to solve for $u_{xx}+u_{yy}$. We start with $\frac{1}{r^2}u_{\theta\theta}$ to get rid of those pesky r^2. If we add u_{rr} and use the fact that $\cos^2(\theta)+\sin^2(\theta)=1$, we get

$$\frac{1}{r^2}u_{\theta\theta}+u_{rr}=u_{xx}+u_{yy}-\frac{1}{r}\cos(\theta)u_x-\frac{1}{r}\sin(\theta)u_y.$$

We're not quite there yet, but all we are lacking is $\frac{1}{r}u_r$. Adding it we obtain the *Laplacian in polar coordinates*:

$$\frac{1}{r^2}u_{\theta\theta}+\frac{1}{r}u_r+u_{rr}=u_{xx}+u_{yy}=\Delta u.$$

Notice that the Laplacian in polar coordinates no longer has constant coefficients.

15.10.2: Series Solution

Let us separate variables as usual. That is let us try $u(r,\theta)=R(r)\Theta(\theta)$. Then

$$0=\Delta u=\frac{1}{r^2}R\Theta''+\frac{1}{r}R'\Theta+R''\Theta.$$

Let us put R on one side and Θ on the other and conclude that both sides must be constant.

$$\begin{aligned}\frac{1}{r^2}R\Theta'' &= -\left(\frac{1}{r}R'+R''\right)\Theta.\\ \frac{\Theta''}{\Theta} &= -\frac{rR'+r^2R''}{R}+-\lambda.\end{aligned} \tag{15.10.6}$$

We get two equations:

$$\begin{aligned}\Theta''+\lambda\Theta &= 0,\\ r^2R''+rR'-\lambda R &= 0.\end{aligned} \tag{15.10.7}$$

Let us first focus on Θ. We know that $u(r,\theta)$ ought to be 2π-periodic in θ, that is, $u(r,\theta)=u(r,\theta+2\pi)$. Therefore, the solution to $\Theta''+\lambda\Theta=0$ must be 2π-periodic. We have seen such a problem in Example 4.1.5. We conclude that $\lambda=n^2$ for a nonnegative integer $n=0,1,2,3,\ldots$. The equation becomes $\Theta''+n^2\Theta=0$. When $n=0$ the equation is just $\Theta''=0$, so we have the general solution $A\theta+B$. As Θ is periodic, $A=0$. For convenience let us write this solution as

$$\Theta_0=\frac{a_0}{2}$$

for some constant a_0. For positive n, the solution to $\Theta''+n^2\Theta=0$ is

$$\Theta_n=a_n\cos(n\theta)+b_n\sin(n\theta),$$

for some constants a_n and b_n.

Next, we consider the equation for R,

$$r^2R''+rR'-n^2R=0.$$

This equation has appeared in exercises before—we solved it in Exercise 2.E.2.1.6 and Exercise 2.E.1.7. The idea is to try a solution r^s and if that does not work out try a solution of the form $r^s\ln r$. When $n=0$ we obtain

$$R_0=Ar^0+Br^0\ln r=A+B\ln r,$$

and if $n>0$, we get

$$R_n=Ar^n+Br^{-n}.$$

The function $u(r,\theta)$ must be finite at the origin, that is, when $r=0$. Therefore, $B=0$ in both cases. Let us set $A=1$ in both cases as well, the constants in Θ_n will pick up the slack so we do not lose anything. Therefore let

$$R_0=1,\quad \text{and}\quad R_n=r^n.$$

Hence our building block solutions are

$$u_0(r,\theta) = \frac{a_0}{2}, \quad u_n(r,\theta) = a_n r^n \cos(n\theta) + b_n r^n \sin(n\theta).$$

Putting everything together our solution is:

$$u(r,\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n r^n \cos(n\theta) + b_n r^n \sin(n\theta).$$

We look at the boundary condition in (15.10.1),

$$g(\theta) = u(1,\theta) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(n\theta) + b_n \sin(n\theta).$$

Therefore, the solution (15.10.1)is to expand $g(\theta)$, which is a 2π-periodic function as a Fourier series, and then the nth coordinate is multiplied by r^n. In other words, to compute a^n and b^n from the formula we can, as usual, compute

$$a_n = \frac{1}{\pi}\int_{\pi}^{-\pi} g(\theta)\cos(n\theta)d\theta, \quad \text{and} \quad b_n = \frac{1}{\pi}\int_{\pi}^{-\pi} g(\theta)\sin(n\theta)d\theta.$$

Example 15.10.1

Suppose we wish to solve

$$\begin{aligned} &\Delta u = 0, \qquad 0 \le r < 1, \quad -\pi < \theta \le \pi, \\ &u(1,\theta) = \cos(10\,\theta), \qquad -\pi < \theta \le \pi. \end{aligned} \tag{15.10.8}$$

The solution is

$$u(r,\theta) = r^{10}\cos(10\theta).$$

See the plot in Figure 15.10.3 The thing to notice in this example is that the effect of a high frequency is mostly felt at the boundary. In the middle of the disc, the solution is very close to zero. That is because r^{10} rather small when r is close to 0.

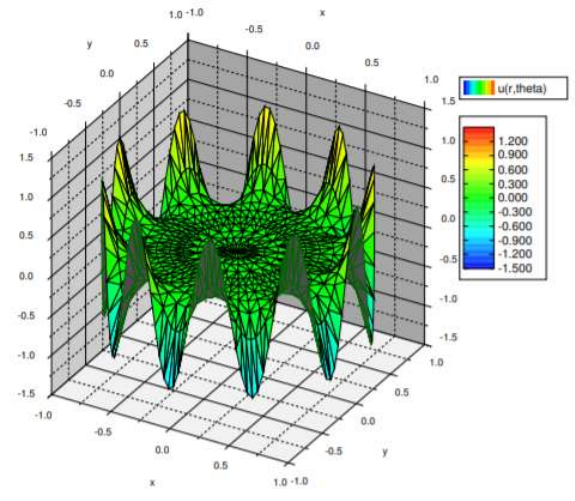


Figure 15.10.3: The solution of the Dirichlet problem in the disc with $\cos(10\theta)$ as boundary data.

Example 15.10.2

Let us solve a more difficult problem. Suppose we have a long rod with circular cross section of radius 1 and we wish to solve the steady state heat problem. If the rod is long enough we simply need to solve the Laplace equation in two dimensions. Let us put the center of the rod at the origin and we have exactly the region we are currently studying—a circle of radius 1. For the boundary conditions, suppose in Cartesian coordinates x and y, the temperature is fixed at 0 when $y < 0$ and at $2y$ when $y > 0$.

We set the problem up. As $y = r\sin(\theta)$, then on the circle of radius 1 we have $2y = 2\sin(\theta)$. So

$$\begin{aligned} &\Delta u = 0, \qquad 0 \le r < 1, \quad -\pi < \theta \le \pi, \\ &u(1,\theta) = \begin{cases} 2\sin(\theta) & \text{if} \quad 0 \le \theta \le \pi, \\ 0 & \text{if} \ -\pi < \theta < 0. \end{cases} \end{aligned} \tag{15.10.9}$$

We must now compute the Fourier series for the boundary condition. By now the reader has plentiful experience in computing Fourier series and so we simply state that

$$u(1,\theta) = \frac{2}{\pi} + \sin(\theta) + \sum_{n=1}^{\infty} \frac{-4}{\pi(4n^2-1)} \cos(2n\theta).$$

? Exercise 15.10.1

Compute the series for $u(1,\theta)$ and verify that it really is what we have just claimed. Hint: Be careful, make sure not to divide by zero.

We now simply write the solution (see Figure 15.10.4) by multiplying by r^n in the right places.

$$u(r,\theta) = \frac{2}{\pi} + r\sin(\theta) + \sum_{n=1}^{\infty} \frac{-4r^{2n}}{\pi(4n^2-1)} \cos(2n\theta).$$

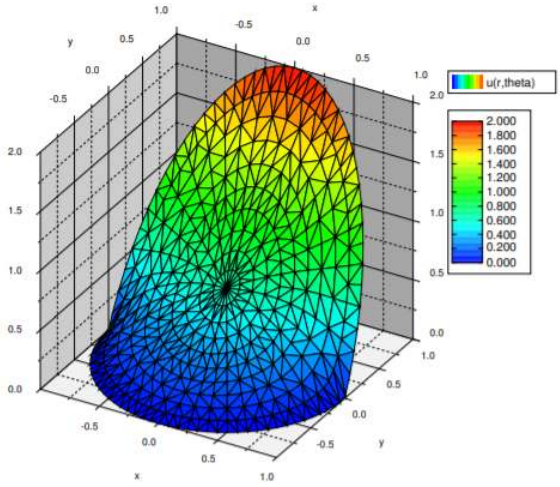


Figure 15.10.4: The solution of the Dirichlet problem with boundary data 0 for $y < 0$ and $2y$ for $y > 0$.

15.10.3: Poisson Kernel

There is another way to solve the Dirichlet problem with the help of an integral kernel. That is, we will find a function $P(r,\theta,\alpha)$ called the *Poisson kernel*[1] such that

$$u(r,\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P(r,\theta,\alpha) g(\alpha) d\alpha.$$

While the integral will generally not be solvable analytically, it can be evaluated numerically. In fact, unless the boundary data is given as a Fourier series already, it will be much easier to numerically evaluate this formula as there is only one integral to evaluate.

The formula also has theoretical applications. For instance, as $P(r,\theta,\alpha)$ will have infinitely many derivatives, then via differentiating under the integral we find that the solution $u(r,\theta)$ has infinitely many derivatives, at least when inside the circle, $r < 1$. By infinitely many derivatives what you should think of is that $u(r,\theta)$ has "no corners" and all of its partial derivatives exist too and also have "no corners".

We will compute the formula for $P(r,\theta,\alpha)$ from the series solution, and this idea can be applied anytime you have a convenient series solution where the coefficients are obtained via integration. Hence you can apply this reasoning to obtain such integral

kernels for other equations, such as the heat equation. The computation is long and tedious, but not overly difficult. Since the ideas are often applied in similar contexts, it is good to understand how this computation works.

What we do is start with the series solution and replace the coefficients with the integrals that compute them. Then we try to write everything as a single integral. We must use a different dummy variable for the integration and hence we use α instead of θ.

$$\begin{aligned} u(r,\theta) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n r^n \cos(n\theta) + b_n r^n \sin(n\theta) \\ &= \underbrace{\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} g(\alpha)d\alpha\right)}_{\frac{a_0}{2}} + \sum_{n=1}^{\infty} \underbrace{\left(\frac{1}{\pi}\int_{-\pi}^{\pi} g(\alpha)\cos(n\alpha)d\alpha\right)}_{a_n} r^n\cos(n\theta) \\ &+ \underbrace{\left(\frac{1}{\pi}\int_{-\pi}^{\pi} g(\alpha)\sin(n\alpha)d\alpha\right)}_{b_n} r^n \sin(n\theta) \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(g(\alpha) + 2\sum_{n=1}^{\infty} g(\alpha)\cos(n\alpha)r^n\cos(n\theta) + g(\alpha)\sin(n\alpha)r^n\sin(n\theta)\right) d\alpha \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \underbrace{\left(1 + 2\sum_{n=1}^{\infty} r^n\left(\cos(n\alpha)\cos(n\theta) + \sin(n\alpha)\sin(n\theta)\right)\right)}_{P(r,\theta,\alpha)} g(\alpha)d\alpha \end{aligned} \tag{15.10.10}$$

OK, so we have what we wanted, the expression in the parentheses is the Poisson kernel, $P(r,\theta,\alpha)$. However, we can do a lot better. It is still given as a series, and we would really like to have a nice simple expression for it. We must work a little harder. The trick is to rewrite everything in terms of complex exponentials. Let us work just on the kernel.

$$\begin{aligned} P(r,\theta,\alpha) &= 1 + 2\sum_{n=1}^{\infty} r^n(\cos(n\alpha)\cos(n\theta) + \sin(n\alpha)\sin(n\theta)) \\ &= 1 + 2\sum_{n=1}^{\infty} r^n \cos(n(\theta-\alpha)) \\ &= 1 + 2\sum_{n=1}^{\infty} r^n (e^{in(\theta-\alpha)} + e^{-in(\theta-\alpha)}) \\ &= 1 + \sum_{n=1}^{\infty} (re^{i(\theta-\alpha)})^n + \sum_{n=1}^{\infty} (re^{-i(\theta-\alpha)})^n. \end{aligned} \tag{15.10.11}$$

In the above expression we recognize the *geometric series*. That is, recall from calculus that as long as $|z| < 1$, then

$$\sum_{n=1}^{\infty} z^n = \frac{z}{1-z}.$$

Note that n starts at 1 and that is why we have the z in the numerator. It is the standard geometric series multiplied by z. Let us continue with the computation.

$$\begin{aligned} P(r,\theta,\alpha) &= 1 + \sum_{n=1}^{\infty} (re^{i(\theta-\alpha)})^n + \sum_{n=1}^{\infty} (re^{-i(\theta-\alpha)})^n \\ &= 1 + \frac{re^{i(\theta-\alpha)}}{1-re^{i(\theta-\alpha)}} + \frac{re^{-i(\theta-\alpha)}}{1-re^{-i(\theta-\alpha)}} \\ &= \frac{(1-re^{i(\theta-\alpha)})(1-re^{-i(\theta-\alpha)}) + (1-re^{-i(\theta-\alpha)})re^{i(\theta-\alpha)} + (1-re^{i(\theta-\alpha)})re^{-i(\theta-\alpha)}}{(1-re^{i(\theta-\alpha)})(1-re^{-i(\theta-\alpha)})} \\ &= \frac{1-r^2}{1-re^{i(\theta-\alpha)} - re^{-i(\theta-\alpha)} + r^2} \\ &= \frac{1-r^2}{1-2r\cos(\theta-\alpha)+r^2}. \end{aligned} \tag{15.10.12}$$

Now that's a formula we can live with. The solution to the Dirichlet problem using the Poisson kernel is

$$u(r,\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1-r^2}{1-2r\cos(\theta-\alpha)+r^2} g(\alpha)d\alpha.$$

Sometimes the formula for the Poisson kernel is given together with the constant $\frac{1}{2\pi}$, in which case we should of course not leave it in front of the integral. Also, often the limits of the integral are given as 0 to 2π; everything inside is 2π-periodic in α, so this does not change the integral.

Let us not leave the Poisson kernel without explaining its geometric meaning. Let s be the distance from (r,θ) to $(1,\alpha)$. You may recall from calculus that this distance s in polar coordinates is given precisely by the square root of $1-2r\cos(\theta-\alpha)+r^2$. That is, the Poisson kernel is really the formula

$$\frac{1-r^2}{s^2}.$$

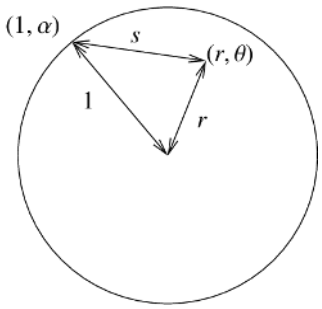


Figure 15.10.5

One final note we make about the formula is to note that it is really a weighted average of the boundary values. First let us look at what happens at the origin, that is when $r=0$.

$$\begin{aligned} u(0,0) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1-0^2}{1-2(0)\cos(\theta-\alpha)+0^2} g(\alpha)d\alpha \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} g(\alpha)d\alpha. \end{aligned} \tag{15.10.13}$$

So $u(0,0)$ is precisely the average value of $g(\theta)$ and therefore the average value of u on the boundary. This is a general feature of harmonic functions, the value at some point p is equal to the average of the values on a circle centered at p.

What the formula says is that the value of the solution at any point in the circle is a weighted average of the boundary data $g(\theta)$. The kernel is bigger when (r,θ) is closer to $(1,\alpha)$. Therefore when computing $u(r,\theta)$ we give more weight to the values $g(\alpha)$ when $(1,\alpha)$ is closer to (r,θ) and less weight to the values $g(\theta)$ when $(1,\alpha)$ far from (r,θ).

15.10.4: Footnotes

[1] Named for the French mathematician Siméon Denis Poisson (1781 – 1840).

This page titled 15.10: Dirichlet Problem in the Circle and the Poisson Kernel is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.10: Dirichlet Problem in the Circle and the Poisson Kernel** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

15.E: Fourier Series and PDEs (Exercises)

These are homework exercises to accompany Libl's "Differential Equations for Engineering" Textmap. This is a textbook targeted for a one semester first course on differential equations, aimed at engineering students. Prerequisite for the course is the basic calculus sequence.

15.E.1: 4.1: Boundary value problems

Hint for the following exercises: Note that when $\lambda > 0$, then $\cos(\sqrt{\lambda}(t-a))$ and $\sin(\sqrt{\lambda}(t-a))$ are also solutions of the homogeneous equation.

Exercise 15.E.4.1.1

Compute all eigenvalues and eigenfunctions of $x'' + \lambda x = 0, x(a) = 0, x(b) = 0$ (assume $a < b$).

Exercise 15.E.4.1.2

Compute all eigenvalues and eigenfunctions of $x'' + \lambda x = 0, x'(a) = 0, x'(b) = 0$ (assume $a < b$).

Exercise 15.E.4.1.3

Compute all eigenvalues and eigenfunctions of $x'' + \lambda x = 0, x'(a) = 0, x(b) = 0$ (assume $a < b$).

Exercise 15.E.4.1.4

Compute all eigenvalues and eigenfunctions of $x'' + \lambda x = 0, x(a) = x(b), x'(a) = x'(b)$ (assume $a < b$).

Exercise 15.E.4.1.5

We have skipped the case of $\lambda < 0$ for the boundary value problem $x'' + \lambda x = 0, x(-\pi) = x(\pi), x'(-\pi) = x'(\pi)$. Finish the calculation and show that there are no negative eigenvalues.

Exercise 15.E.4.1.6

Consider a spinning string of length 2 and linear density 0.1 and tension 3. Find smallest angular velocity when the string pops out.

Answer

$\omega = \pi\sqrt{\frac{15}{2}}$

Exercise 15.E.4.1.7

Suppose $x'' + \lambda x = 0$ and $x(0) = 1, x(1) = 1$. Find all λ for which there is more than one solution. Also find the corresponding solutions (only for the eigenvalues).

Answer

$\lambda_k = 4k^2\pi^2$ for $k = 1,\ 2,\ 3, \ldots$ $x_k = \cos(2k\pi t) + B\sin(2k\pi t)$ (for any B)

Exercise 15.E.4.1.8

Suppose $x'' + x = 0$ and $x(0) = 0, x'(\pi) = 1$. Find all the solution(s) if any exist.

Answer

$x(t) = -\sin(t)$

Exercise 15.E.4.1.9

Consider $x' + \lambda x = 0$ and $x(0) = 0, x(1) = 0$. Why does it not have any eigenvalues? Why does any first order equation with two endpoint conditions such as above have no eigenvalues?

Answer

General solution is $x = Ce^{-\lambda t}$. Since $x(0) = 0$ then $C = 0$, and so $x(t) = 0$. Therefore, the solution is always identically zero. One condition is always enough to guarantee a unique solution for a first order equation.

Exercise 15.E.4.1.10: (challenging)

Suppose $x''' + \lambda x = 0$ and $x(0) = 0, x'(0) = 0, x(1) = 0$. Suppose that $\lambda > 0$. Find an equation that all such eigenvalues must satisfy. Hint: Note that $-\sqrt[3]{\lambda}$ is a root of $r^3 + \lambda = 0$.

Answer

$$\frac{\sqrt{3}}{3} e^{\frac{-3}{2}\sqrt[3]{\lambda}} - \frac{\sqrt{3}}{3}\cos\left(\frac{\sqrt{3}\sqrt[3]{\lambda}}{2}\right) + \sin\left(\frac{\sqrt{3}\sqrt[3]{\lambda}}{2}\right) = 0$$

15.E.2: 4.2: The Trigonometric Series

Exercise 15.E.4.2.1

Suppose $f(t)$ is defined on $[-\pi, \pi]$ as $\sin(5t) + \cos(3t)$. Extend periodically and compute the Fourier series of $f(t)$.

Exercise 15.E.4.2.2

Suppose $f(t)$ is defined on $[-\pi, \pi]$ as $|t|$. Extend periodically and compute the Fourier series of $f(t)$.

Exercise 15.E.4.2.3

Suppose $f(t)$ is defined on $[-\pi, \pi]$ as $|t|^3$. Extend periodically and compute the Fourier series of $f(t)$.

Exercise 15.E.4.2.4

Suppose $f(t)$ is defined on $(-\pi, \pi]$ as

$$f(t) = \begin{cases} -1 & if -\pi < t \leq 0, \\ 1 & if 0 < t \leq \pi. \end{cases} \tag{15.E.1}$$

Extend periodically and compute the Fourier series of $f(t)$.

Exercise 15.E.4.2.5

Suppose $f(t)$ is defined on $(-\pi, \pi]$ as t^3. Extend periodically and compute the Fourier series of $f(t)$.

Exercise 15.E.4.2.6

Suppose $f(t)$ is defined on $[-\pi, \pi]$ as t^2. Extend periodically and compute the Fourier series of $f(t)$.

There is another form of the Fourier series using complex exponentials e^{nt} for $n = \ldots, -2, -1, 0, 1, 2, \ldots$ instead of $\cos(nt)$ and $\sin(nt)$ for positive n. This form may be easier to work with sometimes. It is certainly more compact to write, and there is

only one formula for the coefficients. On the downside, the coefficients are complex numbers.

Exercise 15.E.4.2.7

Let

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt) + b_n \sin(nt). \tag{15.E.2}$$

Use Euler's formula $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ to show that there exist complex numbers c_m such that

$$f(t) = \sum_{m=-\infty}^{\infty} c_m e^{imt}. \tag{15.E.3}$$

Note that the sum now ranges over all the integers including negative ones. Do not worry about convergence in this calculation. Hint: It may be better to start from the complex exponential form and write the series as

$$c_0 + \sum_{m=1}^{\infty} c_m e^{imt} + c_{-m} e^{-imt}. \tag{15.E.4}$$

Exercise 15.E.4.2.8

Suppose $f(t)$ is defined on $[-\pi, \pi]$ as $f(t) = \sin(t)$. Extend periodically and compute the Fourier series.

Answer

$\sin(t)$

Exercise 15.E.4.2.9

Suppose $f(t)$ is defined on $(-\pi, \pi]$ as $f(t) = \sin(\pi t)$. Extend periodically and compute the Fourier series.

Answer

$\sum_{n=1}^{\infty} \frac{(\pi - n)\sin(\pi n + \pi^2) + (\pi + n)\sin(\pi n - \pi^2)}{\pi n^2 - \pi^3} \sin(nt)$

Exercise 15.E.4.2.10

Suppose $f(t)$ is defined on $(-\pi, \pi]$ as $f(t) = \sin^2(t)$. Extend periodically and compute the Fourier series.

Answer

$\frac{1}{2} - \frac{1}{2}\cos(2t)$

Exercise 15.E.4.2.11

Suppose $f(t)$ is defined on $(-\pi, \pi]$ as $f(t) = t^4$. Extend periodically and compute the Fourier series.

Answer

$\frac{\pi^4}{5} + \sum_{n=1}^{\infty} \frac{(-1)^n (8\pi^2 n^2 - 48)}{n^4} \cos(nt)$

15.E.3: 4.3: More on the Fourier series

Exercise 15.E.4.3.1

Let

$$f(t)=\begin{cases}0 & if \quad -1<t\leq 0,\\ t & if \quad 0<t\leq 1,\end{cases} \tag{15.E.5}$$

extended periodically.

a. Compute the Fourier series for $f(t)$.
b. Write out the series explicitly up to the 3rd harmonic.

Exercise 15.E.4.3.2

Let

$$f(t)=\begin{cases}-t & if \quad -1<t\leq 0,\\ t^2 & if \quad 0<t\leq 1,\end{cases} \tag{15.E.6}$$

extended periodically.

a. Compute the Fourier series for $f(t)$.
b. Write out the series explicitly up to the 3rd harmonic.

Exercise 15.E.4.3.3

Let

$$f(t)=\begin{cases}\dfrac{-t}{10} & if \quad -10<t\leq 0,\\ \dfrac{t}{10} & if \quad 0<t\leq 10,\end{cases} \tag{15.E.7}$$

extended periodically (period is 20).

a. Compute the Fourier series for $f(t)$.
b. Write out the series explicitly up to the 3rd harmonic.

Exercise 15.E.4.3.4

Let $f(t)=\sum_{n=1}^{\infty}\dfrac{1}{n^3}\cos(nt)$. Is $f(t)$ continuous and differentiable everywhere? Find the derivative (if it exists everywhere) or justify why $f(t)$ is not differentiable everywhere.

Exercise 15.E.4.3.5

Let $f(t)=\sum_{n=1}^{\infty}\dfrac{(-1)^n}{n}\sin(nt)$. Is $f(t)$ differentiable everywhere? Find the derivative (if it exists everywhere) or justify why $f(t)$ is not differentiable everywhere.

Exercise 15.E.4.3.6

Let

$$f(t)=\begin{cases}0 & if -2<t\leq 0,\\ t & if\, 0<t\leq 1,\\ -t+2 & if\, 1<t\leq 2,\end{cases} \tag{15.E.8}$$

extended periodically.

a. Compute the Fourier series for $f(t)$.
b. Write out the series explicitly up to the 3rd harmonic.

Exercise 15.E.4.3.7

Let

$$f(t) = e^t \quad for\ -1 < t \leq 1 \tag{15.E.9}$$

extended periodically.

a. Compute the Fourier series for $f(t)$.
b. Write out the series explicitly up to the 3rd harmonic.
c. What does the series converge to at $t = 1$.

Exercise 15.E.4.3.8

Let

$$f(t) = t^2 \quad for\ -1 < t \leq 1 \tag{15.E.10}$$

extended periodically.

a. Compute the Fourier series for $f(t)$.
b. By plugging in $t = 0$, evaluate $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} = 1 - \frac{1}{4} + \frac{1}{9} - \cdots$.
c. Now evaluate $\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \cdots$.

Exercise 15.E.4.3.9

Let

$$f(t) = \begin{cases} 0 & \text{if} \quad -3 < t \leq 0, \\ t & \text{if} \quad 0 < t \leq 3, \end{cases} \tag{15.E.11}$$

extended periodically. Suppose $F(t)$ is the function given by the Fourier series of f. Without computing the Fourier series evaluate.

a. $F(2)$
b. $F(-2)$
c. $F(4)$
d. $F(-4)$
e. $F(3)$
f. $F(-9)$

Exercise 15.E.4.3.10

Let

$$f(t) = t^2 \quad for\ -2 < t \leq 2 \tag{15.E.12}$$

extended periodically.

a. Compute the Fourier series for $f(t)$.
b. Write out the series explicitly up to the 3rd harmonic.

Answer

a. $\frac{8}{6} + \sum_{n=1}^{\infty} \frac{16(-1)^n}{\pi^2 n^2} \cos\left(\frac{n\pi}{2} t\right)$

b. $\frac{8}{6}-\frac{16}{\pi^2}\cos\left(\frac{\pi}{2}t\right)+\frac{4}{\pi^2}\cos(\pi t)-\frac{16}{9\pi^2}\cos\left(\frac{3\pi}{2}t\right)+\cdots$

Exercise 15.E.4.3.11

Let

$$f(t)=t \quad \text{for } \lambda<t\leq\lambda \text{ (for some } \lambda) \tag{15.E.13}$$

extended periodically.

a. Compute the Fourier series for $f(t)$.
b. Write out the series explicitly up to the 3rd harmonic.

Answer

a. $\sum_{n=1}^{\infty}\frac{(-1)^{n+1}2\lambda}{n\pi}\sin\left(\frac{n\pi}{\lambda}t\right)$
b. $\frac{2\lambda}{\pi}\sin\left(\frac{\pi}{\lambda}t\right)-\frac{\lambda}{\pi}\sin\left(\frac{2\pi}{\lambda}t\right)+\frac{2\lambda}{3\pi}\sin\left(\frac{3\pi}{\lambda}t\right)-\cdots$

Exercise 15.E.4.3.12

Let

$$f(t)=\frac{1}{2}+\sum_{n=1}^{\infty}\frac{1}{n(n^2+1)}\sin(n\pi t). \tag{15.E.14}$$

Compute $f'(t)$.

Answer

$f'(t)=\sum_{n=1}^{\infty}\frac{\pi}{n^2+1}\cos(n\pi t)$

Exercise 15.E.4.3.13

Let

$$f(t)=\frac{1}{2}+\sum_{n=1}^{\infty}\frac{1}{n^3)}\cos(nt). \tag{15.E.15}$$

a. Find the antiderivative.
b. Is the antiderivative periodic?

Answer

a. $F(t)=\frac{t}{2}+C+\sum_{n=1}^{\infty}\frac{1}{n^4}\sin(nt)$
b. no

Exercise 15.E.4.3.14

Let

$$f(t)=\frac{t}{2} \quad \text{for } -\pi<t\leq\pi \tag{15.E.16}$$

extended periodically.

a. Compute the Fourier series for $f(t)$.
b. Plug in $t=\frac{\pi}{2}$ to find a series representation for $\frac{\pi}{4}$.

c. Using the first 4 terms of the result from part b) approximate $\frac{\pi}{4}$.

Answer

a. $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nt)$
b. f is continuous at $t = \frac{\pi}{2}$ so the Fourier series converges to $f\left(\frac{\pi}{2}\right) = \frac{\pi}{4}$. Obtain
$\frac{\pi}{4} = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{2n-1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$.
c. Using the first 4 terms get $76/105 \approx 0.72$ (quite a bad approximation, you would have to take about 50 terms to start to get to within 0.01 of $\frac{\pi}{4}$).

Exercise 15.E.4.3.15

Let

$$f(t) = \begin{cases} 0 & \text{if} \quad -2 < t \leq 0, \\ 2 & \text{if} \quad 0 < t \leq 2, \end{cases} \tag{15.E.17}$$

extended periodically. Suppose $F(t)$ is the function given by the Fourier series of f. Without computing the Fourier series evaluate.

a. $F(0)$
b. $F(-1)$
c. $F(1)$
d. $F(-2)$
e. $F(4)$
f. $F(-8)$

Answer

a. $F(0) = 1$
b. $F(-1) = 0$
c. $F(1) = 2$
d. $F(-2) = 1$
e. $F(4) = 1$
f. $F(-9) = 0$

15.E.4: 4.4: Sine and Cosine Series

Exercise 15.E.4.4.1

Take $f(t) = (t-1)^2$ defined on $0 \leq t \leq 1$.

a. Sketch the plot of the even periodic extension of f.
b. Sketch the plot of the odd periodic extension of f.

Exercise 15.E.4.4.2

Find the Fourier series of both the odd and even periodic extension of the function $f(t) = (t-1)^2$ for $0 \leq t \leq 1$. Can you tell which extension is continuous from the Fourier series coefficients?

Exercise 15.E.4.4.3

Find the Fourier series of both the odd and even periodic extension of the function $f(t) = t$ for $0 \leq t \leq \pi$.

Exercise 15.E.4.4.4

Find the Fourier series of the even periodic extension of the function $f(t) = \sin t$ for $0 \le t \le \pi$.

Exercise 15.E.4.4.5

Consider

$$x''(t) + 4x(t) = f(t), \tag{15.E.18}$$

where $f(t) = 1$ on $0 < t < 1$.

a. Solve for the Dirichlet conditions $x(0) = 0$, $x(1) = 0$.
b. Solve for the Neumann conditions $x'(0) = 0$, $x'(1) = 0$.

Exercise 15.E.4.4.6

Consider

$$x''(t) + 9x(t) = f(t), \tag{15.E.19}$$

for $f(t) = \sin(2\pi t)$ on $0 < t < 1$.

a. Solve for the Dirichlet conditions $x(0) = 0,\ x(1) = 0$.
b. b) Solve for the Neumann conditions $x'(0) = 0,\ x'(1) = 0$.

Exercise 15.E.4.4.7

Consider

$$x''(t) + 3x(t) = f(t), \quad x(0) = 0, \quad x(1) = 0, \tag{15.E.20}$$

where $f(t) = \sum_{n=1}^{\infty} b_n \sin(n\pi t)$. Write the solution $x(t)$ as a Fourier series, where the coefficients are given in terms of b_n.

Exercise 15.E.4.4.8

Let $f(t) = t^2(2-t)$ for $0 \le t \le 2$. Let $F(t)$ be the odd periodic extension. Compute $F(1),\ F(2),\ F(3),\ F(-1),\ F(\frac{9}{2}),\ F(101),\ F(103)$. Note: Do **not** compute using the sine series.

Exercise 15.E.4.4.9

Let $f(t) = \frac{t}{3}$ on $0 \le t < 3$.

a. Find the Fourier series of the even periodic extension.
b. Find the Fourier series of the odd periodic extension.

Answer

a. $\frac{1}{2} + \sum\limits_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{-4}{\pi^2 n^2} \cos\left(\frac{n\pi}{3} t\right)$
b. $\sum\limits_{n=1}^{\infty} \frac{2(-1)^{n+1}}{\pi n} \sin\left(\frac{n\pi}{3} t\right)$

Exercise 15.E.4.4.10

Let $f(t) = \cos(2t)$ on $0 \le t < \pi$.

a. Find the Fourier series of the even periodic extension.
b. Find the Fourier series of the odd periodic extension.

Answer

a. $\cos(2t)$

b. $\sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{-4n}{\pi n^2 - 4\pi} \sin(nt)$

Exercise 15.E.4.4.11

Let $f(t)$ be defined on $0 \le t < 1$. Now take the average of the two extensions $g(t) = \frac{F_{odd}(t) + F_{even}(t)}{2}$.

a. What is $g(t)$ if $0 \le t < 1$ (Justify!)

b. What is $g(t)$ if $-1 < t < 0$ (Justify!)

Answer

a. $f(t)$

b. 0

Exercise 15.E.4.4.12

Let $f(t) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sin(nt)$. Solve $x'' - x = f(t)$ for the Dirichlet conditions $x(0) = 0$ and $x(\pi) = 0$.

Answer

$\sum_{n=1}^{\infty} \frac{-1}{n^2(1+n^2)} \sin(nt)$

Exercise 15.E.4.4.13: (challenging)

Let $f(t) = t + \sum_{n=1}^{\infty} \frac{1}{2^n} \sin(nt)$. Solve $x'' + \pi x = f(t)$ for the Dirichlet conditions $x(0) = 0$ and $x(\pi) = 1$. Hint: Note that $\frac{t}{\pi}$ satisfies the given Dirichlet conditions.

Answer

$\frac{t}{\pi} + \sum_{n=1}^{\infty} \frac{1}{2^n(\pi - n^2)} \sin(nt)$

15.E.5: 4.5: Applications of Fourier series

Exercise 15.E.4.5.1

Let $F(t) = \frac{1}{2} + \sum_{n=1}^{\infty} \frac{1}{n^2} \cos(n\pi t)$. Find the steady periodic solution to $x'' + 2x = F(t)$. Express your solution as a Fourier series.

Exercise 15.E.4.5.2

Let $F(t) = \sum_{n=1}^{\infty} \frac{1}{n^3} \sin(n\pi t)$. Find the steady periodic solution to $x'' + x' + x = F(t)$. Express your solution as a Fourier series.

Exercise 15.E.4.5.3

Let $F(t) = \sum_{n=1}^{\infty} \frac{1}{n^2} \cos(n\pi t)$. Find the steady periodic solution to $x'' + 4x = F(t)$. Express your solution as a Fourier series.

Exercise 15.E.4.5.4

Let $F(t)=t$ for $-1<t<1$ and extended periodically. Find the steady periodic solution to $x''+x=F(t)$. Express your solution as a series.

Exercise 15.E.4.5.5

Let $F(t)=t$ for $-1<t<1$ and extended periodically. Find the steady periodic solution to $x''+\pi^2 x=F(t)$. Express your solution as a series.

Exercise 15.E.4.5.6

Let $F(t)=\sin(2\pi t)+0.1\cos(10\pi t)$. Find the steady periodic solution to $x''+\sqrt{2}x=F(t)$. Express your solution as a Fourier series.

Answer

$$x=\frac{1}{\sqrt{2}-4\pi^2}\sin(2\pi t)+\frac{0.1}{\sqrt{2}-100\pi^2}\cos(10\pi t)$$

Exercise 15.E.4.5.7

Let $F(t)=\sum_{n=1}^{\infty}e^{-n}\cos(2nt)$. Find the steady periodic solution to $x''+3x=F(t)$. Express your solution as a Fourier series.

Answer

$$x=\sum_{n=1}^{\infty}\frac{e^{-n}}{3-(2n)^2}\cos(2nt)$$

Exercise 15.E.4.5.8

Let $F(t)=|t|$ for $-1\le t\le 1$ extended periodically. Find the steady periodic solution to $x''+\sqrt{3}x=F(t)$. Express your solution as a series.

Answer

$$x=\frac{1}{2\sqrt{3}}+\sum_{\substack{n=1\\ n\text{ odd}}}^{\infty}\frac{-4}{n^2\pi^2(\sqrt{3}-n^2\pi^2)}\cos(n\pi t)$$

Exercise 15.E.4.5.9

Let $F(t)=|t|$ for $-1\le t\le 1$ extended periodically. Find the steady periodic solution to $x''+\pi^2 x=F(t)$. Express your solution as a series.

Answer

$$x=\frac{1}{2\sqrt{3}}-\frac{2}{\pi^3}t\sin(\pi t)+\sum_{\substack{n=3\\ n\text{ odd}}}^{\infty}\frac{-4}{n^2\pi^4(1-n^2)}\cos(n\pi t)$$

15.E.6: 4.6: PDEs, Separation of Variables, and the Heat Equation

Exercise 15.E.4.6.1

Imagine you have a wire of length 2, with $k = 0.001$ and an initial temperature distribution of $u(x,0) = 50x$. Suppose that both the ends are embedded in ice (temperature 0). Find the solution as a series.

Exercise 15.E.4.6.2

Find a series solution of

$$\begin{aligned} u_t &= u_{xx}, \\ u(0,t) &= u(1,t) = 0, \\ u(x,0) &= 100 \quad \text{for } 0 < x < 1. \end{aligned} \tag{15.E.21}$$

Exercise 15.E.4.6.3

Find a series solution of

$$\begin{aligned} u_t &= u_{xx}, \\ u_x(0,t) &= u_x(\pi,t) = 0, \\ u(x,0) &= 3\cos(x) + \cos(3x) \quad \text{for } 0 < x < \pi. \end{aligned} \tag{15.E.22}$$

Exercise 15.E.4.6.4

Find a series solution of

$$\begin{aligned} u_t &= \frac{1}{3} u_{xx}, \\ u_x(0,t) &= u_x(\pi,t) = 0, \\ u(x,0) &= \frac{10x}{\pi} \quad \text{for } 0 < x < \pi. \end{aligned} \tag{15.E.23}$$

Exercise 15.E.4.6.5

Find a series solution of

$$\begin{aligned} u_t &= u_{xx}, \\ u(0,t) &= 0, \quad u(1,t) = 100, \\ u(x,0) &= \sin(\pi x) \quad \text{for } 0 < x < 1. \end{aligned} \tag{15.E.24}$$

Hint: Use the fact that $u(x,t) = 100x$ is a solution satisfying $u_t = u_{xx}$, $u(0,t) = 0$, $u(1,t) = 100$. Then usesuperposition.

Exercise 15.E.4.6.6

Find the steady state temperature solution as a function of x alone, by letting $t \to \infty$ in the solution from exercises 15.E.4 and 15.E.5. Verify that it satisfies the equation $u_{xx} = 0$.

Exercise 15.E.4.6.7

Use separation variables to find a nontrivial solution to $u_{xx} + u_{yy} = 0$, where $u(x,0) = 0$ and $u(0,y) = 0$. Hint: Try $u(x,y) = X(x)Y(y)$.

Exercise 15.E.4.6.8: (challenging)

Suppose that one end of the wire is insulated (say at $x = 0$) and the other end is kept at zero temperature. That is, find a series solution of

$$\begin{aligned} u_t &= k u_{xx}, \\ u_x(0,t) &= u(L,t) = 0, \\ u(x,0) &= f(x) \quad \text{for } 0 < x < L. \end{aligned} \tag{15.E.25}$$

Express any coefficients in the series by integrals of $f(x)$.

Exercise 15.E.4.6.9: (challenging)

Suppose that the wire is circular and insulated, so there are no ends. You can think of this as simply connecting the two ends and making sure the solution matches up at the ends. That is, find a series solution of

$$\begin{aligned} u_t &= k u_{xx}, \\ u(0,t) &= u(L,t), \quad u_x(0,t) = u_x(L,t) \\ u(x,0) &= f(x) \quad \text{for } 0 < x < L. \end{aligned} \tag{15.E.26}$$

Express any coefficients in the series by integrals of $f(x)$.

Exercise 15.E.4.6.10

Consider a wire insulated on both ends, $L = 1$, $k = 1$, and $u(x,0) = \cos^2(\pi x)$.

a. Find the solution $u(x,t)$. Hint: a trig identity.
b. Find the average temperature.
c. Initially the temperature variation is 1 (maximum minus the minimum). Find the time when the variation is $\frac{1}{2}$.

Exercise 15.E.4.6.11

Find a series solution of

$$\begin{aligned} u_t &= 3u_{xx}, \\ u(0,t) &= u(\pi,t) = 0, \\ u(x,0) &= 5\sin(x) + 2\sin(5x) \quad \text{for } 0 < x < \pi. \end{aligned} \tag{15.E.27}$$

Answer

$u(x,t) = 5\sin(x)e^{3t} + 2\sin(5x)e^{-75t}$

Exercise 15.E.4.6.12

Find a series solution of

$$\begin{aligned} u_t &= 0.1u_{xx}, \\ u_x(0,t) &= u_x(\pi,t) = 0, \\ u(x,0) &= 1 + 2\cos(x) \quad \text{for } 0 < x < \pi. \end{aligned} \tag{15.E.28}$$

Answer

$u(x,t) = 1 + 2\cos(x)e^{-0.1t}$

Exercise 15.E.4.6.13

Use separation of variables to find a nontrivial solution to $u_{xt} = u_{xx}$.

Answer

$u(x,t) = e^{\lambda t}e^{\lambda x}$ for some λ

Exercise 15.E.4.6.14

Use separation of variables to find a nontrivial solution to $u_x + u_t = u$. (Hint: try $u(x,t) = X(x) + T(t)$).

Answer

$u(x,t) = Ae^x + Be^t$

Exercise 15.E.4.6.15

Suppose that the temperature on the wire is fixed at 0 at the ends, $L = 1$, $k = 1$, and $u(x,0) = 100\sin(2\pi x)$

a. What is the temperature at $x = \frac{1}{2}$ at any time.
b. What is the maximum and the minimum temperature on the wire at $t = 0$.
c. At what time is the maximum temperature on the wire exactly one half of the initial maximum at $t = 0$.

Answer

a. 0
b. minimum -100, maximum 100
c. $t = \frac{\ln 2}{4\pi^2}$

15.E.7: 4.7: One dimensional wave equation

Exercise 15.E.4.7.1

Solve

$$\begin{aligned} y_{tt} &= 9y_{xx}, \\ y(0,t) &= y(1,t) = 0, \\ y(x,0) &= \sin(3\pi x) + \frac{1}{4}\sin(6\pi x) \qquad \text{for } 0 < x < 1, \\ y_t(x,0) &= 0 \qquad\qquad\qquad\qquad\quad \text{for } 0 < x < 1. \end{aligned} \tag{15.E.29}$$

Exercise 15.E.4.7.2

Solve

$$\begin{aligned} y_{tt} &= 4y_{xx}, \\ y(0,t) &= y(1,t) = 0, \\ y(x,0) &= \sin(3\pi x) + \frac{1}{4}\sin(6\pi x) \qquad \text{for } 0 < x < 1, \\ y_t(x,0) &= \sin(9\pi x) \qquad\qquad\qquad\quad \text{for } 0 < x < 1. \end{aligned} \tag{15.E.30}$$

Exercise 15.E.4.7.3

Derive the solution for a general plucked string of length L, where we raise the string some distance b at the midpoint and let go, and for any constant a (in the equation $y_{tt} = a^2 y_{xx}$).

Exercise 15.E.4.7.4

Imagine that a stringed musical instrument falls on the floor. Suppose that the length of the string is 1 and $a = 1$. When the musical instrument hits the ground the string was in rest position and hence $y(x,0) = 0$. However, the string was moving at some velocity at impact $(t = 0)$), say $y_t(x,0) = -1$. Find the solution $y(x,t)$ for the shape of the string at time t.

Exercise 15.E.4.7.5: (challenging)

Suppose that you have a vibrating string and that there is air resistance proportional to the velocity. That is, you have

$$\begin{aligned} y_{tt} &= a^2 y_{xx} - k y_t, \\ y(0,t) &= y(1,t) = 0, \\ y(x,0) &= f(x) && \text{for } 0 < x < 1, \\ y_t(x,0) &= 0 && \text{for } 0 < x < 1. \end{aligned} \tag{15.E.31}$$

Suppose that $0 < k < 2\pi a$. Derive a series solution to the problem. Any coefficients in the series should be expressed as integrals of $f(x)$.

Exercise 15.E.4.7.6

Suppose you touch the guitar string exactly in the middle to ensure another condition $u\left(\frac{L}{2}, t\right) = 0$ for all time. Which multiples of the fundamental frequency $\frac{\pi a}{L}$ show up in the solution?

Exercise 15.E.4.7.7

Solve

$$\begin{aligned} y_{tt} &= y_{xx}, \\ y(0,t) &= y(\pi,t) = 0, \\ y(x,0) &= \sin(x) && \text{for } 0 < x < \pi, \\ y_t(x,0) &= \sin(x) && \text{for } 0 < x < \pi. \end{aligned} \tag{15.E.32}$$

Answer

$y(x,t) = \sin(x)\left(\sin(t) + \cos(t)\right)$

Exercise 15.E.4.7.8

Solve

$$\begin{aligned} y_{tt} &= 25 y_{xx}, \\ y(0,t) &= y(2,t) = 0, \\ y(x,0) &= 0 && \text{for } 0 < x < 2, \\ y_t(x,0) &= \sin(\pi x) + 0.1\sin(2\pi t) && \text{for } 0 < x < 2. \end{aligned} \tag{15.E.33}$$

Answer

$y(x,t) = \frac{1}{5\pi}\sin(\pi x)\sin(5\pi t) + \frac{1}{100\pi}\sin(2\pi x)\sin(10\pi t)$

Exercise 15.E.4.7.9

Solve

$$\begin{aligned} y_{tt} &= 2 y_{xx}, \\ y(0,t) &= y(\pi,t) = 0, \\ y(x,0) &= x && \text{for } 0 < x < \pi, \\ y_t(x,0) &= 0 && \text{for } 0 < x < \pi. \end{aligned} \tag{15.E.34}$$

Answer

$y(x,t) = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n}\sin(nx)\cos(n\sqrt{2}t)$

Exercise 15.E.4.7.10

Let's see what happens when $a=0$. Find a solution to $y_{tt}=0,\ y(0,t)=y(\pi,t)=0,\ y(x,0)=\sin(2x),\ y_t(x,0)=\sin(x)$.

Answer

$y(x,t)=\sin(2x)+t\sin(x)$

15.E.8: 4.8: D'Alembert solution of the wave equation

Exercise 15.E.4.8.1

Using the d'Alembert solution solve $y_{tt}=4y_{xx},\ 0<x<\pi, t>0, y(0,t)=y(\pi,t)=0, y(x,0)=\sin x,$ and $y_t(x,0)=\sin x$. Hint: Note that $\sin x$ is the odd extension of $y(x,0)$ and $y_t(x,0)$.

Exercise 15.E.4.8.2

Using the d'Alembert solution solve $y_{tt}=2y_{xx},\ 0<x<1, t>0, y(0,t)=y(1,t)=0, y(x,0)=\sin^5(\pi x),$ and $y_t(x,0)=\sin^3(\pi x)$.

Exercise 15.E.4.8.3

Take $y_{tt}=4y_{xx}$, $0<x<\pi, t>0, y(0,t)=y(\pi,t)=0, y(x,0)=x(\pi-x)$, and $y_t(x,0)=0$.

a. Solve using the d'Alembert formula. Hint: You can use the sine series for $y(x,0)$.
b. Find the solution as a function of x for a fixed $t=0.5,\ t=1,$ and $t=2$. Do not use the sine series here.

Exercise 15.E.4.8.4

Derive the d'Alembert solution for $y_{tt}=a^2y_{xx}$, $0<x<\pi, t>0, y(0,t)=y(\pi,t)=0, y(x,0)=f(x),$ and $y_t(x,0)=0$, using the Fourier series solution of the wave equation, by applying an appropriate trigonometric identity. Hint: Do it first for a single term of the Fourier series solution, in particular do it when y is $\sin\left(\frac{n\pi}{L}x\right)\sin\left(\frac{n\pi a}{L}t\right)$.

Exercise 15.E.4.8.5

The d'Alembert solution still works if there are no boundary conditions and the initial condition is defined on the whole real line. Suppose that $y_{tt}=y_{xx}$ (for all x on the real line and $t\geq 0$), $y(x,0)=f(x)$, and $y_t(x,0)$, where

$$f(x)=\begin{cases} 0 & \text{if} \quad x<-1, \\ x+1 & \text{if} \quad -1\leq x<0, \\ -x+1 & \text{if} \quad 0\leq x<1 \\ 0 & \text{if} \quad x>1. \end{cases}$$

Solve using the d'Alembert solution. That is, write down a piecewise definition for the solution. Then sketch the solution for $t=0, t=1/2, t=1$, and $t=2$.

Exercise 15.E.4.8.6

Using the d'Alembert solution solve $y_{tt}=9y_{xx},\ 0<x<1, t>0, y(0,t)=y(1,t)=0, y(x,0)=\sin(2\pi x),$ and $y_t(x,0)=\sin(3\pi x)$.

Answer

$y(x,t)=\frac{\sin(2\pi(x-3t))+\sin(2\pi(3t+x))}{2}+\frac{\cos(3\pi(x-3t))-\cos(3\pi(3t+x))}{18\pi}$

Exercise 15.E.4.8.7

Take $y_{tt} = 4y_{xx}$, $0 < x < 1, t > 0$, $y(0,t) = y(1,t) = 0$, $y(x,0) = x - x^2$, and $y_t(x,0) = 0$. Using the D'Alembert solution find the solution at

a. $t = 0.1$,
b. $t = 1/2$,
c. $t = 1$.

You may have to split your answer up by cases.

Answer

a. $y(x,t) = \begin{cases} x - x^2 - 0.04 & \text{if} \quad 0.2 \le x \le 0.8 \\ 0.6x & \text{if} \quad x \le 0.2 \\ 0.6 - 0.6x & \text{if} \quad x \ge 0.8 \end{cases}$
b. $y(x, \frac{1}{2}) = -x + x^2$
c. $y(x,1) = x - x^2$

Exercise 15.E.4.8.8

Take $y_{tt} = 100y_{xx}$, $0 < x < 4, t > 0$, $y(0,t) = y(4,t) = 0$, $y(x,0) = F(x)$, and $y_t(x,0) = 0$. Suppose that $F(0) = 0$, $F(1) = 2$, $F(2) = 3$, $F(3) = 1$. Using the D'Alembert solution find

a. $y(1,1)$,
b. $y(4,3)$,
c. $y(3,9)$.

Answer

a. $y(1,1) = -\frac{1}{2}$
b. $y(4,3) = 0$
c. $y(3,9) = \frac{1}{2}$

15.E.9: 4.9: Steady state temperature and the Laplacian

Exercise 15.E.4.9.1

Let R be the region described by $0 < x < \pi$ and $0 < y < \pi$. Solve the problem

$$\Delta u = 0, \quad u(x,0) = \sin x, \quad u(x,\pi) = 0, \quad u(0,y) = 0, \quad u(\pi,y) = 0.$$

Exercise 15.E.4.9.2

Let R be the region described by $0 < x < 1$ and $0 < y < 1$. Solve the problem

$$\begin{aligned} u_{xx} + u_{yy} &= 0, \\ u(x,0) &= \sin(\pi x) - \sin(2\pi x), \quad u(x,1) = 0, \\ u(0,y) &= 0, \quad u(1,y) = 0. \end{aligned} \tag{15.E.35}$$

Exercise 15.E.4.9.3

Let R be the region described by $0 < x < 1$ and $0 < y < 1$. Solve the problem

$$\begin{aligned} u_{xx} + u_{yy} &= 0, \\ u(x,0) &= u(x,1) = u(0,y) = u(1,y) = C. \end{aligned} \tag{15.E.36}$$

for some constant C. Hint: Guess, then check your intuition.

Exercise 15.E.4.9.4

Let R be the region described by $0<x<\pi$ and $0<y<\pi$. Solve

$$\Delta u=0,\quad u(x,0)=0,\quad u(x,\pi)=\pi,\quad u(0,y)=y,\quad u(\pi,y)=y.$$

Hint: Try a solution of the form $u(x,y)=X(x)+Y(y)$ (different separation of variables).

Exercise 15.E.4.9.5

Use the solution of Exercise 15.E.4 to solve

$$\Delta u=0,\quad u(x,0)=\sin x,\quad u(x,\pi)=\pi,\quad u(0,y)=y,\quad u(\pi,y)=y.$$

Hint: Use superposition.

Exercise 15.E.4.9.6

Let R be the region described by $0<x<w$ and $0<y<h$. Solve the problem

$$\begin{aligned} u_{xx}+u_{yy}&=0,\\ u(x,0)=0,&\qquad u(x,h)=f(x),\\ u(0,y)=0,&\qquad u(w,y)=0. \end{aligned} \tag{15.E.37}$$

The solution should be in series form using the Fourier series coefficients of $f(x)$.

Exercise 15.E.4.9.7

Let R be the region described by $0<x<w$ and $0<y<h$. Solve the problem

$$\begin{aligned} u_{xx}+u_{yy}&=0,\\ u(x,0)=0,&\qquad u(x,h)=0,\\ u(0,y)=f(y),&\qquad u(w,y)=0. \end{aligned} \tag{15.E.38}$$

The solution should be in series form using the Fourier series coefficients of $f(y)$.

Exercise 15.E.4.9.8

Let R be the region described by $0<x<w$ and $0<y<h$. Solve the problem

$$\begin{aligned} u_{xx}+u_{yy}&=0,\\ u(x,0)=0,&\qquad u(x,h)=0,\\ u(0,y)=0,&\qquad u(w,y)=f(y). \end{aligned} \tag{15.E.39}$$

The solution should be in series form using the Fourier series coefficients of $f(y)$.

Exercise 15.E.4.9.9

Let R be the region described by $0<x<1$ and $0<y<1$. Solve the problem

$$\begin{aligned} u_{xx}+u_{yy}&=0,\\ u(x,0)=\sin(9\pi x),&\qquad u(x,1)=\sin(2\pi x),\\ u(0,y)=0,&\qquad u(1,y)=0. \end{aligned} \tag{15.E.40}$$

Hint: Use superposition.

Exercise 15.E.4.9.10

Let R be the region described by $0<x<1$ and $0<y<1$. Solve the problem

$$\begin{aligned} u_{xx}+u_{yy} &= 0, \\ u(x,0)=\sin(\pi x), &\qquad u(x,1)=\sin(\pi x), \\ u(0,y)=\sin(\pi y), &\qquad u(1,y)=\sin(\pi y). \end{aligned} \tag{15.E.41}$$

Hint: Use superposition.

Exercise 15.E.4.9.11: (challenging)

Using only your intuition find $u(1/2,1/2)$ for the problem $\Delta u=0$, where $u(0,y)=u(1,y)=100$ for $0<y<1$, and $u(x,0)=u(x,1)=0$ for $0<x<1$. Explain.

Exercise 15.E.4.9.12

Let R be the region described by $0<x<1$ and $0<y<1$. Solve the problem

$$\Delta u=0, \quad u(x,0)=\sum_{n=1}^{\infty}\sin(n\pi x), \quad u(x,1)=0, \quad u(0,y)=0, \quad u(1,y)=0.$$

Answer

$$u(x,y)=\sum_{n=1}^{\infty}\frac{1}{n^2}\sin(n\pi x)\left(\frac{\sinh(n\pi(1-y))}{\sinh(n\pi)}\right)$$

Exercise 15.E.4.9.13

Let R be the region described by $0<x<1$ and $0<y<2$. Solve the problem

$$\Delta u=0, \quad u(x,0)=0.1\sin(\pi x), \quad u(x,2)=0, \quad u(0,y)=0, \quad u(1,y)=0.$$

Answer

$$u(x,y)=0.1\sin(\pi x)\left(\frac{\sinh(\pi(2-y))}{\sinh(2\pi)}\right)$$

15.E.10: 4.10: Dirichlet problem in the circle and the Poisson kernel

Exercise 15.E.4.10.1

Using series solve $\Delta u=0, u(1,\theta)=|\theta|$ for $-\pi<\theta\leq\pi$.

Exercise 15.E.4.10.2

Using series solve $\Delta u=0, u(1,\theta)=g(\theta)$ for the following data. Hint: trig identities.

a. $g(\theta)=1/2+3\sin(\theta)+\cos(3\theta)$
b. $g(\theta)=\cos(3\theta)+3\sin(3\theta)+\sin(9\theta)$
c. $g(\theta)=2\cos(\theta+1)$
d. $g(\theta)=\sin^2(\theta)$

Exercise 15.E.4.10.3

Using the Poisson kernel, give the solution to $\Delta u=0$, where $u(1,\theta)$ is zero for θ outside the interval $[-\pi/4,\pi/4]$and $u(1,\theta)$ is 1 for θ on the interval $[-\pi/4,\pi/4]$

Exercise 15.E.4.10.4

a. Draw a graph for the Poisson kernel as a function of α when $r = 1/2$ and $\theta = 0$.
b. Describe what happens to the graph when you make r bigger (as it approaches 1).
c. Knowing that the solution $u(r, \theta)$ is the weighted average of $g(\theta)$ with Poisson kernel as the weight, explain what your answer to part b means.

Exercise 15.E.4.10.5

Take the function $g(\theta)$ to be the function $xy = \cos(\theta)\sin(\theta)$ on the boundary. Use the series solution to find a solution to the Dirichlet problem $\Delta u = 0, u(1, \theta) = g(\theta)$. Now convert the solution to Cartesian coordinates x and y. Is this solution surprising? Hint: use your trig identities.

Exercise 15.E.4.10.6

Carry out the computation we needed in the separation of variables and solve $r^2R'' + rR' - n^2R = 0$, for $n = 0, 1, 2, 3, \ldots$.

Exercise 15.E.4.10.7: (challenging)

Derive the series solution to the Dirichlet problem if the region is a circle of radius ρ rather than 1. That is, solve $\Delta u = 0, u(\rho, \theta) = g(\theta)$.

Exercise 15.E.4.10.8: (challenging)

1. Find the solution for $\Delta u = 0$, $u(1, \theta) = x^2y^3 + 5x^2$. Write the answer in Cartesian coordinates.
2. Now solve $\Delta u = 0$, $u(1, \theta) = x^k y^l$. Write the solution in Cartesian coordinates.
3. Suppose you have a polynomial $P(x, y) = \sum_{j=0}^{m} \sum_{k=0}^{n} c_{j,k}x^j y^k$, solve $\Delta u = 0$, $u(1, \theta) = P(x, y)$ (that is, write down the formula for the answer). Write the answer in Cartesian coordinates.

Notice the answer is again a polynomial in x and y. See also Exercise 15.E.5.

Exercise 15.E.4.10.9

Using series solve $\Delta u = 0,\ u(1, \theta) = 1 + \sum_{n=1}^{\infty} \frac{1}{n^2}\sin(n\theta)$.

Answer

$$u = 1 \sum_{n=1}^{\infty} \frac{1}{n^2} r^n \sin(n\theta)$$

Exercise 15.E.4.10.10

Using the series solution find the solution to $\Delta u = 0,\ u(1, \theta) = 1 - \cos(\theta)$. Express the solution in Cartesian coordinates (that is, using x and y).

Answer

$$u = 1 - x$$

Exercise 15.E.4.10.11

a. Try and guess a solution to $\Delta u = -1,\ u(1, \theta) = 0$. Hint: try a solution that only depends on r. Also first, don't worry about the boundary condition.
b. Now solve $\Delta u = -1,\ u(1, \theta) = \sin(2\theta)$ using superposition.

Answer

a. $u = \frac{-1}{4} r^2 + \frac{1}{4}$
b. $u = \frac{-1}{4} r^2 + \frac{1}{4} + r^2 \sin(2\theta)$

? Exercise 15.E. 4.10.12: (challenging)

Derive the Poisson kernel solution if the region is a circle of radius ρ rather than 1. That is, solve $\Delta u = 0,\ u(\rho, \theta) = g(\theta)$.

Answer

$$u(r, \theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\rho^2 - r^2}{\rho - 2r\rho \cos(\theta - \alpha) + r^2} g(\alpha) d\alpha$$

This page titled 15.E: Fourier Series and PDEs (Exercises) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **4.E: Fourier series and PDEs (Exercises)** has no license indicated.

16: Power series methods

Contributors and Attributions

- Jiří Lebl (Oklahoma State University).These pages were supported by NSF grants DMS-0900885 and DMS-1362337.

This page titled 16: Power series methods is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **7: Power series methods** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

16.1: Power Series

Many functions can be written in terms of a power series

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

If we assume that a solution of a differential equation is written as a power series, then perhaps we can use a method reminiscent of undetermined coefficients. That is, we will try to solve for the numbers a_k. Before we can carry out this process, let us review some results and concepts about power series.

16.1.1: Definition

As we said, a power series is an expression such as

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k = a_0 + a_1 (x - x_0) + a_2 (x - x_0)^2 + a_3 (x - x_0)^3 + \cdots, \tag{16.1.1}$$

where $a_0, a_1, a_2, \ldots, a_k, \ldots$ and x_0 are constants. Let

$$S_n(x) = \sum_{k=0}^{n} a_k (x - x_0)^k = a_0 + a_1 (x - x_0) + a_2 (x - x_0)^2 + a_3 (x - x_0)^3 + \cdots + a_n (x - x_0)^n,$$

denote the so-called *partial sum*. If for some x, the limit

$$\lim_{n \to \infty} S_n(x) = \lim_{n \to \infty} \sum_{k=0}^{n} a_k (x - x_0)^k$$

exists, then we say that the series (16.1.1) converges at x. Note that for $x = x_0$, the series always converges to a_0. When (16.1.1) converges at any other point $x \neq x_0$, we say that (16.1.1) is a convergent power series. In this case we write

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k = \lim_{n \to \infty} \sum_{k=0}^{n} a_k (x - x_0)^k.$$

If the series does not converge for any point $x \neq x_0$, we say that the series is divergent.

Example 16.1.1

The series

$$\sum_{k=0}^{\infty} \frac{1}{k!} x^k = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots$$

is convergent for any x. Recall that $k! = 1 \cdot 2 \cdot 3 \cdots k$ is the factorial. By convention we define $0! = 1$. In fact, you may recall that this series converges to e^x.

We say that (16.1.1) converges absolutely at x whenever the limit

$$\lim_{n \to \infty} \sum_{k=0}^{n} |a_k| \, |x - x_0|^k$$

exists. That is, the series $\sum_{k=0}^{\infty} |a_k| \, |x - x_0|^k$ is convergent. If (16.1.1) converges absolutely at x, then it converges at x. However, the opposite implication is not true.

Example 16.1.2

The series

$$\sum_{k=1}^{\infty} \frac{1}{k} x^k$$

converges absolutely for all x in the interval $(-1, 1)$.

It converges at $x = -1$, as

$\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ converges (conditionally)

by the alternating series test.

But the power series does not converge absolutely at $x = -1$, because $\sum_{k=1}^{\infty} \frac{1}{k}$ does not converge. The series diverges at $x = 1$.

16.1.2: Radius of Convergence

If a power series converges absolutely at some x_1, then for all x such that $|x - x_0| \leq |x_1 - x_0|$ (that is, x is closer than x_1 to x_0) we have $|a_k(x - x_0)^k| \leq |a_k(x_1 - x_0)^k|$ for all k. As the numbers $|a_k(x_1 - x_0)^k|$ sum to some finite limit, summing smaller positive numbers $|a_k(x - x_0)^k|$ must also have a finite limit. Hence, the series must converge absolutely at x.

Theorem 16.1.1

For a power series (16.1.1), there exists a number ρ (we allow $\rho = \infty$) called the radius of convergence such that the series converges absolutely on the interval $(x_0 - \rho,\ x_0 + \rho)$ and diverges for $x < x_0 - \rho$ and $x > x_0 + \rho$. We write $\rho = \infty$ if the series converges for all x.

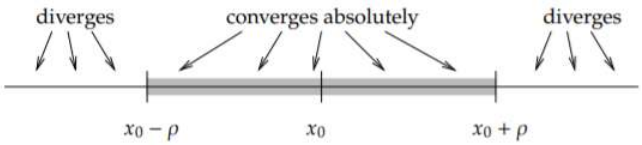


Figure 16.1.1: Convergence of a power series.

See Figure 16.1.1. In Example 16.1.1 the radius of convergence is $\rho = \infty$ as the series converges everywhere. In Example 16.1.2 the radius of convergence is $\rho = 1$. We note that $\rho = 0$ is another way of saying that the series is divergent. A useful test for convergence of a series is the *ratio test*. Suppose that

$$\sum_{k=0}^{\infty} c_k$$

is a series such that the limit

$$L = \lim_{n \to \infty} \left| \frac{c_{k+1}}{c_k} \right|$$

exists. Then the series converges absolutely if $L < 1$ and diverges if $L > 1$.

Let us apply this test to the series (16.1.1). That is we let $c_k = a_k(x - x_0)^k$ in the test. Compute

$$L = \lim_{n \to \infty} \left| \frac{c_{k+1}}{c_k} \right| = \lim_{n \to \infty} \left| \frac{a_{k+1}(x - x_0)^{k+1}}{a_k(x - x_0)^k} \right| = \lim_{n \to \infty} \left| \frac{a_{k+1}}{a_k} \right| |x - x_0|.$$

Define A by

$$A = \lim_{n \to \infty} \left| \frac{a_{k+1}}{a_k} \right|.$$

Then if $1 > L = A|x - x_0|$ the series (16.1.1) converges absolutely. If $A = 0$, then the series always converges. If $A > 0$, then the series converges absolutely if $|x - x_0| < \frac{1}{A}$, and diverges if $|x - x_0| > \frac{1}{A}$. That is, the radius of convergence is $\frac{1}{A}$.

A similar test is the *root test*. Suppose

$$L = \lim_{k\to\infty} \sqrt[3]{|c_k|}$$

exists. Then $\sum_{k=0}^{\infty} c_k$ converges absolutely if $L < 1$ and diverges if $L > 1$. We can use the same calculation as above to find A. Let us summarize.

Theorem 16.1.2: Ratio and Root Tests for Power Series

Let

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

be a power series such that

$$A = \lim_{n\to\infty} \left| \frac{a_{k+1}}{a_k} \right| \quad \text{or} \quad A = \lim_{k\to\infty} \sqrt[3]{|a_k|}$$

exists. If $A = 0$, then the radius of convergence of the series is ∞. Otherwise the radius of convergence is $\frac{1}{A}$.

Below is a video on finding the interval of convergence of a power series.



Below is another video on finding the interval of convergence of a power series.



Example 16.1.3

Suppose we have the series

$$\sum_{k=0}^{\infty} 2^{-k}(x-1)^k.$$

First we compute,

$$A = \lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty}\left|\frac{2^{-k-1}}{2^{-k}}\right| = 2^{-1} = \frac{1}{2}.$$

Therefore the radius of convergence is 2, and the series converges absolutely on the interval $(-1, 3)$. And we could just as well have used the root test:

$$A = \lim_{k\to\infty}\lim_{k\to\infty}\sqrt[k]{|a_k|} = \lim_{k\to\infty}\sqrt[k]{|2^{-k}|} = \lim_{k\to\infty} 2^{-1} = \frac{1}{2}.$$

Below is a video on the interval of convergence of a power series not centered at the origin.



Example 16.1.4

Consider

$$\sum_{k=0}^{\infty}\frac{1}{k^k}x^k.$$

Compute the limit for the root test,

$$A = \lim_{k\to\infty}\sqrt[k]{|a_k|} = \lim_{k\to\infty}\sqrt[k]{\left|\frac{1}{k^k}\right|} = \lim_{k\to\infty}\sqrt[k]{\left|\frac{1}{k}\right|^k} = \lim_{k\to\infty}\frac{1}{k} = 0.$$

So the radius of convergence is ∞: the series converges everywhere. The ratio test would also work here.

The root or the ratio test does not always apply. That is the limit of $\left|\frac{a_{k+1}}{a_k}\right|$ or $\sqrt[k]{|a_k|}$ might not exist. There exist more sophisticated ways of finding the radius of convergence, but those would be beyond the scope of this chapter. The two methods above cover many of the series that arise in practice. Often if the root test applies, so does the ratio test, and vice versa, though the limit might be easier to compute in one way than the other.

16.1.3: Analytic Functions

Functions represented by power series are called analytic functions. Not every function is analytic, although the majority of the functions you have seen in calculus are. An analytic function $f(x)$ is equal to its Taylor series[1] near a point x_0. That is, for x near x_0 we have

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k, \tag{16.1.2}$$

where $f^{(k)}(x_0)$ denotes the kth derivative of $f(x)$ at the point x_0.

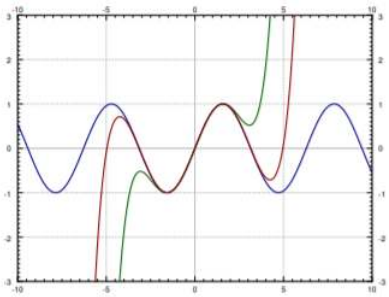


Figure 16.1.2: The sine function and its Taylor approximations around $x_o = 0$ of 5th and 9th degree.

For example, sine is an analytic function and its Taylor series around $x_0 = 0$ is given by

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}.$$

In Figure 16.1.2 we plot $\sin(x)$ and the truncations of the series up to degree 5 and 9. You can see that the approximation is very good for x near 0, but gets worse for x further away from 0. This is what happens in general. To get a good approximation far away from x_0 you need to take more and more terms of the Taylor series.

16.1.4: Manipulating Power Series

One of the main properties of power series that we will use is that we can differentiate them term by term. That is, suppose that $\sum a_k (x - x_0)^k$ is a convergent power series. Then for x in the radius of convergence we have

$$\frac{d}{dx}\left[\sum_{k=0}^{\infty} a_k (x - x_0)^k\right] = \sum_{k=1}^{\infty} k a_k (x - x_0)^{k-1}.$$

Notice that the term corresponding to $k = 0$ disappeared as it was constant. The radius of convergence of the differentiated series is the same as that of the original.

Example 16.1.5

Let us show that the exponential $y = e^x$ solves $y' = y$. First write

$$y = e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$

Now differentiate

$$y' = \sum_{k=1}^{\infty} k \frac{1}{k!} x^{k-1} = \sum_{k=1}^{\infty} \frac{1}{(k-1)!} x^{k-1}.$$

We *reindex* the series by simply replacing k with $k+1$. The series does not change, what changes is simply how we write it. After reindexing the series starts at $k = 0$ again.

$$\sum_{k=1}^{\infty} \frac{1}{(k-1)!} x^{k-1} = \sum_{k+1=1}^{\infty} \frac{1}{((k+1)-1)!} x^{(k+1)-1} = \sum_{k=0}^{\infty} \frac{1}{k!} x^k.$$

That was precisely the power series for e^x that we started with, so we showed that $\frac{d}{dx}[e^x] = e^x$.

Convergent power series can be added and multiplied together, and multiplied by constants using the following rules. First, we can add series by adding term by term,

$$\left(\sum_{k=0}^{\infty} a_k (x-x_0)^k\right) + \left(\sum_{k=0}^{\infty} b_k (x-x_0)^k\right) = \sum_{k=0}^{\infty} (a_k + b_k)(x-x_0)^k.$$

We can multiply by constants,

$$\alpha \left(\sum_{k=0}^{\infty} a_k (x-x_0)^k\right) = \sum_{k=0}^{\infty} \alpha a_k (x-x_0)^k.$$

We can also multiply series together,

$$\left(\sum_{k=0}^{\infty} a_k (x-x_0)^k\right)\left(\sum_{k=0}^{\infty} b_k (x-x_0)^k\right) = \sum_{k=0}^{\infty} c_k (x-x_0)^k,$$

where $c_k = a_0 b_k + a_1 b_{k-1} + \cdots + a_k b_0$. The radius of convergence of the sum or the product is at least the minimum of the radii of convergence of the two series involved.

Below is a video on differentiation and integration using power series.



16.1.5: Power Series for Rational Functions

Polynomials are simply finite power series. That is, a polynomial is a power series where the a_k are zero for all k large enough. We can always expand a polynomial as a power series about any point x_0 by writing the polynomial as a polynomial in $(x-x_0)$. For example, let us write $2x^2 - 3x + 4$ as a power series around $x_0 = 1$:

$$2x^2 - 3x + 4 = 3 + (x-1) + 2(x-1)^2.$$

In other words $a_0 = 3$, $a_1 = 1$, $a_2 = 2$, and all other $a_k = 0$. To do this, we know that $a_k = 0$ for all $k \geq 3$ as the polynomial is of degree 2.

We write $a_0 + a_1(x-1) + a_2(x-1)^2$, we expand, and we solve for a_0, a_1, and a_2. We could have also differentiated at $x = 1$and used the Taylor series formula (16.1.2).

Let us look at rational functions, that is, ratios of polynomials. An important fact is that a series for a function only defines the function on an interval even if the function is defined elsewhere. For example, for $-1 < x < 1$ we have

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \cdots$$

This series is called the *geometric series*. The ratio test tells us that the radius of convergence is 1. The series diverges for $x \leq -1$ and $x \geq 1$, even though $\frac{1}{1-x}$ is defined for all $x \neq 1$.

We can use the geometric series together with rules for addition and multiplication of power series to expand rational functions around a point, as long as the denominator is not zero at x_0. Note that as for polynomials, we could equivalently use the Taylor series expansion (16.1.2).

Below is a video on finding a power series to represent a rational function.



Example 16.1.5

Expand $\frac{x}{1+2x+x^2}$ as a power series around the origin $(x_0 = 0)$ and find the radius of convergence. First, write $1+2x+x^2 = (1+x)^2 = \big(1-(-x)\big)^2$. Now we compute

$$\begin{aligned} \frac{x}{1+2x+x^2} &= x\left(\frac{1}{1-(-x)}\right)^2 \\ &= x\left(\sum_{k=0}^{\infty}(-1)^k x^k\right)^2 \\ &= x\left(\sum_{k=0}^{\infty} c_k x^k\right) \\ &= \sum_{k=0}^{\infty} c_k x^{k+1}, \end{aligned} \tag{16.1.3}$$

where using the formula for the product of series we obtain, $c_0 = 1$, $c_1 = -1-1 = -2$, $c_2 = 1+1+1 = 3$, etc . . ..

Therefore

$$\frac{x}{1+2x+x^2} = \sum_{k=1}^{\infty}(-1)^{k+1} k x^k = x - 2x^2 + 3x^3 - 4x^4 + \cdots$$

The radius of convergence is at least 1. We use the ratio test

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty}\left|\frac{(-1)^{k+2}(k+1)}{(-1)^{k+1}k}\right| = \lim_{k\to\infty}\frac{k+1}{k} = 1.$$

So the radius of convergence is actually equal to 1.

When the rational function is more complicated, it is also possible to use method of partial fractions. For example, to find the Taylor series for $\frac{x^3+x}{x^2-1}$, we write

$$\frac{x^3+x}{x^2-1} = x + \frac{1}{1+x} - \frac{1}{1-x} = x + \sum_{k=0}^{\infty}(-1)^k x^k - \sum_{k=0}^{\infty} x^k = -x + \sum_{\substack{k=3 \\ k\,\text{odd}}}^{\infty}(-2)x^k.$$

16.1.6: Footnotes

[1] Named after the English mathematician Sir Brook Taylor (1685–1731).

This page titled 16.1: Power Series is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **7.1: Power Series** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

16.2: Series Solutions of Linear Second Order ODEs

Suppose we have a linear second order homogeneous ODE of the form

$$p(x)y'' + q(x)y' + r(x)y = 0 \tag{16.2.1}$$

Suppose that $p(x)$, $q(x)$, and $r(x)$ are polynomials. We will try a solution of the form

$$y = \sum_{k=0}^{\infty} a_k (x - x_o)^2$$

and solve for the a_k to try to obtain a solution defined in some interval around x_o.

Definition: Ordinary and Singular Points

The point x_o is called an *ordinary point* if $p(x_o) \neq 0$ in linear second order homogeneous ODE of the form in Equation 16.2.1. That is, the functions

$$\frac{q(x)}{p(x)} \quad \text{and} \quad \frac{r(x)}{p(x)}$$

are defined for x near x_o.

If $p(x_0) = 0$, then we say x_o is a *singular point*.

Handling singular points is harder than ordinary points and so we now focus only on ordinary points.

Example 16.2.1: Expansion around an Ordinary Point

Let us start with a very simple example

$$y'' - y = 0$$

Let us try a power series solution near $x_o = 0$, which is an ordinary point.

Solution

Every point is an ordinary point in fact, as the equation is constant coefficient. We already know we should obtain exponentials or the hyperbolic sine and cosine, but let us pretend we do not know this.

We try

$$y = \sum_{k=0}^{\infty} a_k x^k$$

If we differentiate, the $k = 0$ term is a constant and hence disappears. We therefore get

$$y' = \sum_{k=1}^{\infty} k\, a_k\, x^{k-1}$$

We differentiate yet again to obtain (now the $k = 1$ term disappears)

$$y'' = \sum_{k=2}^{\infty} k\,(k-1)\, a_k\, x^{k-2}$$

We reindex the series (replace k with $k+2$) to obtain

$$y'' = \sum_{k=0}^{\infty} (k+2)\,(k+1)\, a_{k+2}\, x^k$$

Now we plug y and y'' into the differential equation.

$$\begin{aligned} 0 &= y'' - y \\ &= \left(\sum_{k=0}^{\infty} (k+2)(k+1) a_{k+2} x^k \right) - \left(\sum_{k=0}^{\infty} a_k x^k \right) \\ &= \sum_{k=0}^{\infty} \left((k+2)(k+1) a_{k+2} x^k - a_k x^k \right) \\ &= \sum_{k=0}^{\infty} \left((k+2)(k+1) a_{k+2} - a_k \right) x^k . \end{aligned}$$

As $y'' - y$ is supposed to be equal to 0, we know that the coefficients of the resulting series must be equal to 0. Therefore,

$$(k+2)(k+1) a_{k+2} - a_k = 0 \tag{16.2.2}$$

or

$$a_{k+2} = \frac{a_k}{(k+2)(k+1)} .$$

The above equation is called a *recurrence relation* for the coefficients of the power series. It did not matter what a_0 or a_1 was. They can be arbitrary. But once we pick a_0 and a_1, then all other coefficients are determined by the recurrence relation.

Let us see what the coefficients must be. First, a_0 and a_1 are arbitrary

$$a_2 = \frac{a_0}{2}, \quad a_3 = \frac{a_1}{(3)(2)}, \quad a_4 = \frac{a_2}{(4)(3)} = \frac{a_0}{(4)(3)(2)}, \quad a_5 = \frac{a_3}{(5)(4)} = \frac{a_1}{(5)(4)(3)(2)}, \quad \ldots$$

So we note that for even k, that is $k = 2n$ we get

$$a_k = a_{2n} = \frac{a_o}{(2n)!}$$

and for odd k that is $k = 2n+1$ we have

$$a_k = a_{2n+1} = \frac{a_1}{(2n+1)!}$$

Let us write down the series

$$y = \sum_{k=0}^{\infty} a_k x^k = \sum_{n=0}^{\infty} \left(\frac{a_0}{(2n)!} x^{2n} + \frac{a_1}{(2n+1)!} x^{2n+1} \right) = a_0 \sum_{n=0}^{\infty} \frac{1}{(2n)!} x^{2n} + a_1 \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} x^{2n+1} .$$

We recognize the two series as the hyperbolic sine and cosine. Therefore,

$$y = a_o \cosh x + a_1 \sinh x$$

Of course, in general we will not be able to recognize the series that appears, since usually there will not be any elementary function that matches it. In that case we will be content with the series.

Example 16.2.2

Let us do a more complex example. Suppose we wish to solve *Airy's equation*[1], that is

$$y'' - xy = 0$$

near the point $x_0 = 0$, which is an ordinary point.

We try

$$y = \sum_{k=0}^{\infty} a_k x^k$$

We differentiate twice (as above) to obtain

$$y'' = \sum_{k=2}^{\infty} k(k-1)a_k x^{k-2}$$

We plug y into the equation

$$\begin{aligned} 0 = y'' - xy &= \left(\sum_{k=2}^{\infty} k\,(k-1)\,a_k x^{k-2}\right) - x\left(\sum_{k=0}^{\infty} a_k x^k\right) \\ &= \left(\sum_{k=2}^{\infty} k\,(k-1)\,a_k x^{k-2}\right) - \left(\sum_{k=0}^{\infty} a_k x^{k+1}\right). \end{aligned} \tag{16.2.3}$$

We reindex to make things easier to sum

$$\begin{aligned} 0 = y'' - xy &= \left(2a_2 + \sum_{k=1}^{\infty} (k+2)\,(k+1)\,a_{k+2} x^k\right) - \left(\sum_{k=1}^{\infty} a_{k-1} x^k\right). \\ &= 2a_2 + \sum_{k=1}^{\infty} \Big((k+2)\,(k+1)\,a_{k+2} - a_{k-1}\Big)\, x^k. \end{aligned} \tag{16.2.4}$$

Again $y'' - xy$ is supposed to be 0 so first we notice that $a_2 = 0$ and also

$$(k+2)\,(k+1)\,a_{k+2} - a_{k-1} = 0, \qquad \text{or} \qquad a_{k+2} = \frac{a_{k-1}}{(k+2)(k+1)}.$$

Now we jump in steps of three. First we notice that since $a_2 = 0$ we must have that, $a_5 = 0$, $a_8 = 0$, $a_{11} = 0$, etc .... In general $a_{3n+2} = 0$. The constants a_0 and a_1 are arbitrary and we obtain

$$a_3 = \frac{a_0}{(3)(2)}, \quad a_4 = \frac{a_1}{(4)(3)}, \quad a_6 = \frac{a_3}{(6)(5)} = \frac{a_0}{(6)(5)(3)(2)}, \quad a_7 = \frac{a_4}{(7)(6)} = \frac{a_1}{(7)(6)(4)(3)}, \quad \ldots$$

For a_k where k is a multiple of 3, that is $k = 3n$ we notice that

$$a_{3n} = \frac{a_0}{(2)(3)(5)(6)\cdots(3n-1)(3n)}.$$

For a_k where $k = 3n+1$, we notice

$$a_{3n+1} = \frac{a_1}{(3)(4)(6)(7)\cdots(3n)(3n+1)}.$$

In other words, if we write down the series for y we notice that it has two parts

$$\begin{aligned} y &= \left(a_0 + \frac{a_0}{6}x^3 + \frac{a_0}{180}x^6 + \cdots + \frac{a_0}{(2)(3)(5)(6)\cdots(3n-1)(3n)}x^{3n} + \cdots\right) \\ &\quad + \left(a_1 x + \frac{a_1}{12}x^4 + \frac{a_1}{504}x^7 + \cdots + \frac{a_1}{(3)(4)(6)(7)\cdots(3n)(3n+1)}x^{3n+1} + \cdots\right) \\ &= a_0\left(1 + \frac{1}{6}x^3 + \frac{1}{180}x^6 + \cdots + \frac{1}{(2)(3)(5)(6)\cdots(3n-1)(3n)}x^{3n} + \cdots\right) \\ &\quad + a_1\left(x + \frac{1}{12}x^4 + \frac{1}{504}x^7 + \cdots + \frac{1}{(3)(4)(6)(7)\cdots(3n)(3n+1)}x^{3n+1} + \cdots\right). \end{aligned} \tag{16.2.5}$$

We define

$$\begin{aligned} y_1(x) &= 1 + \frac{1}{6}x^3 + \frac{1}{180}x^6 + \cdots + \frac{1}{(2)(3)(5)(6)\cdots(3n-1)(3n)}x^{3n} + \cdots, \\ y_2(x) &= x + \frac{1}{12}x^4 + \frac{1}{504}x^7 + \cdots + \frac{1}{(3)(4)(6)(7)\cdots(3n)(3n+1)}x^{3n+1} + \cdots, \end{aligned} \tag{16.2.6}$$

and write the general solution to the equation as $y(x) = a_0 y_1(x) + a_1 y_2(x)$. Notice from the power series that $y_1(0) = 1$ and $y_2(0) = 0$. Also, $y_1'(0) = 0$ and $y_2'(0) = 1$. Therefore $y(x)$ is a solution that satisfies the initial conditions $y(0) = a_0$ and $y'(0) = a_1$.

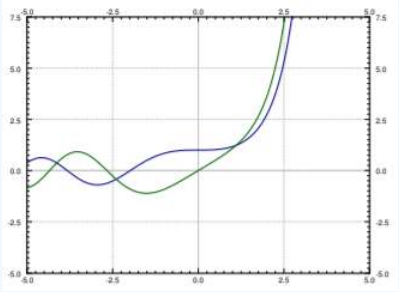


Figure 16.2.1: The two solutions y_1 and y_2 to Airy's equation.

The functions y_1 and y_2 cannot be written in terms of the elementary functions that you know. See Figure 16.2.1 for the plot of the solutions y_1 and y_2. These functions have many interesting properties. For example, they are oscillatory for negative x (like solutions to $y''+y=0$) and for positive x they grow without bound (like solutions to $y''-y=0$).

Sometimes a solution may turn out to be a polynomial.

Example 16.2.3: Hermite Equation

Let us find a solution to the so-called Hermite's equation of order n[2] is the equation

$$y''-2xy'+2ny=0.$$

Find a solution around the point $x_0=0$.

Solution

We try

$$y=\sum_{k=0}^{\infty}a_kx^k.$$

We differentiate (as above) to obtain

$$\begin{aligned} y' &= \sum_{k=1}^{\infty}ka_kx^{k-1}, \\ y'' &= \sum_{k=2}^{\infty}k\left(k-1\right)a_kx^{k-2}. \end{aligned} \tag{16.2.7}$$

Now we plug into the equation

$$\begin{aligned} 0 &= y''-2xy'+2ny \\ &= \left(\sum_{k=2}^{\infty}k\left(k-1\right)a_kx^{k-2}\right)-2x\left(\sum_{k=1}^{\infty}ka_kx^{k-1}\right)+2n\left(\sum_{k=0}^{\infty}a_kx^k\right) \\ &= \left(\sum_{k=2}^{\infty}k\left(k-1\right)a_kx^{k-2}\right)-\left(\sum_{k=1}^{\infty}2ka_kx^k\right)+\left(\sum_{k=0}^{\infty}2na_kx^k\right) \\ &= \left(2a_2+\sum_{k=1}^{\infty}(k+2)\left(k+1\right)a_{k+2}x^k\right)-\left(\sum_{k=1}^{\infty}2ka_kx^k\right)+\left(2na_0+\sum_{k=1}^{\infty}2na_kx^k\right) \\ &= 2a_2+2na_0+\sum_{k=1}^{\infty}\left(\left(k+2\right)\left(k+1\right)a_{k+2}-2ka_k+2na_k\right)x^k. \end{aligned} \tag{16.2.8}$$

As $y''-2xy'+2ny=0$ we have

$$(k+2)\,(k+1)\,a_{k+2}+(-2k+2n)a_k=0, \qquad \text{or} \qquad a_{k+2}=\frac{(2k-2n)}{(k+2)(k+1)}a_k.$$

This recurrence relation actually includes $a_2=-na_0$ (which comes about from $2a_2+2na_0=0$). Again a_0 and a_1 are arbitrary.

$$\begin{aligned} a_2 &= \frac{-2n}{(2)(1)}a_0, \qquad a_3=\frac{2(1-n)}{(3)(2)}a_1, \\ a_4 &= \frac{2(2-n)}{(4)(3)}a_2=\frac{2^2(2-n)(-n)}{(4)(3)(2)(1)}a_0, \\ a_5 &= \frac{2(3-n)}{(5)(4)}a_3=\frac{2^2(3-n)(1-n)}{(5)(4)(3)(2)}a_1, \quad \ldots \end{aligned} \tag{16.2.9}$$

Let us separate the even and odd coefficients. We find that

$$\begin{aligned} a_{2m} &= \frac{2^m(-n)(2-n)\cdots(2m-2-n)}{(2m)!}, \\ a_{2m+1} &= \frac{2^m(1-n)(3-n)\cdots(2m-1-n)}{(2m+1)!}. \end{aligned} \tag{16.2.10}$$

Let us write down the two series, one with the even powers and one with the odd.

$$\begin{aligned} y_1(x) &= 1+\frac{2(-n)}{2!}x^2+\frac{2^2(-n)(2-n)}{4!}x^4+\frac{2^3(-n)(2-n)(4-n)}{6!}x^6+\cdots, \\ y_2(x) &= x+\frac{2(1-n)}{3!}x^3+\frac{2^2(1-n)(3-n)}{5!}x^5+\frac{2^3(1-n)(3-n)(5-n)}{7!}x^7+\cdots. \end{aligned} \tag{16.2.11}$$

We then write

$$y(x)=a_0y_1(x)+a_1y_2(x).$$

We also notice that if n is a positive even integer, then $y_1(x)$ is a polynomial as all the coefficients in the series beyond a certain degree are zero. If n is a positive odd integer, then $y_2(x)$ is a polynomial. For example, if $n=4$, then

$$y_1(x)=1+\frac{2(-4)}{2!}x^2+\frac{2^2(-4)(2-4)}{4!}x^4=1-4x^2+\frac{4}{3}x^4.$$

16.2.1: Footnotes

[1] Named after the English mathematician Sir George Biddell Airy (1801 – 1892).

[2] Named after the French mathematician Charles Hermite (1822–1901).

This page titled 16.2: Series Solutions of Linear Second Order ODEs is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **7.2: Series Solutions of Linear Second Order ODEs** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

16.3: Singular Points and the Method of Frobenius

16.3.1: Examples

While behavior of ODEs at singular points is more complicated, certain singular points are not especially difficult to solve. Let us look at some examples before giving a general method. We may be lucky and obtain a power series solution using the method of the previous section, but in general we may have to try other things.

Example 16.3.1

Let us first look at a simple first order equation

$$2xy' - y = 0. \tag{16.3.1}$$

Note that $x = 0$ is a singular point. If we only try to plug in

$$y = \sum_{k=0}^{\infty} a_k x^k, \tag{16.3.2}$$

we obtain

$$\begin{aligned} 0 = 2xy' - y &= 2x \left(\sum_{k=1}^{\infty} k a_k x^{k-1} \right) - \left(\sum_{k=0}^{\infty} a_k x^k \right) \\ &= a_0 + \sum_{k=1}^{\infty} (2ka_k - a_k) \, x^k. \end{aligned} \tag{16.3.3}$$

First, $a_0 = 0$. Next, the only way to solve $0 = 2ka_k - a_k = (2k-1)\, a_k$ for $k = 1, 2, 3, \ldots$ is for $a_k = 0$ for all k. Therefore we only get the trivial solution $y = 0$. We need a nonzero solution to get the general solution.

Let us try $y = x^r$ for some real number r. Consequently our solution---if we can find one---may only make sense for positive x. Then $y' = rx^{r-1}$. So

$$0 = 2xy' - y = 2xrx^{r-1} - x^r = (2r-1)x^r.$$

Therefore $r = \dfrac{1}{2}$, or in other words $y = x^{1/2}$. Multiplying by a constant, the general solution for positive x is

$$y = Cx^{1/2}.$$

If $C \neq 0$ then the derivative of the solution "blows up" at $x = 0$ (the singular point). There is only one solution that is differentiable at $x = 0$ and that's the trivial solution $y = 0$.

Not every problem with a singular point has a solution of the form $y = x^r$, of course. But perhaps we can combine the methods. What we will do is to try a solution of the form

$$y = x^r f(x)$$

where $f(x)$ is an analytic function.

Example 16.3.2

Suppose that we have the equation

$$4x^2y'' - 4x^2y' + (1 - 2x)y = 0, \tag{16.3.4}$$

and again note that $x = 0$ is a singular point. Let us try

$$y = x^r \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} a_k x^{k+r}, \tag{16.3.5}$$

where r is a real number, not necessarily an integer. Again if such a solution exists, it may only exist for positive x. First let us find the derivatives

$$y' = \sum_{k=0}^{\infty} (k+r)\, a_k x^{k+r-1},$$
$$y'' = \sum_{k=0}^{\infty} (k+r)\,(k+r-1)\, a_k x^{k+r-2}. \tag{16.3.6}$$

Plugging Equations 16.3.5 - 16.3.6 into our original differential equation (Equation 16.3.4) we obtain

$$\begin{aligned}
0 &= 4x^2 y'' - 4x^2 y' + (1-2x)y \\
&= 4x^2 \left(\sum_{k=0}^{\infty} (k+r)\,(k+r-1)\, a_k x^{k+r-2} \right) - 4x^2 \left(\sum_{k=0}^{\infty} (k+r)\, a_k x^{k+r-1} \right) + (1-2x) \left(\sum_{k=0}^{\infty} a_k x^{k+r} \right) \\
&= \left(\sum_{k=0}^{\infty} 4(k+r)\,(k+r-1)\, a_k x^{k+r} \right) - \left(\sum_{k=0}^{\infty} 4(k+r)\, a_k x^{k+r+1} \right) + \left(\sum_{k=0}^{\infty} a_k x^{k+r} \right) - \left(\sum_{k=0}^{\infty} 2a_k x^{k+r+1} \right) \\
&= \left(\sum_{k=0}^{\infty} 4(k+r)\,(k+r-1)\, a_k x^{k+r} \right) - \left(\sum_{k=1}^{\infty} 4(k+r-1)\, a_{k-1} x^{k+r} \right) + \left(\sum_{k=0}^{\infty} a_k x^{k+r} \right) - \left(\sum_{k=1}^{\infty} 2a_{k-1} x^{k+r} \right) \\
&= 4r(r-1)\, a_0 x^r + a_0 x^r + \sum_{k=1}^{\infty} \left(4(k+r)\,(k+r-1)\, a_k - 4(k+r-1)\, a_{k-1} + a_k - 2a_{k-1} \right) x^{k+r} \\
&= \left(4r(r-1)+1 \right)\, a_0 x^r + \sum_{k=1}^{\infty} \left(\left(4(k+r)\,(k+r-1)+1 \right)\, a_k - \left(4(k+r-1)+2 \right)\, a_{k-1} \right) x^{k+r}.
\end{aligned} \tag{16.3.7}$$

To have a solution we must first have $\left(4r(r-1)+1 \right)\, a_0 = 0$. Supposing that $a_0 \neq 0$ we obtain

$$4r(r-1)+1 = 0.$$

This equation is called the *indicial equation*. This particular indicial equation has a double root at $r = \frac{1}{2}$.

OK, so we know what r has to be. That knowledge we obtained simply by looking at the coefficient of x^r. All other coefficients of x^{k+r} also have to be zero so

$$\left(4(k+r)\,(k+r-1)+1 \right)\, a_k - \left(4(k+r-1)+2 \right)\, a_{k-1} = 0.$$

If we plug in $r = \frac{1}{2}$ and solve for a_k we get

$$a_k = \frac{4(k+\frac{1}{2}-1)+2}{4(k+\frac{1}{2})\,(k+\frac{1}{2}-1)+1}\, a_{k-1} = \frac{1}{k}\, a_{k-1}.$$

Let us set $a_0 = 1$. Then

$$a_1 = \frac{1}{1} a_0 = 1, \qquad a_2 = \frac{1}{2} a_1 = \frac{1}{2}, \qquad a_3 = \frac{1}{3} a_2 = \frac{1}{3 \cdot 2}, \qquad a_4 = \frac{1}{4} a_3 = \frac{1}{4 \cdot 3 \cdot 2}, \qquad \ldots$$

Extrapolating, we notice that

$$a_k = \frac{1}{k(k-1)(k-2)\cdots 3 \cdot 2} = \frac{1}{k!}.$$

In other words,

$$y = \sum_{k=0}^{\infty} a_k x^{k+r} = \sum_{k=0}^{\infty} \frac{1}{k!} x^{k+1/2} = x^{1/2} \sum_{k=0}^{\infty} \frac{1}{k!} x^k = x^{1/2} e^x.$$

That was lucky! In general, we will not be able to write the series in terms of elementary functions. We have one solution, let us call it $y_1 = x^{1/2} e^x$. But what about a second solution? If we want a general solution, we need two linearly independent solutions. Picking a_0 to be a different constant only gets us a constant multiple of y_1, and we do not have any other r to try; we only have one solution to the indicial equation. Well, there are powers of x floating around and we are taking derivatives, perhaps the logarithm (the antiderivative of x^{-1}) is around as well. It turns out we want to try for another solution of the form

$$y_2 = \sum_{k=0}^{\infty} b_k x^{k+r} + (\ln x) y_1,$$

which in our case is

$$y_2 = \sum_{k=0}^{\infty} b_k x^{k+1/2} + (\ln x) x^{1/2} e^x.$$

We now differentiate this equation, substitute into the differential equation and solve for b_k. A long computation ensues and we obtain some recursion relation for b_k. The reader can (and should) try this to obtain for example the first three terms

$$b_1 = b_0 - 1, \qquad b_2 = \frac{2b_1 - 1}{4}, \qquad b_3 = \frac{6b_2 - 1}{18}, \qquad \dots$$

We then fix b_0 and obtain a solution y_2. Then we write the general solution as $y = Ay_1 + By_2$.

16.3.2: Method of Frobenius

Before giving the general method, let us clarify when the method applies. Let

$$p(x)y'' + q(x)y' + r(x)y = 0$$

be an ODE. As before, if $p(x_0) = 0$, then x_0 is a singular point. If, furthermore, the limits

$$\lim_{x \to x_0} (x - x_0) \frac{q(x)}{p(x)} \qquad \text{and} \qquad \lim_{x \to x_0} (x - x_0)^2 \frac{r(x)}{p(x)}$$

both exist and are finite, then we say that x_0 is a *regular singular point.*

Example 16.3.3: Expansion around a regular singular point

Often, and for the rest of this section, $x_0 = 0$. Consider

$$x^2 y'' + x(1+x)y' + (\pi + x^2)y = 0.$$

Write

$$\begin{aligned} \lim_{x \to 0} x \frac{q(x)}{p(x)} &= \lim_{x \to 0} x \frac{x(1+x)}{x^2} = \lim_{x \to 0} (1+x) = 1, \\ \lim_{x \to 0} x^2 \frac{r(x)}{p(x)} &= \lim_{x \to 0} x^2 \frac{(\pi + x^2)}{x^2} = \lim_{x \to 0} (\pi + x^2) = \pi \end{aligned} \tag{16.3.8}$$

So $x = 0$ is a regular singular point.

On the other hand if we make the slight change

$$x^2 y'' + (1+x)y' + (\pi + x^2)y = 0,$$

then

$$\lim_{x \to 0} x \frac{q(x)}{p(x)} = \lim_{x \to 0} x \frac{(1+x)}{x^2} = \lim_{x \to 0} \frac{1+x}{x} = \text{DNE}.$$

Here DNE stands for *does not exist*. The point 0 is a singular point, but not a regular singular point.

Below is part 1 of a video on the method of Frobenius.



Below is part 2 of a video on the method of Frobenius.



Let us now discuss the general *Method of Frobenius*[1]. Let us only consider the method at the point $x = 0$ for simplicity. The main idea is the following theorem.

Theorem 16.3.1

Method of Frobenius

Suppose that

$$p(x)y'' + q(x)y' + r(x)y = 0 \tag{16.3.9}$$

has a regular singular point at $x = 0$, then there exists at least one solution of the form

$$y = x^r \sum_{k=0}^{\infty} a_k x^k.$$

A solution of this form is called a *Frobenius-type solution*.

The method usually breaks down like this.

i. We seek a Frobenius-type solution of the form

$$y = \sum_{k=0}^{\infty} a_k x^{k+r}.$$

We plug this y into equation (16.3.9). We collect terms and write everything as a single series.

ii. The obtained series must be zero. Setting the first coefficient (usually the coefficient of x^r) in the series to zero we obtain the *indicial equation*, which is a quadratic polynomial in r.

iii. If the indicial equation has two real roots r_1 and r_2 such that $r_1 - r_2$ is not an integer, then we have two linearly independent Frobenius-type solutions. Using the first root, we plug in

$$y_1 = x^{r_1} \sum_{k=0}^{\infty} a_k x^k,$$

and we solve for all a_k to obtain the first solution. Then using the second root, we plug in

$$y_2 = x^{r_2} \sum_{k=0}^{\infty} b_k x^k,$$

and solve for all b_k to obtain the second solution.

iv. If the indicial equation has a doubled root r, then there we find one solution

$$y_1 = x^{r} \sum_{k=0}^{\infty} a_k x^k,$$

and then we obtain a new solution by plugging

$$y_2 = x^{r} \sum_{k=0}^{\infty} b_k x^k + (\ln x) y_1,$$

into Equation (16.3.9) and solving for the constants b_k.

v. If the indicial equation has two real roots such that $r_1 - r_2$ is an integer, then one solution is

$$y_1 = x^{r_1} \sum_{k=0}^{\infty} a_k x^k,$$

and the second linearly independent solution is of the form

$$y_2 = x^{r_2} \sum_{k=0}^{\infty} b_k x^k + C(\ln x) y_1,$$

where we plug y_2 into (16.3.9) and solve for the constants b_k and C.

vi. Finally, if the indicial equation has complex roots, then solving for a_k in the solution

$$y = x^{r_1} \sum_{k=0}^{\infty} a_k x^k$$

results in a complex-valued function---all the a_k are complex numbers. We obtain our two linearly independent solutions[2] by taking the real and imaginary parts of y.

The main idea is to find at least one Frobenius-type solution. If we are lucky and find two, we are done. If we only get one, we either use the ideas above or even a different method such as reduction of order (Exercise 2.1.8) to obtain a second solution.

Below is a video on using the method of Frobenious to solve a differential equation.



Below is another video on using the method of Frobenious to solve a differential equation.



16.3.3: Bessel Functions

An important class of functions that arises commonly in physics are the *Bessel functions*[3]. For example, these functions appear when solving the wave equation in two and three dimensions. First we have *Bessel's equation* of order p:

$$x^2y'' + xy' + \left(x^2 - p^2\right)y = 0.$$

We allow p to be any number, not just an integer, although integers and multiples of $\frac{1}{2}$ are most important in applications. When we plug

$$y = \sum_{k=0}^{\infty} a_k x^{k+r}$$

into Bessel's equation of order p we obtain the indicial equation

$$r(r-1) + r - p^2 = (r-p)(r+p) = 0.$$

Therefore we obtain two roots $r_1 = p$ and $r_2 = -p$. If p is not an integer following the method of Frobenius and setting $a_0 = 1$, we obtain linearly independent solutions of the form

$$\begin{aligned} y_1 &= x^p \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{2^{2k}k!(k+p)(k-1+p)\cdots(2+p)(1+p)}, \\ y_2 &= x^{-p} \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{2^{2k}k!(k-p)(k-1-p)\cdots(2-p)(1-p)}. \end{aligned} \tag{16.3.10}$$

? Exercise 16.3.1

a. Verify that the indicial equation of Bessel's equation of order p is $(r-p)(r+p) = 0$.
b. Suppose that p is not an integer. Carry out the computation to obtain the solutions y_1 and y_2 above.

Bessel functions will be convenient constant multiples of y_1 and y_2. First we must define the *gamma function*

$$\Gamma(x) = \int_0^{\infty} t^{x-1}e^{-t}\,dt.$$

Notice that $\Gamma(1) = 1$. The gamma function also has a wonderful property

$$\Gamma(x+1) = x\Gamma(x).$$

From this property, one can show that $\Gamma(n) = (n-1)!$ when n is an integer, so the gamma function is a continuous version of the factorial. We compute:

$$\begin{aligned} \Gamma(k+p+1) &= (k+p)(k-1+p)\cdots(2+p)(1+p)\Gamma(1+p), \\ \Gamma(k-p+1) &= (k-p)(k-1-p)\cdots(2-p)(1-p)\Gamma(1-p). \end{aligned} \tag{16.3.11}$$

? Exercise 16.3.2

Verify the above identities using $\Gamma(x+1) = x\Gamma(x)$.

We define the *Bessel functions of the first kind* of order p and $-p$ as

$$\begin{aligned} J_p(x) &= \frac{1}{2^p\Gamma(1+p)}y_1 = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(k+p+1)}\left(\frac{x}{2}\right)^{2k+p}, \\ J_{-p}(x) &= \frac{1}{2^{-p}\Gamma(1-p)}y_2 = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!\Gamma(k-p+1)}\left(\frac{x}{2}\right)^{2k-p}. \end{aligned} \tag{16.3.12}$$

As these are constant multiples of the solutions we found above, these are both solutions to Bessel's equation of order p. The constants are picked for convenience.

When p is not an integer, J_p and J_{-p} are linearly independent. When n is an integer we obtain

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(k+n)!}\left(\frac{x}{2}\right)^{2k+n}.$$

In this case it turns out that

$$J_n(x) = (-1)^n J_{-n}(x),$$

and so we do not obtain a second linearly independent solution. The other solution is the so-called *Bessel function of second kind*. These make sense only for integer orders n and are defined as limits of linear combinations of $J_p(x)$ and $J_{-p}(x)$ as p approaches n in the following way:

$$Y_n(x) = \lim_{p\to n} \frac{\cos(p\pi)J_p(x) - J_{-p}(x)}{\sin(p\pi)}.$$

As each linear combination of $J_p(x)$ and $J_{-p}(x)$ is a solution to Bessel's equation of order p, then as we take the limit as p goes to n, $Y_n(x)$ is a solution to Bessel's equation of order n. It also turns out that $Y_n(x)$ and $J_n(x)$ are linearly independent. Therefore when n is an integer, we have the general solution to Bessel's equation of order n

$$y = AJ_n(x) + BY_n(x),$$

for arbitrary constants A and B. Note that $Y_n(x)$ goes to negative infinity at $x = 0$. Many mathematical software packages have these functions $J_n(x)$ and $Y_n(x)$ defined, so they can be used just like say $\sin(x)$ and $\cos(x)$. In fact, they have some similar properties. For example, $-J_1(x)$ is a derivative of $J_0(x)$, and in general the derivative of $J_n(x)$ can be written as a linear combination of $J_{n-1}(x)$ and $J_{n+1}(x)$. Furthermore, these functions oscillate, although they are not periodic. See Figure 16.3.1 for graphs of Bessel functions.

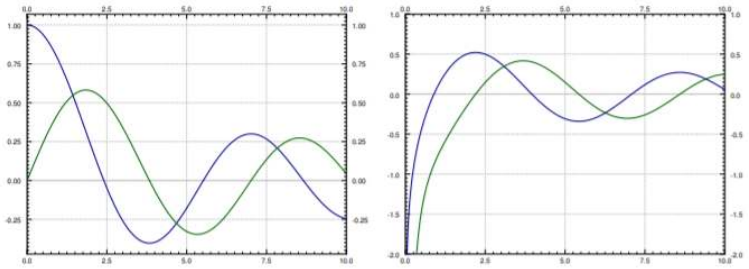


Figure 16.3.1: Plot of the $J_0(x)$ and $J_1(x)$ in the first graph and $Y_0(x)$ and $Y_1(x)$ in the second graph.

✓ Example 16.3.4: Using Bessel functions to Solve a ODE

Other equations can sometimes be solved in terms of the Bessel functions. For example, given a positive constant λ,

$$xy'' + y' + \lambda^2 xy = 0,$$

can be changed to $x^2y'' + xy' + \lambda^2 x^2 y = 0$. Then changing variables $t = \lambda x$ we obtain via chain rule the equation in y and t:

$$t^2y'' + ty' + t^2y = 0,$$

which can be recognized as Bessel's equation of order 0. Therefore the general solution is $y(t) = AJ_0(t) + BY_0(t)$, or in terms of x:

$$y = AJ_0(\lambda x) + BY_0(\lambda x).$$

This equation comes up for example when finding fundamental modes of vibration of a circular drum, but we digress.

16.3.4: Footnotes

[1] Named after the German mathematician Ferdinand Georg Frobenius (1849 – 1917).

[2] See Joseph L. Neuringera, The Frobenius method for complex roots of the indicial equation, International Journal of Mathematical Education in Science and Technology, Volume 9, Issue 1, 1978, 71–77.

[3] Named after the German astronomer and mathematician Friedrich Wilhelm Bessel (1784 – 1846).

This page titled 16.3: Singular Points and the Method of Frobenius is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **7.3: Singular Points and the Method of Frobenius** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

16.E: Power series methods (Exercises)

16.E.1: 7.1: Power Series

Exercise 16.E.7.1.1

Is the power series $\sum_{k=0}^{\infty} e^k x^k$ convergent? If so, what is the radius of convergence?

Exercise 16.E.7.1.2

Is the power series $\sum_{k=0}^{\infty} kx^k$ convergent? If so, what is the radius of convergence?

Exercise 16.E.7.1.3

Is the power series $\sum_{k=0}^{\infty} k!x^k$ convergent? If so, what is the radius of convergence?

Exercise 16.E.7.1.4

Is the power series $\sum_{k=0}^{\infty} \frac{1}{(2k)!}(x-10)^k$ convergent? If so, what is the radius of convergence?

Exercise 16.E.7.1.5

Determine the Taylor series for $\sin x$ around the point $x_0 = \pi$.

Exercise 16.E.7.1.6

Determine the Taylor series for $\ln x$ around the point $x_0 = 1$, and find the radius of convergence.

Exercise 16.E.7.1.7

Determine the Taylor series and its radius of convergence of $\dfrac{1}{1+x}$ around $x_0 = 0$.

Exercise 16.E.7.1.8

Determine the Taylor series and its radius of convergence of $\dfrac{x}{4-x^2}$ around $x_0 = 0$. Hint: You will not be able to use the ratio test.

Exercise 16.E.7.1.9

Expand $x^5 + 5x + 1$ as a power series around $x_0 = 5$.

Exercise 16.E.7.1.10

Suppose that the ratio test applies to a series $\sum_{k=0}^{\infty} a_k x^k$. Show, using the ratio test, that the radius of convergence of the differentiated series is the same as that of the original series.

Exercise 16.E.7.1.11

Suppose that f is an analytic function such that $f^{(n)}(0) = n$. Find $f(1)$.

Exercise 16.E.7.1.12

Is the power series $\sum_{n=1}^{\infty}(0.1)^n x^n$ convergent? If so, what is the radius of convergence?

Answer

Yes. Radius of convergence is 10.

Exercise 16.E.7.1.13: (challenging)

Is the power series $\sum_{n=1}^{\infty}\frac{n!}{n^n}x^n$ convergent? If so, what is the radius of convergence?

Answer

Yes. Radius of convergence is e.

Exercise 16.E.7.1.14

Using the geometric series, expand $\frac{1}{1-x}$ around $x_0=2$. For what x does the series converge?

Answer

$\frac{1}{1-x}=-\frac{1}{1-(2-x)}$ so $\frac{1}{1-x}=\sum_{n=0}^{\infty}(-1)^{n+1}(x-2)^n$, which converges for $1<x<3$.

Exercise 16.E.7.1.15: (challenging)

Find the Taylor series for x^7e^x around $x_0=0$.

Answer

$\sum_{n=7}^{\infty}\frac{1}{(n-7)!}x^n$

Exercise 16.E.7.1.16: (challenging)

Imagine f and g are analytic functions such that $f^{(k)}(0)=g^{(k)}(0)$ for all large enough k. What can you say about $f(x)-g(x)$?

Answer

$f(x)-g(x)$ is a polynomial. Hint: Use Taylor series.

16.E.2: 7.2: Series solutions of linear second order ODEs

In the following exercises, when asked to solve an equation using power series methods, you should find the first few terms of the series, and if possible find a general formula for the kth coefficient.

Exercise 16.E.7.2.1

Use power series methods to solve $y''+y=0$ at the point $x_0=1$.

Exercise 16.E.7.2.2

Use power series methods to solve $y''+4xy=0$ at the point $x_0=0$.

Exercise 16.E.7.2.3

Use power series methods to solve $y'' - xy = 0$ at the point $x_0 = 1$.

Exercise 16.E.7.2.4

Use power series methods to solve $y'' + x^2 y = 0$ at the point $x_0 = 0$.

Exercise 16.E.7.2.5

The methods work for other orders than second order. Try the methods of this section to solve the first order system $y' - xy = 0$ at the point $x_0 = 0$.

Exercise 16.E.7.2.6

Chebyshev's equation of order p:

a. Solve $(1-x^2)y'' - xy' + p^2 y = 0$ using power series methods at $x_0 = 0$.
b. For what p is there a polynomial solution?

Exercise 16.E.7.2.7

Find a polynomial solution to $(x^2+1)y'' - 2xy' + 2y = 0$ using power series methods.

Exercise 16.E.7.2.8

a. Use power series methods to solve $(1-x)y'' + y = 0$ at the point $x_0 = 0$.
b. Use the solution to part a) to find a solution for $xy'' + y = 0$ around the point $x_0 = 1$.

Exercise 16.E.7.2.9

Use power series methods to solve $y'' + 2x^3 y = 0$ at the point $x_0 = 0$.

Answer

$a_2 = 0$, $a_3 = 0$, $a_4 = 0$, recurrence relation (for $k \geq 5$): $a_k = \frac{-2a_{k-5}}{k(k-1)}$, so
$y(x) = a_0 + a_1 x - \frac{a_0}{10}x^5 - \frac{a_1}{15}x^6 + \frac{a_0}{450}x^{10} + \frac{a_1}{825}x^{11} - \frac{a_0}{47250}x^{15} - \frac{a_1}{99000}x^{16} + \cdots$

Exercise 16.E.7.2.10: (challenging)

We can also use power series methods in nonhomogeneous equations.

a. Use power series methods to solve $y'' - xy = \frac{1}{1-x}$ at the point $x_0 = 0$. Hint: Recall the geometric series.
b. Now solve for the initial condition $y(0) = 0$, $y'(0) = 0$.

Answer

a. $a_2 = \frac{1}{2}$, and for $k \geq 1$ we have $a_k = \frac{a_{k-3}+1}{k(k-1)}$, so
$y(x) = a_0 + a_1 x + \frac{1}{2}x^2 + \frac{a_0+1}{6}x^3 + \frac{a_1+1}{12}x^4 + \frac{3}{40}x^5 + \frac{a_0+2}{30}x^6 + \frac{a_1+2}{42}x^7 + \frac{5}{112}x^8 + \frac{a_0+3}{72}x^9 + \frac{a_1+3}{90}x^{10} + \cdots$
b. $y(x) = \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{12}x^4 + \frac{3}{40}x^5 + \frac{1}{15}x^6 + \frac{1}{21}x^7 + \frac{5}{112}x^8 + \frac{1}{24}x^9 + \frac{1}{30}x^{10} + \cdots$

https://math.libretexts.org/@go/page/128104

Exercise 16.E.7.2.11

Attempt to solve $x^2y'' - y = 0$ at $x_0 = 0$ using the power series method of this section (x_0 is a singular point). Can you find at least one solution? Can you find more than one solution?

Answer

Applying the method of this section directly we obtain $a_k = 0$ for all k and so $y(x) = 0$ is the only solution we find.

16.E.3: 7.3: Singular points and the method of Frobenius

Exercise 16.E.7.3.1

Find a particular (Frobenius-type) solution of $x^2y'' + xy' + (1+x)y = 0$.

Exercise 16.E.7.3.2

Find a particular (Frobenius-type) solution of $xy'' - y = 0$.

Exercise 16.E.7.3.3

Find a particular (Frobenius-type) solution of $y'' + \frac{1}{x}y' - xy = 0$.

Exercise 16.E.7.3.4

Find the general solution of $2xy'' + y' - x^2y = 0$.

Exercise 16.E.7.3.5

Find the general solution of $x^2y'' - xy' - y = 0$.

Exercise 16.E.7.3.6

In the following equations classify the point $x = 0$ as ordinary, regular singular, or singular but not regular singular.

a. $x^2(1+x^2)y'' + xy = 0$
b. $x^2y'' + y' + y = 0$
c. $xy'' + x^3y' + y = 0$
d. $xy'' + xy' - e^xy = 0$
e. $x^2y'' + x^2y' + x^2y = 0$

Exercise 16.E.7.3.7

In the following equations classify the point $x = 0$ as ordinary, regular singular, or singular but not regular singular.

a. $y'' + y = 0$
b. $x^3y'' + (1+x)y = 0$
c. $xy'' + x^5y' + y = 0$
d. $\sin(x)y'' - y = 0$
e. $\cos(x)y'' - \sin(x)y = 0$

Answer

a. ordinary,
b. singular but not regular singular,
c. regular singular,

d. regular singular,
e. ordinary.

Exercise 16.E.7.3.8

Find the general solution of $x^2y'' - y = 0$.

Answer

$y = Ax^{\frac{1+\sqrt{5}}{2}} + Bx^{\frac{1-\sqrt{5}}{2}}$

Exercise 16.E.7.3.9

Find a particular solution of $x^2y'' + (x - \frac{3}{4})y = 0$.

Answer

$y = x^{3/2}\sum_{k=0}^{\infty} \frac{(-1)^{-1}}{k!(k+2)!}x^k$ (Note that for convenience we did not pick $a_0 = 1$.)

Exercise 16.E.7.3.10: (tricky)

Find the general solution of $x^2y'' - xy' + y = 0$.

Answer

$y = Ax + Bx\ln(x)$

This page titled 16.E: Power series methods (Exercises) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- 7.E: Power series methods (Exercises) has no license indicated.

17: Nonlinear Systems

Contributors and Attributions

- Jiří Lebl (Oklahoma State University).

These pages were supported by NSF grants DMS-0900885 and DMS-1362337.

This page titled 17: Nonlinear Systems is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **8: Nonlinear Systems** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

17.1: Linearization, Critical Points, and Equilibria

Except for a few brief detours in Chapter 1, we considered mostly linear equations. Linear equations suffice in many applications, but in reality most phenomena require nonlinear equations. Nonlinear equations, however, are notoriously more difficult to understand than linear ones, and many strange new phenomena appear when we allow our equations to be nonlinear.

Not to worry, we did not waste all this time studying linear equations. Nonlinear equations can often be approximated by linear ones if we only need a solution "locally," for example, only for a short period of time, or only for certain parameters. Understanding linear equations can also give us qualitative understanding about a more general nonlinear problem. The idea is similar to what you did in calculus in trying to approximate a function by a line with the right slope.

In Section 2.4 we looked at the pendulum of mass m and length L. The goal was to solve for the angle $\theta(t)$ as a function of the time t. The equation for the setup is the nonlinear equation

$$\theta'' + \frac{g}{L}\sin\theta = 0.$$

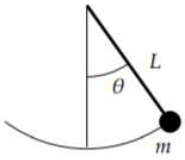


Figure 17.1.1

Instead of solving this equation, we solved the rather easier linear equation

$$\theta'' + \frac{g}{L}\theta = 0.$$

While the solution to the linear equation is not exactly what we were looking for, it is rather close to the original, as long as the angle θ is small and the time period involved is short.

You might ask: Why don't we just solve the nonlinear problem? Well, it might be very difficult, impractical, or impossible to solve analytically,depending on the equation in question. We may not even be interested in the actual solution, we might only be interested in some qualitative idea of what the solution is doing. For example, what happens as time goes to infinity?

17.1.1: Autonomous Systems and Phase Plane Analysis

We restrict our attention to a two dimensional autonomous system

$$x' = f(x,y), \qquad y' = g(x,y),$$

where $f(x,y)$ and $g(x,y)$ are functions of two variables, and the derivatives are taken with respect to time t. Solutions are functions $x(t)$ and $y(t)$ such that

$$x'(t) = f\big(x(t),y(t)\big), \qquad y'(t) = g\big(x(t),y(t)\big).$$

The way we will analyze the system is very similar to Section 1.6, where we studied a single autonomous equation. The ideas in two dimensions are the same, but the behavior can be far more complicated.

It may be best to think of the system of equations as the single vector equation

$$\begin{bmatrix} x \\ y \end{bmatrix}' = \begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}. \tag{17.1.1}$$

As in Section 3.1 we draw the *phase portrait* (or *phase diagram*), where each point (x,y) corresponds to a specific state of the system. We draw the *vector field* given at each point (x,y) by the vector $\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}$. And as before if we find solutions, we draw the trajectories by plotting all points $\big(x(t),y(t)\big)$ for a certain range of t.

Example 17.1.1

Consider the second order equation $x'' = -x + x^2$. Write this equation as a first order nonlinear system

$$x' = y, \qquad y' = -x + x^2.$$

The phase portrait with some trajectories is drawn in Figure 17.1.2

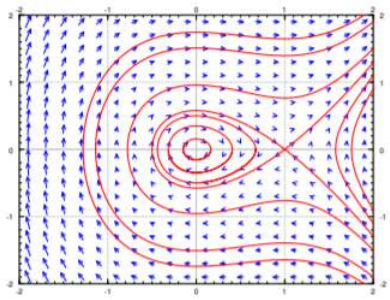


Figure 17.1.2: Phase portrait with some trajectories of $x' = y$, $y' = -x + x^2$.

From the phase portrait it should be clear that even this simple system has fairly complicated behavior. Some trajectories keep oscillating around the origin, and some go off towards infinity. We will return to this example often, and analyze it completely in this (and the next) section.

If we zoom into the diagram near a point where $\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}$ is not zero, then nearby the arrows point generally in essentially that same direction and have essentially the same magnitude. In other words the behavior is not that interesting near such a point. We are of course assuming that $f(x, y)$ and $g(x, y)$ are continuous.

Let us concentrate on those points in the phase diagram above where the trajectories seem to start, end, or go around. We see two such points: $(0, 0)$ and $(1, 0)$. The trajectories seem to go around the point $(0, 0)$, and they seem to either go in or out of the point $(1, 0)$. These points are precisely those points where the derivatives of both x and y are zero. Let us define the *critical points* as the points (x, y) such that

$$\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix} = \vec{0}.$$

In other words, the points where both $f(x, y) = 0$ and $g(x, y) = 0$.

The critical points are where the behavior of the system is in some sense the most complicated. If $\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}$ is zero, then nearby, the vector can point in any direction whatsoever. Also, the trajectories are either going towards, away from, or around these points, so if we are looking for long term behavior of the system, we should look at what happens there.

Critical points are also sometimes called *equilibria*, since we have so-called *equilibrium solutions* at critical points. If (x_0, y_0) is a critical point, then we have the solutions

$$x(t) = x_0, \quad y(t) = y_0.$$

In Example 17.1.1, there are two equilibrium solutions:

$$x(t) = 0, \quad y(t) = 0, \qquad \text{and} \qquad x(t) = 1, \quad y(t) = 0.$$

Compare this discussion on equilibria to the discussion in Section 1.6. The underlying concept is exactly the same.

17.1.2: Linearization

In Section 3.5 we studied the behavior of a homogeneous linear system of two equations near a critical point. For a linear system of two variables the only critical point is generally the origin $(0, 0)$. Let us put the understanding we gained in that section to good use understanding what happens near critical points of nonlinear systems.

In calculus we learned to estimate a function by taking its derivative and linearizing. We work similarly with nonlinear systems of ODE. Suppose (x_0, y_0) is a critical point. First change variables to (u, v), so that $(u, v) = (0, 0)$ corresponds to (x_0, y_0). That is,

$$u = x - x_0, \qquad v = y - y_0.$$

Next we need to find the derivative. In multivariable calculus you may have seen that the several variables version of the derivative is the *Jacobian matrix*[1]. The Jacobian matrix of the vector-valued function $\begin{bmatrix} f(x,y) \\ g(x,y) \end{bmatrix}$ at (x_0, y_0) is

$$\begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \\ \frac{\partial g}{\partial x}(x_0, y_0) & \frac{\partial g}{\partial y}(x_0, y_0) \end{bmatrix}.$$

This matrix gives the best linear approximation as u and v (and therefore x and y) vary. We define the *linearization* of the equation (17.1.1) as the linear system

$$\begin{bmatrix} u \\ v \end{bmatrix}' = \begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \\ \frac{\partial g}{\partial x}(x_0, y_0) & \frac{\partial g}{\partial y}(x_0, y_0) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.$$

Example 17.1.2

Let us keep with the same equations as Example 17.1.1: $x' = y$, $y' = -x + x^2$. There are two critical points, $(0, 0)$and $(1, 0)$. The Jacobian matrix at any point is

$$\begin{bmatrix} \frac{\partial f}{\partial x}(x, y) & \frac{\partial f}{\partial y}(x, y) \\ \frac{\partial g}{\partial x}(x, y) & \frac{\partial g}{\partial y}(x, y) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 + 2x & 0 \end{bmatrix}.$$

Therefore at $(0, 0)$, we have $u = x$ and $v = y$, and the linearization is

$$\begin{bmatrix} u \\ v \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix},$$

where $u = x$ and $v = y$.

At the point $(1, 0)$, we have $u = x - 1$ and $v = y$, and the linearization is

$$\begin{bmatrix} u \\ v \end{bmatrix}' = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}.$$

The phase diagrams of the two linearizations at the point $(0, 0)$ and $(1, 0)$ are given in Figure 17.1.3. Note that the variables are now u and v. Compare Figure 17.1.3 with Figure 17.1.2, and look especially at the behavior near the critical points.

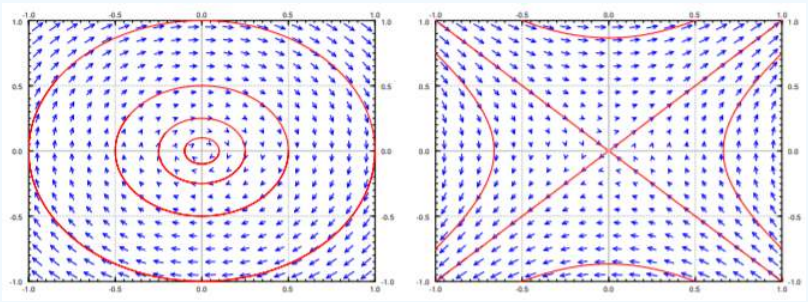


Figure 17.1.3: Phase diagram with some trajectories of linearizations at the critical points $(0, 0)$ (left) and $(1, 0)$ (right) of $x' = y$, $y' = -x + x^2$.

17.1.3: Footnotes

[1] Named for the German mathematician Carl Gustav Jacob Jacobi (1804–1851).

This page titled 17.1: Linearization, Critical Points, and Equilibria is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **8.1: Linearization, Critical Points, and Equilibria** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

17.2: Stability and Classification of Isolated Critical Points

17.2.1: Isolated Critical Points and Almost Linear Systems

A *critical point* is *isolated* if it is the only critical point in some small "neighborhood" of the point. That is, if we zoom in far enough it is the only critical point we see. In the above example, the critical point was isolated. If on the other hand there would be a whole curve of critical points, then it would not be isolated.

A system is called *almost linear* (at a critical point (x_0, y_0)) if the critical point is isolated and the Jacobian at the point is invertible, or equivalently if the linearized system has an isolated critical point. In such a case, the nonlinear terms will be very small and the system will behave like its linearization, at least if we are close to the critical point.

In particular the system we have just seen in Examples 8.1.1 and 8.1.2 has two isolated critical points $(0, 0)$ and $(0, 1)$, and is almost linear at both critical points as both of the Jacobian matrices $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ are *invertible*.

On the other hand a system such as $x' = x^2, y' = y^2$ has an isolated critical point at $(0, 0)$, however the Jacobian matrix

$$\begin{bmatrix} 2x & 0 \\ 0 & 2y \end{bmatrix}$$

is zero when $(x, y) = (0, 0)$. Therefore the system is not almost linear. Even a worse example is the system $x' = x, y' = x^2$, which does not have an isolated critical point, as x' and y' are both zero whenever $x = 0$, that is, the entire y axis.

Fortunately, most often critical points are isolated, and the system is almost linear at the critical points. So if we learn what happens here, we have figured out the majority of situations that arise in applications.

17.2.2: Stability and Classification of Isolated Critical Points

Once we have an isolated critical point, the system is almost linear at that critical point, and we computed the associated linearized system, we can classify what happens to the solutions. We more or less use the classification for linear two-variable systems from Section 3.5, with one minor caveat. Let us list the behaviors depending on the eigenvalues of the Jacobian matrix at the critical point in Table 17.2.1. This table is very similar to Table 3.5.1, with the exception of missing "center" points. We will discuss centers later, as they are more complicated.

Table 17.2.1: Behavior of an almost linear system near an isolated critical point.

Eigenvalues of the Jacobian matrix	Behavior	Stability
real and both positive	source / unstable node	unstable
real and both negative	sink / stable node	asymptotically stable
real and opposite signs	saddle	unstable
complex with positive real part	spiral source	unstable
complex with negative real part	spiral sink	asymptotically stable

In the new third column, we have marked points as *asymptotically stable* or *unstable*. Formally, a *stable critical point* (x_0, y_0) is one where given any small distance ϵ to (x_0, y_0),and any initial condition within a perhaps smaller radius around (x_0, y_0),the trajectory of the system will never go further away from (x_0, y_0) than ϵ. An *unstable critical point* is one that is not stable. Informally, a point is stable if we start close to a critical point and follow a trajectory we will either go towards, or at least not get away from, this critical point.

A stable critical point (x_0, y_0) is called *asymptotically stable* if given any initial condition sufficiently close to (x_0, y_0) and any solution $\big(x(t), y(t)\big)$ given that condition, then

$$\lim_{t\to\infty} \big(x(t), y(t)\big) = (x_0, y_0).$$

That is, the critical point is asymptotically stable if any trajectory for a sufficiently close initial condition goes towards the critical point (x_0, y_0).

Example 17.2.1

Consider $x' = -y - x^2$, $y' = -x + y^2$. See Figure 17.2.1 for the phase diagram. Let us find the critical points. These are the points where $-y - x^2 = 0$ and $-x + y^2 = 0$. The first equation means $y = -x^2$, and so $y^2 = x^4$. Plugging into the second equation we obtain $-x + x^4 = 0$. Factoring we obtain $x(1 - x^3) = 0$. Since we are looking only for real solutions we get either $x = 0$ or $x = 1$. Solving for the corresponding y using $y = -x^2$,we get two critical points, one being $(0, 0)$ and the other being $(1, -1)$. Clearly the critical points are isolated. Let us compute the Jacobian matrix:

$$\begin{bmatrix} -2x & -1 \\ -1 & 2y \end{bmatrix}.$$

At the point $(0, 0)$ we get the matrix $\left[\begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix}\right]$ and so the two eigenvalues are 1 and -1. As the matrix is invertible, the system is almost linear at $(0, 0)$. As the eigenvalues are real and of opposite signs, we get a saddle point, which is an unstable equilibrium point.

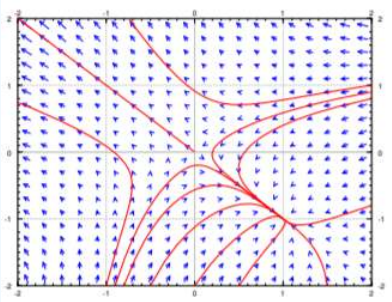


Figure 17.2.1: The phase portrait with few sample trajectories of $x' = -y - x^2$, $y' = -x + y^2$.

At the point $(1, -1)$ we get the matrix $\left[\begin{smallmatrix} -2 & -1 \\ -1 & -2 \end{smallmatrix}\right]$ and computing the eigenvalues we get -1, -3.The matrix is invertible, and so the system is almost linear at $(1, -1)$. As we have real eigenvalues both negative, the critical point is a sink, and therefore an asymptotically stable equilibrium point. That is, if we start with any point (x_i, y_i) close to $(1, -1)$ as an initial condition and plot a trajectory, it will approach $(1, -1)$. In other words,

$$\lim_{t\to\infty} \big(x(t), y(t)\big) = (1, -1).$$

As you can see from the diagram, this behavior is true even for some initial points quite far from $(1, -1)$,but it is definitely not true for all initial points.

Example 17.2.2

Let us look at $x' = y + y^2 e^x$, $y' = x$. First let us find the critical points. These are the points where $y + y^2 e^x = 0$ and $x = 0$. Simplifying we get $0 = y + y^2 = y(y + 1)$. So the critical points are $(0, 0)$ and $(0, -1)$,and hence are isolated. Let us compute the Jacobian matrix:

$$\begin{bmatrix} y^2 e^x & 1 + 2ye^x \\ 1 & 0 \end{bmatrix}.$$

At the point $(0, 0)$ we get the matrix $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ and so the two eigenvalues are 1 and -1. As the matrix is invertible, the system is almost linear at $(0, 0)$. And, as the eigenvalues are real and of opposite signs, we get a saddle point, which is an unstable equilibrium point.

At the point $(0, -1)$ we get the matrix $\left[\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right]$ whose eigenvalues are $\frac{1}{2} \pm i\frac{\sqrt{3}}{2}$. The matrix is invertible, and so the system is almost linear at $(0, -1)$. As we have complex eigenvalues with positive real part, the critical point is a spiral source, and therefore an unstable equilibrium point.

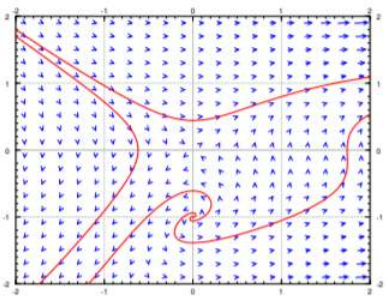


Figure 17.2.2: The phase portrait with few sample trajectories of $x' = y + y^2 e^x$, $y' = x$.

See Figure 17.2.2 for the phase diagram. Notice the two critical points, and the behavior of the arrows in the vector field around these points.

17.2.3: Trouble with Centers

Recall, a linear system with a center meant that trajectories traveled in closed elliptical orbits in some direction around the critical point. Such a critical point we would call a *center* or a *stable center*. It would not be an asymptotically stable critical point, as the trajectories would never approach the critical point, but at least if you start sufficiently close to the critical point, you will stay close to the critical point. The simplest example of such behavior is the linear system with a center. Another example is the critical point $(0, 0)$ in Example 8.1.1.

The trouble with a center in a nonlinear system is that whether the trajectory goes towards or away from the critical point is governed by the sign of the real part of the eigenvalues of the Jacobian. Since this real part is zero at the critical point itself, it can have either sign nearby, meaning the trajectory could be pulled towards or away from the critical point.

Example 17.2.3

An easy example where such a problematic behavior is exhibited is the system $x' = y, y' = -x + y^3$. The only critical point is the origin $(0, 0)$. The Jacobian matrix is

$$\begin{bmatrix} 0 & 1 \\ -1 & 3y^2 \end{bmatrix}.$$

At $(0, 0)$ the Jacobian matrix is $\left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$,which has eigenvalues $\pm i$. Therefore, the linearization has a center.

Using the quadratic equation, the eigenvalues of the Jacobian matrix at any point (x, y) are

$$\lambda = \frac{3}{2}y^2 \pm i\frac{\sqrt{4 - 9y^4}}{2}.$$

At any point where $y \neq 0$ (so at most points near the origin), the eigenvalues have a positive real part (y^2 can never be negative). This positive real part will pull the trajectory away from the origin. A sample trajectory for an initial condition near the origin is given in Figure 17.2.3

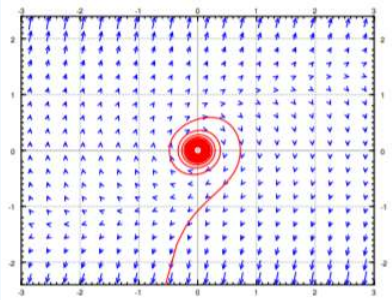


Figure 17.2.3: An unstable critical point (spiral source) at the origin for $x' = y$, $y' = -x + y^3$, even if the linearization has a center.

The moral of the example is that further analysis is needed when the linearization has a center. The analysis will in general be more complicated than in the above example, and is more likely to involve case-by-case consideration. Such a complication should not be surprising to you. By now in your mathematical career, you have seen many places where a simple test is inconclusive, perhaps starting with the second derivative test for maxima or minima, and requires more careful, and perhaps ad hoc analysis of the situation.

17.2.4: Conservative Equations

An equation of the form

$$x'' + f(x) = 0$$

for an arbitrary function $f(x)$ is called a *conservative equation*. For example the pendulum equation is a conservative equation. The equations are conservative as there is no friction in the system so the energy in the system is "conserved." Let us write this equation as a system of nonlinear ODE.

$$x' = y, \qquad y' = -f(x).$$

These types of equations have the advantage that we can solve for their trajectories easily. The trick is to first think of y as a function of x for a moment. Then use the chain rule

$$x'' = y' = y\frac{dy}{dx},$$

where the prime indicates a derivative with respect to t. We obtain $y\frac{dy}{dx} + f(x) = 0$. We integrate with respect to x to get $\int y\frac{dy}{dx}\,dx + \int f(x)\,dx = C$. In other words

$$\frac{1}{2}y^2 + \int f(x)\,dx = C.$$

We obtained an implicit equation for the trajectories, with different C giving different trajectories. The value of C is conserved on any trajectory. This expression is sometimes called the *Hamiltonian* or the energy of the system. If you look back to Section 1.8, you will notice that $y\frac{dy}{dx} + f(x) = 0$ is an exact equation, and we just found a potential function.

Example 17.2.4

Let us find the trajectories for the equation $x'' + x - x^2 = 0$, which is the equation from Example 8.1.1. The corresponding first order system is

$$x' = y, \qquad y' = -x + x^2.$$

Trajectories satisfy

$$\frac{1}{2}y^2 + \frac{1}{2}x^2 - \frac{1}{3}x^3 = C.$$

We solve for y

$$y = \pm\sqrt{-x^2 + \frac{2}{3}x^3 + 2C}\,.$$

Plotting these graphs we get exactly the trajectories in Figure 8.1.2. In particular we notice that near the origin the trajectories are *closed curves*: they keep going around the origin, never spiraling in or out. Therefore we discovered a way to verify that the critical point at $(0,0)$ is a stable center. The critical point at $(0,1)$ is a saddle as we already noticed. This example is typical for conservative equations.

Consider an arbitrary conservative equation. The trajectories are given by

$$y = \pm\sqrt{-2\int f(x)\,dx + 2C}.$$

So all trajectories are mirrored across the x-axis. In particular, there can be no spiral sources nor sinks. All critical points occur when $y = 0$ (the x-axis), that is when $x' = 0$. The critical points are simply those points on the x-axis where $f(x) = 0$. The Jacobian matrix is

$$\begin{bmatrix} 0 & 1 \\ -f'(x) & 0 \end{bmatrix}.$$

So the critical point is almost linear if $f'(x) \neq 0$ at the critical point. Let J denote the Jacobian matrix, then the eigenvalues of J are solutions to

$$0 = \det(J - \lambda I) = \lambda^2 + f'(x).$$

Therefore $\lambda = \pm\sqrt{-f'(x)}$. In other words, either we get real eigenvalues of opposite signs, or we get purely imaginary eigenvalues. There are only two possibilities for critical points, either an *unstable saddle point*, or a *stable center*. There are never any asymptotically stable points, sinks, or sources.

This page titled 17.2: Stability and Classification of Isolated Critical Points is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **8.2: Stability and Classification of Isolated Critical Points** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

17.3: Applications of Nonlinear Systems

In this section we study two very standard examples of nonlinear systems. First, we look at the nonlinear pendulum equation. We saw the pendulum equation's linearization before, but we noted it was only valid for small angles and short times. Now we find out what happens for large angles. Next, we look at the predator-prey equation, which finds various applications in modeling problems in biology, chemistry, economics, and elsewhere.

17.3.1: Pendulum

The first example we study is the pendulum equation $\theta'' + \frac{g}{L}\sin\theta = 0$. Here, θ is the angular displacement, g is the gravitational acceleration, and L is the length of the pendulum. In this equation we disregard friction, so we are talking about an idealized pendulum.

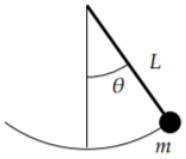


Figure 17.3.1

This equation is a conservative equation, so we can use our analysis of conservative equations from the previous section. Let us change the equation to a two-dimensional system in variables (θ, ω) by introducing the new variable ω:

$$\begin{bmatrix} \theta \\ \omega \end{bmatrix}' = \begin{bmatrix} \omega \\ -\frac{g}{L}\sin\theta \end{bmatrix}.$$

The critical points of this system are when $\omega = 0$ and $-\frac{g}{L}\sin\theta = 0$, or in other words if $\sin\theta = 0$. So the critical points are when $\omega = 0$ and θ is a multiple of π. That is, the points are $\ldots(-2\pi, 0), (-\pi, 0), (0, 0), (\pi, 0), (2\pi, 0)\ldots$. While there are infinitely many critical points, they are all isolated. Let us compute the Jacobian matrix:

$$\begin{bmatrix} \frac{\partial}{\partial\theta}\left(\omega\right) & \frac{\partial}{\partial\omega}\left(\omega\right) \\ \frac{\partial}{\partial\theta}\left(-\frac{g}{L}\sin\theta\right) & \frac{\partial}{\partial\omega}\left(-\frac{g}{L}\sin\theta\right) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{g}{L}\cos\theta & 0 \end{bmatrix}.$$

For conservative equations, there are two types of critical points. Either stable centers, or saddle points. The eigenvalues of the Jacobian matrix are $\lambda = \pm\sqrt{-\frac{g}{L}\cos\theta}$.

The eigenvalues are going to be real when $\cos\theta < 0$. This happens at the odd multiples of π. The eigenvalues are going to be purely imaginary when $\cos\theta > 0$. This happens at the even multiples of π. Therefore the system has a stable center at the points $\ldots(-2\pi, 0), (0, 0), (2\pi, 0)\ldots$, and it has an unstable saddle at the points $\ldots(-3\pi, 0), (-\pi, 0), (\pi, 0), (3\pi, 0)\ldots$. Look at the phase diagram in Figure 17.3.2, where for simplicity we let $\frac{g}{L} = 1$.

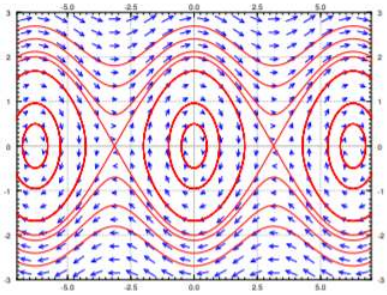


Figure 17.3.2: Phase plane diagram and some trajectories of the nonlinear pendulum equation.

In the linearized equation we have only a single critical point, the center at $(0, 0)$. Now we see more clearly what we meant when we said the linearization is good for small angles. The horizontal axis is the deflection angle. The vertical axis is the angular velocity of the pendulum. Suppose we start at $\theta = 0$ (no deflection), and we start with a small angular velocity ω. Then the

trajectory keeps going around the critical point $(0,0)$ in an approximate circle. This corresponds to short swings of the pendulum back and forth. When θ stays small, the trajectories really look like circles and hence are very close to our linearization.

When we give the pendulum a big enough push, it goes across the top and keeps spinning about its axis. This behavior corresponds to the wavy curves that do not cross the horizontal axis in the phase diagram. Let us suppose we look at the top curves, when the angular velocity ω is large and positive. Then the pendulum is going around and around its axis. The velocity is going to be large when the pendulum is near the bottom, and the velocity is the smallest when the pendulum is close to the top of its loop.

At each critical point, there is an equilibrium solution. Consider the solution $\theta = 0$; the pendulum is not moving and is hanging straight down. This is a stable place for the pendulum to be, hence this is a *stable* equilibrium.

The other type of equilibrium solution is at the unstable point, for example $\theta = \pi$. Here the pendulum is upside down. Sure you can balance the pendulum this way and it will stay, but this is an *unstable* equilibrium. Even the tiniest push will make the pendulum start swinging wildly.

See Figure 17.3.3 for a diagram. The first picture is the stable equilibrium $\theta = 0$. The second picture corresponds to those in the phase diagram around $\theta = 0$ when the angular velocity is small. The next picture is the unstable equilibrium $\theta = \pi$. The last picture corresponds to the wavy lines for large angular velocities.

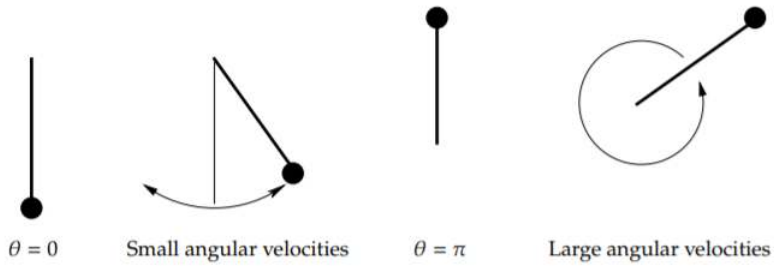


Figure 17.3.3: Various possibilities for the motion of the pendulum.

The quantity

$$\frac{1}{2}\omega^2 - \frac{g}{L}\cos\theta$$

is conserved by any solution. This is the energy or the Hamiltonian of the system.

We have a conservative equation and so (exercise) the trajectories are given by

$$\omega = \pm\sqrt{\frac{2g}{L}\cos\theta + C},$$

for various values of C. Let us look at the initial condition of $(\theta_0, 0)$, that is, we take the pendulum to angle θ_0, and just let it go (initial angular velocity 0). We plug the initial conditions into the above and solve for C to obtain

$$C = -\frac{2g}{L}\cos\theta_0.$$

Thus the expression for the trajectory is

$$\omega = \pm\sqrt{\frac{2g}{L}}\sqrt{\cos\theta - \cos\theta_0}.$$

Let us figure out the period. That is, the time it takes for the pendulum to swing back and forth. We notice that the oscillation about the origin in the phase plane is symmetric about both the θ and the ω axis. That is, in terms of θ, the time it takes from θ_0 to $-\theta_0$ is the same as it takes from $-\theta_0$ back to θ_0. Furthermore, the time it takes from $-\theta_0$ to 0 is the same as to go from 0 to θ_0. Therefore, let us find how long it takes for the pendulum to go from angle 0 to angle θ_0, which is a quarter of the full oscillation and then multiply by 4.

We figure out this time by finding $\frac{dt}{d\theta}$ and integrating from 0 to ω_0. The period is four times this integral. Let us stay in the region where ω is positive. Since $\omega = \frac{d\theta}{dt}$, inverting we get

$$\frac{dt}{d\theta} = \sqrt{\frac{L}{2g}} \frac{1}{\sqrt{\cos\theta - \cos\theta_0}}.$$

Therefore the period T is given by

$$T = 4\sqrt{\frac{L}{2g}} \int_0^{\theta_0} \frac{1}{\sqrt{\cos\theta - \cos\theta_0}}\, d\theta.$$

The integral is an improper integral, and we cannot in general evaluate it symbolically. We must resort to numerical approximation if we want to compute a particular T.

Recall from Section 2.4, the linearized equation $\theta'' + \frac{g}{L}\theta = 0$ has period

$$T_{\text{linear}} = 2\pi\sqrt{\frac{L}{g}}.$$

We plot T, T_{linear}, and the relative error $\frac{T - T_{\text{linear}}}{T}$ in Figure 17.3.4 The relative error says how far is our approximation from the real period percentage-wise. Note that T_{linear} is simply a constant, it does not change with the initial angle θ_0. The actual period T gets larger and larger as θ_0 gets larger. Notice how the relative error is small when θ_0 is small. It is still only 15% when $\theta_0 = \frac{\pi}{2}$, that is, a 90 degree angle. The error is 3.8% when starting at $\frac{\pi}{4}$, a 45 degree angle. At a 5 degree initial angle, the error is only 0.048%.

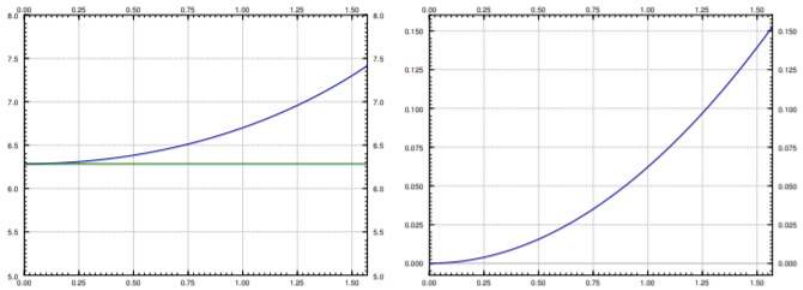


Figure 17.3.4: The plot of T and T_{linear} with $\frac{g}{L} = 1$ (left), and the plot of the relative error $\frac{T - T_{\text{linear}}}{T}$ (right), for θ_0 between 0 and $\frac{\pi}{2}$.

While it is not immediately obvious from the formula, it is true that

$$\lim_{\theta_0 \uparrow \pi} T = \infty.$$

That is, the period goes to infinity as the initial angle approaches the unstable equilibrium point. So if we put the pendulum almost upside down it may take a very long time before it gets down. This is consistent with the limiting behavior, where the exactly upside down pendulum never makes an oscillation, so we could think of that as infinite period.

17.3.2: Predator-Prey or Lotka-Volterra Systems

One of the most common simple applications of nonlinear systems are the so-called *predator-prey* or*Lotka-Volterra*[1] systems. For example, these systems arise when two species interact, one as the prey and one as the predator. It is then no surprise that the equations also see applications in economics. The system also arises in chemical reactions. In biology, this system of equations explains the natural periodic variations of populations of different species in nature. Before the application of differential equations, these periodic variations in the population baffled biologists.

We keep with the classical example of hares and foxes in a forest, it is the easiest to understand.

$$\begin{aligned} x &= \#\text{ of hares (the prey)}, \\ y &= \#\text{ of foxes (the predator)}. \end{aligned} \tag{17.3.1}$$

When there are a lot of hares, there is plenty of food for the foxes, so the fox population grows. However, when the fox population grows, the foxes eat more hares, so when there are lots of foxes, the hare population should go down, and vice versa. The Lotka–Volterra model proposes that this behavior is described by the system of equations

$$\begin{aligned} x' &= (a-by)x, \\ y' &= (cx-d)y, \end{aligned} \tag{17.3.2}$$

where a, b, c, d are some parameters that describe the interaction of the foxes and hares[2]. In this model, these are all positive numbers.

Let us analyze the idea behind this model. The model is a slightly more complicated idea based on the exponential population model. First expand,

$$x' = (a-by)x = ax - byx.$$

The hares are expected to simply grow exponentially in the absence of foxes, that is where the ax term comes in, the growth in population is proportional to the population itself. We are assuming the hares always find enough food and have enough space to reproduce. However, there is another component $-byx$, that is, the population also is decreasing proportionally to the number of foxes. Together we can write the equation as $(a-by)x$, so it is like exponential growth or decay but the constant depends on the number of foxes.

The equation for foxes is very similar, expand again

$$y' = (cx-d)y = cxy - dy.$$

The foxes need food (hares) to reproduce: the more food, the bigger the rate of growth, hence the cxy term. On the other hand, there are natural deaths in the fox population, and hence the $-dy$ term.

Without further delay, let us start with an explicit example. Suppose the equations are

$$x' = (0.4 - 0.01y)x, \qquad y' = (0.003x - 0.3)y.$$

See Figure 17.3.5 for the phase portrait. In this example it makes sense to also plot x and y as graphs with respect to time. Therefore the second graph in Figure 17.3.5 is the graph of x and y on the vertical axis (the prey x is the thinner line with taller peaks), against time on the horizontal axis. The particular solution graphed was with initial conditions of 20 foxes and 50 hares.

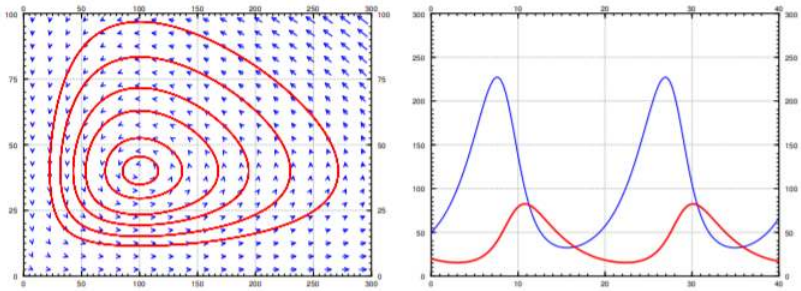


Figure 17.3.5: The phase portrait (left) and graphs of x and y for a sample solution (right).

Let us analyze what we see on the graphs. We work in the general setting rather than putting in specific numbers. We start with finding the critical points. Set $(a-by)x = 0$, and $(cx-d)y = 0$. The first equation is satisfied if either $x = 0$ or $y = \frac{a}{b}$. If $x = 0$, the second equation implies $y = 0$. If $y = \frac{a}{b}$, the second equation implies $x = \frac{d}{c}$. There are two equilibria: at $(0,0)$ when there are no animals at all, and at $(\frac{d}{c}, \frac{a}{b})$. In our specific example $x = \frac{d}{c} = 100$, and $y = \frac{a}{b} = 40$. This is the point where there are 100 hares and 40 foxes.

We compute the Jacobian matrix:

$$\begin{bmatrix} a-by & -bx \\ cy & cx-d \end{bmatrix}.$$

At the origin $(0,0)$ we get the matrix $\begin{bmatrix} a & 0 \\ 0 & -d \end{bmatrix}$, so the eigenvalues are a and $-d$, hence real and of opposite signs. So the critical point at the origin is a saddle. This makes sense. If you started with some foxes but no hares, then the foxes would go extinct, that is, you would approach the origin. If you started with no foxes and a few hares, then the hares would keep multiplying without check, and so you would go away from the origin.

OK, how about the other critical point at $(\frac{d}{c}, \frac{a}{b})$. Here the Jacobian matrix becomes

$$\begin{bmatrix} 0 & -\frac{bd}{c} \\ \frac{ac}{b} & 0 \end{bmatrix}.$$

The eigenvalues satisfy $\lambda^2 + ad = 0$. In other words, $\lambda = \pm i\sqrt{ad}$. The eigenvalues being purely imaginary, we are in the case where we cannot quite decide using only linearization. We could have a stable center, spiral sink, or a spiral source. That is, the equilibrium could be asymptotically stable, stable, or unstable. Of course I gave you a picture above that seems to imply it is a stable center. But never trust a picture only. Perhaps the oscillations are getting larger and larger, but only *very* slowly. Of course this would be bad as it would imply something will go wrong with our population sooner or later. And I only graphed a very specific example with very specific trajectories.

How can we be sure we are in the stable situation? As we said before, in the case of purely imaginary eigenvalues, we have to do a bit more work. Previously we found that for conservative systems, there was a certain quantity that was conserved on the trajectories, and hence the trajectories had to go in closed loops. We can use a similar technique here. We just have to figure out what is the conserved quantity. After some trial and error we find the constant

$$C = \frac{y^a x^d}{e^{cx+by}} = y^a x^d e^{-cx-by}$$

is conserved. Such a quantity is called the *constant of motion*. Let us check C really is a constant of motion. How do we check, you say? Well, a constant is something that does not change with time, so let us compute the derivative with respect to time:

$$C' = ay^{a-1}y'x^d e^{-cx-by} + y^a dx^{d-1}x' e^{-cx-by} + y^a x^d e^{-cx-by}(-cx' - by').$$

Our equations give us what x' and y' are so let us plug those in:

$$\begin{aligned} C' &= ay^{a-1}(cx-d)yx^d e^{-cx-by} + y^a dx^{d-1}(a-by)xe^{-cx-by} \\ &\quad + y^a x^d e^{-cx-by}\big(-c(a-by)x - b(cx-d)y\big) \\ &= y^a x^d e^{-cx-by}\Big(a(cx-d) + d(a-by) + \big(-c(a-by)x - b(cx-d)y\big)\Big) \\ &= 0. \end{aligned} \tag{17.3.3}$$

So along the trajectories C is constant. In fact, the expression $C = \frac{y^a x^d}{e^{cx+by}}$ gives us an implicit equation for the trajectories. In any case, once we have found this constant of motion, it must be true that the trajectories are simple curves, that is, the level curves of $\frac{y^a x^d}{e^{cx+by}}$. It turns out, the critical point at $(\frac{d}{c}, \frac{a}{b})$ is a maximum for C (left as an exercise). So $(\frac{d}{c}, \frac{a}{b})$ is a stable equilibrium point, and we do not have to worry about the foxes and hares going extinct or their populations exploding.

One blemish on this wonderful model is that the number of foxes and hares are discrete quantities and we are modeling with continuous variables. Our model has no problem with there being 0.1 fox in the forest for example, while in reality that makes no sense. The approximation is a reasonable one as long as the number of foxes and hares are large, but it does not make much sense for small numbers. One must be careful in interpreting any results from such a model.

An interesting consequence (perhaps counterintuitive) of this model is that adding animals to the forest might lead to extinction, because the variations will get too big, and one of the populations will get close to zero. For example, suppose there are 20 foxes and 50 hares as before, but now we bring in more foxes, bringing their number to 200. If we run the computation, we find the number of hares will plummet to just slightly more than 1 hare in the whole forest. In reality that most likely means the hares die out, and then the foxes will die out as well as they will have nothing to eat.

Showing that a system of equations has a stable solution can be a very difficult problem. When Isaac Newton put forth his laws of planetary motions, he proved that a single planet orbiting a single sun is a stable system. But any solar system with more than 1 planet proved very difficult indeed. In fact, such a system behaves chaotically (see Section 8.5), meaning small changes in initial conditions lead to very different long-term outcomes. From numerical experimentation and measurements, we know the earth will not fly out into the empty space or crash into the sun, for at least some millions of years or so. But we do not know what happens beyond that.

17.3.3: Footnotes

[1] Named for the American mathematician, chemist, and statistician Alfred James Lotka (1880–1949) and the Italian mathematician and physicist Vito Volterra (1860–1940).

[2] This interaction does not end well for the hare.

This page titled 17.3: Applications of Nonlinear Systems is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **8.3: Applications of Nonlinear Systems** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

17.4: Limit cycles

For nonlinear systems, trajectories do not simply need to approach or leave a single point. They may in fact approach a larger set, such as a circle or another closed curve.

Example 17.4.1: The Van der Pol oscillator

The *Van der Pol oscillator*[1] is the following equation

$$x'' - \mu(1-x^2)x' + x = 0,$$

where μ is some positive constant. The Van der Pol oscillator originated with electrical circuits, but finds applications in diverse fields such as biology, seismology, and other physical sciences.

For simplicity, let us use $\mu = 1$. A phase diagram is given in the left hand plot in Figure 17.4.1. Notice how the trajectories seem to very quickly settle on a closed curve. On the right hand plot we have the plot of a single solution for $t = 0$ to $t = 30$ with initial conditions $x(0) = 0.1$ and $x'(0) = 0.1$. Notice how the solution quickly tends to a periodic solution.

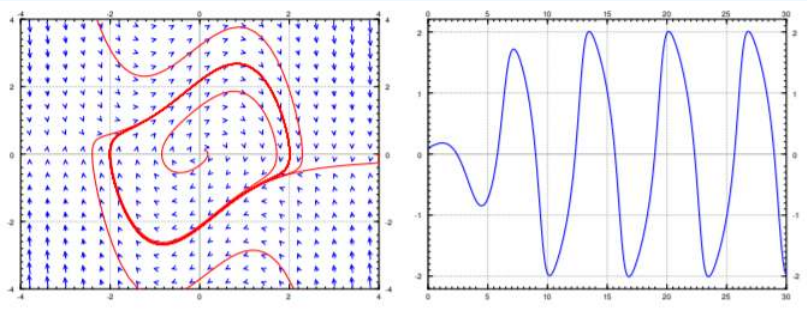


Figure 17.4.1: The phase portrait (left) and a graph of a sample solution of the Van der Pol oscillator.

The Van der Pol oscillator is an example of so-called **relaxation oscillation.** The word relaxation comes from the sudden jump (the very steep part of the solution). For larger μ the steep part becomes even more pronounced, for small μ the limit cycle looks more like a circle. In fact setting $\mu = 0$, we get $x'' + x = 0$, which is a linear system with a center and all trajectories become circles.

The closed curve in the phase portrait above is called a **limit cycle**. A limit cycle is a closed trajectory such that at least one other trajectory spirals into it (or spirals out of it). If all trajectories that start near the limit cycle spiral into it, the limit cycle is called **asymptotically stable**. The limit cycle in the Van der Pol oscillator is asymptotically stable.

Given a limit cycle on an autonomous system, any solution that starts on it is periodic. In fact, this is true for any trajectory that is a closed curve (a so-called *closed trajectory*). Such a curve is called a *periodic orbit.* More precisely, if $\big(x(t), y(t)\big)$ is a solution such that for some t_0 the point $\big(x(t_0), y(t_0)\big)$ lies on a periodic orbit, then both $x(t)$ and $y(t)$ are periodic functions (with the same period). That is, there is some number P such that $x(t) = x(t+P)$ and $y(t) = y(t+P)$.

Consider the system

$$x' = f(x,y), \quad y' = g(x,y), \tag{17.4.1}$$

where the functions f and g have continuous derivatives in some region R in the plane.

Theorem 17.4.1: Poincarè-Bendixson Theorem

Suppose R is a closed bounded region (a region in the plane that includes its boundary and does not have points arbitrarily far from the origin). Suppose $\big(x(t), y(t)\big)$ is a solution of (17.4.1) in R that exists for all $t \geq t_0$. Then either the solution is a periodic function, or the solution spirals towards a periodic solution in R.

The main point of the theorem is that if you find one solution that exists for all t large enough (that is, as t goes to infinity) and stays within a bounded region, then you have found either a periodic orbit, or a solution that spirals towards a limit cycle or tends to a critical point. That is, in the long term, the behavior is very close to a periodic function. Note that a constant solution at a critical point is periodic (with any period). The theorem is more a qualitative statement rather than something to help us in computations. In practice it is hard to find analytic solutions and so hard to show rigorously that they exist for all time. But if we think the solution exists we numerically solve for a large time to approximate the limit cycle. Another caveat is that the theorem only works in two dimensions. In three dimensions and higher, there is simply too much room.

The theorem applies to all solutions in the Van der Pol oscillator. Solutions that start at any point except the origin $(0,0)$ will tend to the periodic solution around the limit cycle, and if the initial condition of $(0,0)$ will lead to the constant solution $x=0$, $y=0$.

Example 17.4.2

Consider

$$x' = y + (x^2+y^2-1)^2 x, \qquad y' = -x + (x^2+y^2-1)^2 y.$$

A vector field along with solutions with initial conditions $(1.02, 0)$, $(0.9, 0)$, and $(0.1, 0)$ are drawn in Figure 17.4.2

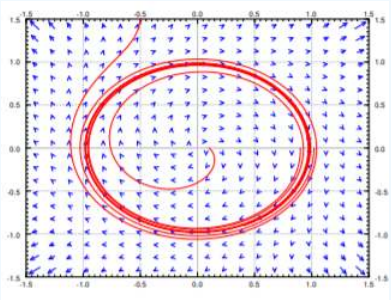


Figure 17.4.2: Unstable limit cycle example.

Notice that points on the unit circle (distance one from the origin) satisfy $x^2+y^2-1=0$. And $x(t)=\sin(t)$, $y=\cos(t)$ is a solution of the system. Therefore we have a closed trajectory. For points off the unit circle, the second term in x' pushes the solution further away from the y-axis than the system $x'=y$, $y'=-x$, and y' pushes the solution further away from the x-axis than the linear system $x'=y$, $y'=-x$. In other words for all other initial conditions the trajectory will spiral out.

This means that for initial conditions inside the unit circle, the solution spirals out towards the periodic solution on the unit circle, and for initial conditions outside the unit circle the solutions spiral off towards infinity. Therefore the unit circle is a limit cycle, but not an asymptotically stable one. The Poincaré–Bendixson Theorem applies to the initial points inside the unit circle, as those solutions stay bounded, but not to those outside, as those solutions go off to infinity.

A very similar analysis applies to the system

$$x' = y + (x^2+y^2-1)x, \qquad y' = -x + (x^2+y^2-1)y.$$

We still obtain a closed trajectory on the unit circle, and points outside the unit circle spiral out to infinity, but now points inside the unit circle spiral towards the critical point at the origin. So this system does not have a limit cycle, even though it has a closed trajectory.

Due to the Picard theorem (3.1.1) we find that no matter where we are in the plane we can always find a solution a little bit further in time, as long as f and g have continuous derivatives. So if we find a closed trajectory in an autonomous system, then for every initial point inside the closed trajectory, the solution will exist for all time and it will stay bounded (it will stay inside the closed trajectory). So the moment we found the solution above going around the unit circle, we knew that for every initial point inside the circle, the solution exists for all time and the Poincaré–Bendixson theorem applies.

Let us next look for conditions when limit cycles (or periodic orbits) do not exist. We assume the equation (17.4.1) is defined on a *simply connected region*, that is, a region with no holes we can go around. For example the entire plane is a simply connected region, and so is the inside of the unit disc. However, the entire plane minus a point is not a simply connected domain as it has a at the origin.

Theorem 17.4.2: Bendixson-Dulac Theorem

Suppose f and g are defined in a simply connected region R. If the expression[4]

$$\frac{\partial f}{\partial x}+\frac{\partial g}{\partial y}$$

is either always positive or always negative on R (except perhaps a small set such as on isolated points or curves) then the system (17.4.1) has no closed trajectory inside R.

The theorem gives us a way of ruling out the existence of a closed trajectory, and hence a way of ruling out limit cycles. The exception about points or lines really means that we can allow the expression to be zero at a few points, or perhaps on a curve, but not on any larger set.

Example 17.4.3

Let us look at $x'=y+y^2e^x$, $y'=x$ in the entire plane (see Example 8.2.2.) The entire plane is simply connected and so we can apply the theorem. We compute $\frac{\partial f}{\partial x}+\frac{\partial g}{\partial y}=y^2e^x+0$. The function y^2e^x is always positive except on the line $y=0$. Therefore, via the theorem, the system has no closed trajectories.

In some books (or the internet) the theorem is not stated carefully and it concludes there are no periodic solutions. That is not quite right. The above example has two critical points and hence it has constant solutions, and constant functions are periodic. The conclusion of the theorem should be that there exist no trajectories that form closed curves. Another way to state the conclusion of the theorem would be to say that there exist no nonconstant periodic solutions that stay in R.

Example 17.4.4

Let us look at a somewhat more complicated example. Take the system $x'=-y-x^2$, $y'=-x+y^2$ (see Example 8.2.1). We compute $\frac{\partial f}{\partial x}+\frac{\partial g}{\partial y}=2x+2y$. This expression takes on both signs, so if we are talking about the whole plane we cannot simply apply the theorem. However, we could apply it on the set where $x+y>0$. Via the theorem, there is no closed trajectory in that set. Similarly, there is no closed trajectory in the set $x+y<0$. We cannot conclude (yet) that there is no closed trajectory in the entire plane. Perhaps half of it is in the set where $x+y>0$ and the other half is in the set where $x+y<0$.

The key is to look at the set $x+y=0$, or $x=-y$. Let us make a substitution $x=z$ and $y=-z$ (so that $x=-y$). Both equations become $z'=z-z^2$. So any solution of $z'=z-z^2$, gives us a solution $x(t)=z(t)$, $y(t)=-z(t)$. In particular, any solution that starts out on the line $x+y=0$, stays on the line $x+y=0$. In other words, there cannot be a closed trajectory that starts on the set where $x+y>0$ and goes through the set where $x+y<0$, as it would have to pass through $x+y=0$.

Example 17.4.5

Consider $x'=y+(x^2+y^2-1)x$, $y'=-x+(x^2+y^2-1)y$, and consider the region R given by $x^2+y^2>\frac{1}{2}$. That is, R is the region outside a circle of radius $\frac{1}{\sqrt{2}}$ centered at the origin. Then there is a closed trajectory in R, namely $x=\cos(t)$, $y=\sin(t)$. Furthermore,

$$\frac{\partial f}{\partial x}+\frac{\partial g}{\partial x}=4x^2+4y^2-2,$$

which is always positive on R. So what is going on? The Bendixson–Dulac theorem does not apply since the region R is not simply connected—it has a hole, the circle we cut out!

17.4.1: Footnotes

[1] Named for the Dutch physicist Balthasar van der Pol (1889–1959).

[2] Ivar Otto Bendixson (1861–1935) was a Swedish mathematician.

[3] Henri Dulac (1870–1955) was a French mathematician.

[4] Usually the expression in the Bendixson–Dulac Theorem is $\frac{\partial(\varphi f)}{\partial x} + \frac{\partial(\varphi g)}{\varphi y}$ for some continuously differentiable function φ. For simplicity, let us just consider the case $\varphi = 1$.

This page titled 17.4: Limit cycles is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **8.4: Limit cycles** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

17.5: Chaos

You have surely heard the story about the flap of a butterfly wing in the Amazon causing hurricanes in the North Atlantic. In a prior section, we mentioned that a small change in initial conditions of the planets can lead to very different configuration of the planets in the long term. These are examples of chaotic systems. Mathematical chaos is not really chaos, there is precise order behind the scenes. Everything is still deterministic. However a chaotic system is extremely sensitive to initial conditions. This also means even small errors induced via numerical approximation create large errors very quickly, so it is almost impossible to numerically approximate for long times. This is large part of the trouble as chaotic systems cannot be in general solved analytically.

Take the weather for example. As a small change in the initial conditions (the temperature at every point of the atmosphere for example) produces drastically different predictions in relatively short time, we cannot accurately predict weather. This is because we do not actually know the exact initial conditions, we measure temperatures at a few points with some error and then we somehow estimate what is in between. There is no way we can accurately measure the effects of every butterfly wing. Then we will solve numerically introducing new errors. That is why you should not trust weather prediction more than a few days out.

The idea of chaotic behavior was first noticed by Edward Lorenz in the 1960s when trying to model thermally induced air convection (movement). The equations Lorentz was looking at form the relatively simple looking system:

$$x' = -10x + 10y, \quad y' = 28x - y - xz, \quad z' = -\frac{8}{3}z + xy.$$

A small change in the initial conditions yield a very different solution after a reasonably short time.



Figure 17.5.1

A very simple example the reader can experiment with, which displays chaotic behavior, is a double pendulum. The equations that govern this system are somewhat complicated and their derivation is quite tedious, so we will not bother to write them down. The idea is to put a pendulum on the end of another pendulum. If you look at the movement of the bottom mass, the movement will appear chaotic. This type of system is a basis for a whole number of office novelty desk toys. It is very simple to build a version. Take a piece of a string, and tie two heavy nuts at different points of the string; one at the end, and one a bit above. Now give the bottom nut a little push, as long as the swings are not too big and the string stays tight, you have a double pendulum system.

17.5.1: Duffing Equation and Strange Attractors

Let us study the so-called *Duffing equation*:

$$x'' + ax' + bx + cx^3 = C\cos(\omega t).$$

Here a, b, c, C, and ω are constants. You will recognize that except for the cx^3 term, this equation looks like a forced mass-spring system. The cx^3 term comes up when the spring does not exactly obey Hooke's law (which no real-world spring actually does obey exactly). When c is not zero, the equation does not have a nice closed form solution, so we have to resort to numerical solutions as is usual for nonlinear systems. Not all choices of constants and initial conditions will exhibit chaotic behavior. Let us study

$$x'' + 0.05x' + x^3 = 8\cos(t).$$

The equation is not autonomous, so we will not be able to draw the vector field in the phase plane. We can still draw the trajectories however. In Figure $17.5.2$we plot trajectories for t going from 0 to 15, for two very close initial conditions $(2, 3)$ and $(2, 2.9)$, and also the solutions in the (x, t) space. The two trajectories are close at first, but after a while diverge significantly. This sensitivity to initial conditions is precisely what we mean by the system behaving chaotically.

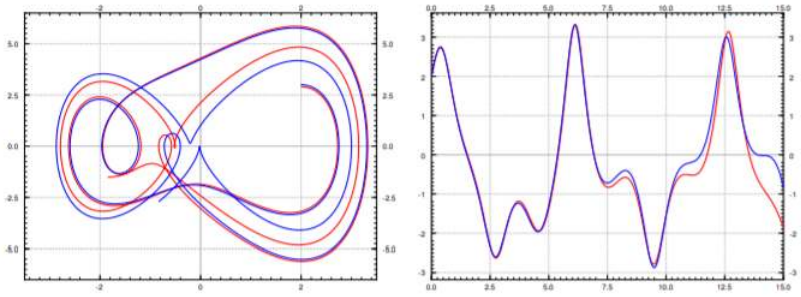


Figure 17.5.2: On left, two trajectories in phase space for $0 \le t \le 15$, for the Duffing equation one with initial conditions $(2, 3)$ and the other with $(2, 2.9)$. On right the two solutions in (x, t)-space.

Let us see the long term behavior. In Figure 17.5.3 we plot the behavior of the system for initial conditions $(2, 3)$, but for much longer period of time. Note that for this period of time it was necessary to use a ridiculously large number of steps[1] in the numerical algorithm used to produce the graph, as even small errors quickly propagate. From the graph it is hard to see any particular pattern in the shape of the solution except that it seems to oscillate, but each oscillation appears quite unique. The oscillation is expected due to the forcing term.

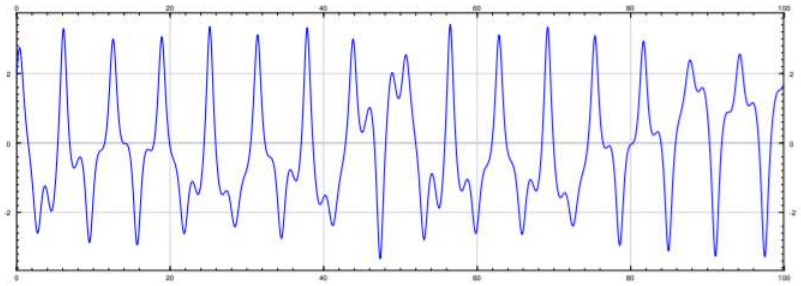


Figure 17.5.3: The solution to the given Duffing equation for t from 0 to 100.

In general it is very difficult to analyze chaotic systems, or to find the order behind the madness, but let us try to do something that we did for the standard mass-spring system. One way we analyzed what happens is that we figured out what was the long term behavior (not dependent on initial conditions). From the figure above it is clear that we will not get a nice description of the long term behavior, but perhaps we can figure out some order to what happens on each "oscillation" and what do these oscillations have in common.

The concept we will explore is that of a *Poincaré section*[2]. Instead of looking at t in a certain interval, we will look at where the system is at a certain sequence of points in time. Imagine flashing a strobe at a certain fixed frequency and drawing the points where the solution is during the flashes. The right strobing frequency depends on the system in question. The correct frequency to use for the forced Duffing equation (and other similar systems) is the frequency of the forcing term. For the Duffing equation above, find a solution $\big(x(t), y(t)\big)$, and look at the points

$$\big(x(0), y(0)\big), \quad \big(x(2\pi), y(2\pi)\big), \quad \big(x(4\pi), y(4\pi)\big), \quad \big(x(6\pi), y(6\pi)\big), \quad \ldots$$

As we are really not interested in the transient part of the solution, that is, the part of the solution that depends on the initial condition we skip some number of steps in the beginning. For example, we might skip the first 100 such steps and start plotting points at $t = 100(2\pi)$, that is

$$\big(x(200\pi), y(200\pi)\big), \quad \big(x(202\pi), y(202\pi)\big), \quad \big(x(204\pi), y(204\pi)\big), \quad \big(x(206\pi), y(206\pi)\big), \quad \ldots$$

The plot of these points is the Poincaré section. After plotting enough points, a curious pattern emerges in Figure 17.5.4 (the left hand picture), a so-called *strange attractor*.

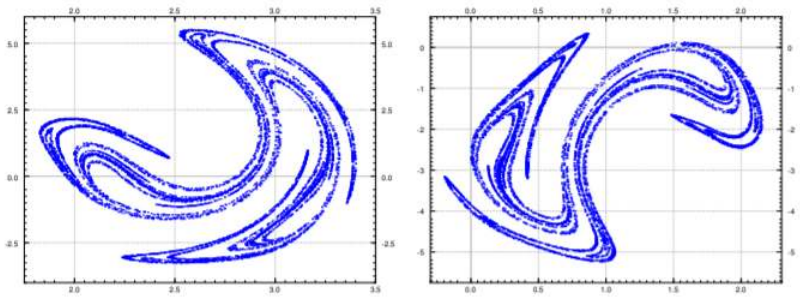


Figure 17.5.4: Strange attractor. The left plot is with no phase shift, the right plot has phase shift $\frac{\pi}{4}$.

If we have a sequence of points, then an *attractor* is a set towards which the points in the sequence eventually get closer and closer to, that is, they are attracted. The Poincaré section above is not really the attractor itself, but as the points are very close to it, we can see its shape. The strange attractor in the figure is a very complicated set, and it in fact has fractal structure, that is, if you would zoom in as far as you want, you would keep seeing the same complicated structure.

The initial condition does not really make any difference. If we started with different initial condition, the points would eventually gravitate towards the attractor, and so as long as we throw away the first few points, we always get the same picture.

An amazing thing is that a chaotic system such as the Duffing equation is not random at all. There is a very complicated order to it, and the strange attractor says something about this order. We cannot quite say what state the system will be in eventually, but given a fixed strobing frequency we can narrow it down to the points on the attractor.

If you would use a phase shift, for example $\frac{\pi}{4}$, and look at the times

$$\frac{\pi}{4}, \quad 2\pi+\frac{\pi}{4}, \quad 4\pi+\frac{\pi}{4}, \quad 6\pi+\frac{\pi}{4}, \quad \ldots$$

you would obtain a slightly different looking attractor. The picture is the right hand side of Figure 17.5.4 It is as if we had rotated, distorted slightly, and then moved the original. Therefore for each phase shift you can find the set of points towards which the system periodically keeps coming back to.

You should study the pictures and notice especially the scales---where are these attractors located in the phase plane. Notice the regions where the strange attractor lives and compare it to the plot of the trajectories in Figure 17.5.2

Let us compare the discussion in this section to the discussion in Section 2.6 about forced oscillations. Take the equation

$$x''+2px'+\omega_0^2 x = \frac{F_0}{m}\cos(\omega t).$$

This is like the Duffing equation, but with no x^3 term. The steady periodic solution is of the form

$$x = C\cos(\omega t+\gamma).$$

Strobing using the frequency ω we would obtain a single point in the phase space. So the attractor in this setting is a single point---an expected result as the system is not chaotic. In fact it was the opposite of chaotic. Any difference induced by the initial conditions dies away very quickly, and we settle into always the same steady periodic motion.

17.5.2: Lorenz System

In two dimensions to have the kind of chaotic behavior we are looking for, we have to study forced, or non-autonomous, systems such as the Duffing equation. Due to the Poincaré-Bendoxson Theorem, if an autonomous two-dimensional system has a solution that exists for all time in the future and does not go towards infinity, then we obtain a limit cycle or a closed trajectory. Hardly the chaotic behavior we are looking for.

In three dimensions even autonomous systems can be chaotic. Let us very briefly return to the *Lorenz system*

$$x' = -10x+10y, \qquad y' = 28x-y-xz, \qquad z' = -\frac{8}{3}z+xy.$$

The Lorenz system is an autonomous system in three dimensions exhibiting chaotic behavior. See the Figure 17.5.5 for a sample trajectory, which is now a curve in three-dimensional space.

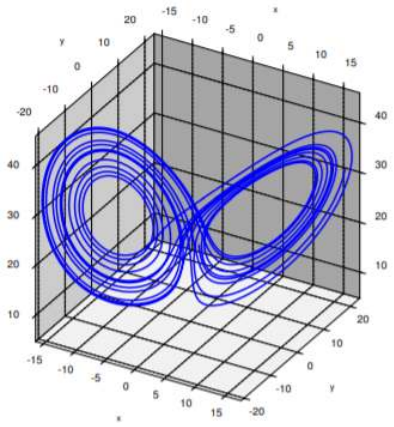


Figure 17.5.5: A trajectory in the Lorenz system.

The solutions will tend to an *attractor* in space, the so-called *Lorenz attractor*. In this case no strobing is necessary. Again we cannot quite see the attractor itself, but if we try to follow a solution for long enough, as in the figure, we will get a pretty good picture of what the attractor looks like. The Lorenz attractor is also a strange attractor and has a complicated fractal structure. And, just as for the Duffing equation, what we want to draw is not the whole trajectory, but start drawing the trajectory after a while, once it is close to the attractor.

The path is not just a repeating figure-eight. The trajectory will spin some seemingly random number of times on the left, then spin a number of times on the right, and so on. As this system arose in weather prediction, one can perhaps imagine a few days of warm weather and then a few days of cold weather, where it is not easy to predict when the weather will change, just as it is not really easy to predict far in advance when the solution will jump onto the other side. See Figure 17.5.6 for a plot of the x component of the solution drawn above. A negative x corresponds to the left "loop" and a positive x corresponds to the right "loop".

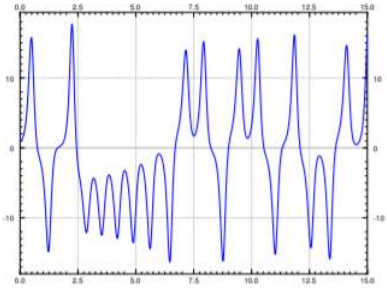


Figure 17.5.6: Graph of the $x(t)$ component of the solution.

Most of the mathematics we studied in this book is quite classical and well understood. On the other hand, chaos, including the Lorenz system, continues to be the subject of current research. Furthermore, chaos has found applications not just in the sciences, but also in art.

17.5.3: Footnotes

[1] In fact for reference, 30,000 steps were used with the Runge–Kutta algorithm, see exercises in Section 1.7.

[2] Named for the French polymath Jules Henri Poincaré (1854-1912).

This page titled 17.5: Chaos is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **8.5: Chaos** by Jiří Lebl is licensed CC BY-SA 4.0. Original source: https://www.jirka.org/diffyqs.

17.6: Appendix B- Table of Laplace Transforms

The function u is the Heaviside function, δ is the Dirac delta function, and

$$\Gamma(t)=\int_0^\infty e^{-\tau}\tau^{t-1}\,d\tau, \qquad \operatorname{erf}(t)=\frac{2}{\sqrt{\pi}}\int_0^t e^{-\tau^2}\,d\tau, \qquad \operatorname{erfc}(t)=1-\operatorname{erf}(t). \tag{17.6.1}$$

Table 17.6.1

$f(t)$	$F(s)=\mathcal{L}\{f(t)\}=\int_0^\infty e^{-st}f(t)dt$
C	$\frac{C}{s}$
t	$\frac{1}{s^2}$
t^2	$\frac{2}{s^3}$
t^n	$\frac{n!}{s^{n+1}}$
$t^p \quad (p>0)$	$\frac{\Gamma(p+1)}{s^{p+1}}$
e^{-at}	$\frac{1}{s+a}$
$\sin(\omega t)$	$\frac{\omega}{s^2+\omega^2}$
$\cos(\omega t)$	$\frac{s}{s^2+\omega^2}$
$\sinh(\omega t)$	$\frac{\omega}{s^2-\omega^2}$
$\cosh(\omega t)$	$\frac{s}{s^2-\omega^2}$
$u(t-a)$	$\frac{e^{-as}}{s}$
$\delta(t)$	1
$\delta(t-a)$	e^{-as}
$\operatorname{erf}\left(\frac{t}{2a}\right)$	$\frac{1}{s}e^{(as)^2}\operatorname{erfc}(as)$
$\frac{1}{\sqrt{\pi t}}\exp\left(\frac{-a^2}{4t}\right) \quad (a\geq 0)$	$\frac{e^{-as}}{\sqrt{s}}$
$\frac{1}{\sqrt{\pi t}}-ae^{a^2t}\operatorname{erfc}(a\sqrt{t}) \quad (a>0)$	$\frac{1}{\sqrt{s}+a}$
$af(t)+bg(t)$	$aF(s)+bG(s)$
$f(at) \quad (a>0)$	$\frac{1}{a}F\left(\frac{s}{a}\right)$
$f(t-a)u(t-a)$	$e^{-as}F(s)$
$e^{-at}f(t)$	$F(s+a)$
$g'(t)$	$sG(s)-g(0)$
$g''(t)$	$s^2G(s)-sg(0)-g'(0)$
$g'''(t)$	$s^3G(s)-s^2g(0)-sg'(0)-g''(0)$
$g^{(n)}(t)$	$s^nG(s)-s^{n-1}g(0)-\cdots-g^{(n-1)}(0)$
$(f*g)(t)=\int_0^t f(\tau)g(t-\tau)d\tau$	$F(s)G(s)$
$tf(t)$	$-F'(s)$
$t^nf(t)$	$(-1)^nF^{(n)}(s)$
$\int_0^t f(\tau)d\tau$	$\frac{1}{s}F(s)$

$f(t)$	$F(s)=\mathcal{L}\{f(t)\}=\int_0^\infty e^{-st}f(t)dt$
$\frac{f(t)}{t}$	$\int_s^\infty F(\sigma)d\sigma$

17.6: Appendix B- Table of Laplace Transforms is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

17.E: Nonlinear Equations (Exercises)

17.E.1: 8.1: Linearization, critical points, and equilibria

Exercise 17.E.8.1.1

Sketch the phase plane vector field for:

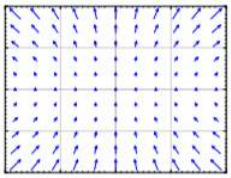
a. $x' = x^2,\ \ y' = y^2$,
b. $x' = (x-y)^2,\ \ y' = -x$,
c. $x' = e^y,\ \ y' = e^x$.

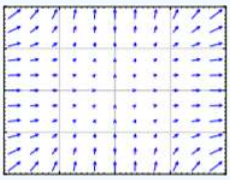
Exercise 17.E.8.1.2

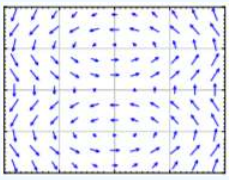
Match systems

1. $x' = x^2$, $y' = y^2$,
2. $x' = xy$, $y' = 1 + y^2$,
3. $x' = \sin(\pi y)$, $y' = x$,

to the vector fields below. Justify.

a. 

b. 

c. 

Exercise 17.E.8.1.3

Find the critical points and linearizations of the following systems.

a. $x' = x^2 - y^2$, $y' = x^2 + y^2 - 1$,
b. $x' = -y$, $y' = 3x + yx^2$,
c. $x' = x^2 + y$, $y' = y^2 + x$.

Exercise 17.E.8.1.4

For the following systems, verify they have critical point at $(0,0)$, and find the linearization at $(0,0)$.

a. $x' = x + 2y + x^2 - y^2$, $y' = 2y - x^2$
b. $x' = -y$, $y' = x - y^3$
c. $x' = ax + by + f(x,y)$, $y' = cx + dy + g(x,y)$, where $f(0,0) = 0$, $g(0,0) = 0$, and all first partial derivatives of f and g are also zero at $(0,0)$, that is, $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = \frac{\partial g}{\partial x}(0,0) = \frac{\partial g}{\partial y}(0,0) = 0$.

Exercise 17.E.8.1.5

Take $x' = (x-y)^2$, $y' = (x+y)^2$.

a. Find the set of critical points.
b. Sketch a phase diagram and describe the behavior near the critical point(s).
c. Find the linearization. Is it helpful in understanding the system?

Exercise 17.E.8.1.6

Take $x' = x^2$, $y' = x^3$.

a. Find the set of critical points.
b. Sketch a phase diagram and describe the behavior near the critical point(s).
c. Find the linearization. Is it helpful in understanding the system?

Exercise 17.E.8.1.7

Find the critical points and linearizations of the following systems.

a. $x' = \sin(\pi y) + (x-1)^2$, $y' = y^2 - y$,
b. $x' = x + y + y^2$, $y' = x$,
c. $x' = (x-1)^2 + y$, $y' = x^2 + y$.

Answer

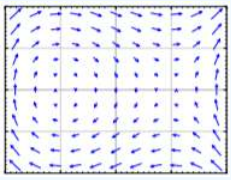
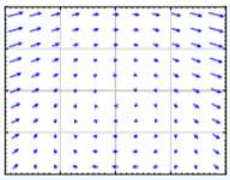
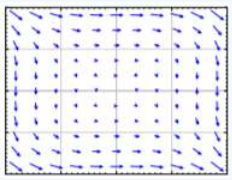
a. Critical points $(0,0)$ and $(0,1)$. At $(0,0)$ using $u = x$, $v = y$ the linearization is $u' = -2u - \left(\frac{1}{\pi}\right) v$, $v' = -v$. At $(0,1)$ using $u = x$, $v = y - 1$ the linearization is $u' = -2u + \left(\frac{1}{\pi}\right) v$ $v' = v$.
b. Critical point $(0,0)$. Using $u = x$, $v = y$ the linearization is $u' = u + v$, $v' = u$.
c. Critical point $\left(\frac{1}{2}, -\frac{1}{4}\right)$. Using $u = x - \frac{1}{2}$, $v = y + \frac{1}{4}$ the linearization is $u' = -u + v$, $v' = u + v$.

Exercise 17.E.8.1.8

Match systems

a. $x' = y^2$, $y' = -x^2$,
b. $x' = y$, $y' = (x-1)(x+1)$,
c. $x' = y + x^2$, $y' = -x$,

to the vector fields below. Justify.

a. 
b. 
c. 

Answer

a. is c),
b. is a),
c. is b)

Exercise 17.E.8.1.9

The idea of critical points and linearization works in higher dimensions as well. You simply make the Jacobian matrix bigger by adding more functions and more variables. For the following system of 3 equations find the critical points and their linearizations:

$x' = x + z^2, \quad y' = z^2 - y, \quad z' = z + x^2.$

Answer

Critical points are $(0,0,0)$, and $(-1,1,-1)$. The linearization at the origin using variables $u = x$, $v = y$, $w = z$ is $u' = u$, $v' = -v$, $z' = w$. The linearization at the point $(-1,1,-1)$ using variables $u = x+1$, $v = y-1$, $w = z+1$ is $u' = u-2$, $v' = -v-2w$, $w' = w-2u$.

Exercise 17.E.8.1.10

Any two-dimensional non-autonomous system $x' = f(x,y,t)$, $y' = g(x,y,t)$ can be written as a three-dimensional autonomous system (three equations). Write down this autonomous system using the variables u, v, w.

Answer

$u' = f(u,v,w)$, $v' = g(u,v,w)$, $w' = 1$.

17.E.2: 8.2: Stability and classification of isolated critical points

Exercise 17.E.8.2.1

For the systems below, find and classify the critical points, also indicate if the equilibria are stable, asymptotically stable, or unstable.

a. $x' = -x + 3x^2, y' = -y$
b. $x' = x^2 + y^2 - 1$, $y' = x$
c. $x' = ye^x$, $y' = y - x + y^2$

Exercise 17.E.8.2.2

Find the implicit equations of the trajectories of the following conservative systems. Next find their critical points (if any) and classify them.

a. $x'' + x + x^3 = 0$
b. $\theta'' + \sin\theta = 0$
c. $z'' + (z-1)(z+1) = 0$
d. $x'' + x^2 + 1 = 0$

Exercise 17.E.8.2.3

Find and classify the critical point(s) of $x' = -x^2$, $y' = -y^2$.

Exercise 17.E.8.2.4

Suppose $x' = -xy, y' = x^2 - 1 - y$.

a. Show there are two spiral sinks at $(-1, 0)$ and $(1, 0)$.
b. For any initial point of the form $(0, y_0)$,find what is the trajectory.
c. Can a trajectory starting at (x_0, y_0) where $x_0 > 0$ spiral into the critical point at $(-1, 0)$? Why or why not?

Exercise 17.E.8.2.5

In the example $x' = y, y' = y^3 - x$ show that for any trajectory, the distance from the origin is an increasing function. Conclude that the origin behaves like is a spiral source. Hint: Consider $f(t) = \big(x(t)\big)^2 + \big(y(t)\big)^2$ and show it has positive derivative.

Exercise 17.E.8.2.6

Suppose f is always positive. Find the trajectories of $x'' + f(x') = 0$. Are there any critical points?

Exercise 17.E.8.2.7

Suppose that $x' = f(x, y), y' = g(x, y)$. Suppose that $g(x, y) > 1$ for all x and y. Are there any critical points? What can we say about the trajectories at t goes to infinity?

Exercise 17.E.8.2.8

For the systems below, find and classify the critical points.

a. $x' = -x + x^2, y' = y$
b. $x' = y - y^2 - x, y' = -x$
c. $x' = xy, y' = x + y - 1$

Answer

a. $(0, 0)$: saddle (unstable), $(1, 0)$: source (unstable),
b. $(0, 0)$: spiral sink (asymptotically stable), $(0, 1)$: saddle (unstable),
c. $(1, 0)$: saddle (unstable), $(0, 1)$: source (unstable)

Exercise 17.E.8.2.9

Find the implicit equations of the trajectories of the following conservative systems. Next find their critical points (if any) and classify them.

a. $x'' + x^2 = 4$
b. $x'' + e^x = 0$
c. $x'' + (x + 1)e^x = 0$

Answer

a. $\frac{1}{2}y^2 + \frac{1}{3}x^3 - 4x = C$, critical points: $(-2, 0)$, an unstable saddle, and $(2, 0)$, a stable center.
b. $\frac{1}{2}y^2 + e^x = C$, no critical points.
c. $\frac{1}{2}y^2 + xe^x = C$, critical point at $(-1, 0)$ is a stable center.

Exercise 17.E.8.2.10

The conservative system $x'' + x^3 = 0$ is not almost linear. Classify its critical point(s) nonetheless.

Answer

Critical point at $(0,0)$. Trajectories are $y = \pm\sqrt{2C - \left(\frac{1}{2}\right)x^4}$, for $C > 0$, these give closed curves around the origin, so the critical point is a stable center.

Exercise 17.E.8.2.11

Derive an analogous classification of critical points for equations in one dimension, such as $x' = f(x)$ based on the derivative. A point x_0 is critical when $f(x_0) = 0$ and almost linear if in addition $f'(x_0) \neq 0$. Figure out if the critical point is stable or unstable depending on the sign of $f'(x_0)$. Explain. Hint: see Section 1.6.

Answer

A critical point x_0 is stable if $f'(x_0) < 0$ and unstable when $f'(x_0) < 0$.

17.E.3: 8.3: Applications of nonlinear systems

Exercise 17.E.8.3.1

Take the *damped nonlinear pendulum equation* $\theta'' + \mu\theta' + (\frac{g}{L})\sin\theta = 0$ for some $\mu > 0$ (that is, there is some friction).

a. Suppose $\mu = 1$ and $\frac{g}{L} = 1$ for simplicity, find and classify the critical points.
b. Do the same for any $\mu > 0$ and any g and L, but such that the damping is small, in particular, $\mu^2 < 4(\frac{g}{L})$.
c. Explain what your findings mean, and if it agrees with what you expect in reality.

Exercise 17.E.8.3.2

Suppose the hares do not grow exponentially, but logistically. In particular consider

$$x' = (0.4 - 0.01y)x - \gamma x^2, \quad y' = (0.003x - 0.3)y. \tag{17.E.1}$$

For the following two values of γ, find and classify all the critical points in the positive quadrant, that is, for $x \geq 0$ and $y \geq 0$. Then sketch the phase diagram. Discuss the implication for the long term behavior of the population.

a. $\gamma = 0.001$,
b. $\gamma = 0.01$.

Exercise 17.E.8.3.3

a. Suppose x and y are positive variables. Show $\frac{yx}{e^{x+y}}$ attains a maximum at $(1,1)$.
b. Suppose a, b, c, d are positive constants, and also suppose x and y are positive variables. Show $\frac{y^a x^d}{e^{cx+by}}$ attains a maximum at $(\frac{d}{c}, \frac{a}{b})$.

Exercise 17.E.8.3.4

Suppose that for the pendulum equation we take a trajectory giving the spinning-around motion, for example $\omega = \sqrt{\frac{2g}{L}\cos\theta + \frac{2g}{L} + \omega_0^2}$. This is the trajectory where the lowest angular velocity is ω_0^2. Find an integral expression for how long it takes the pendulum to go all the way around.

Exercise 17.E.8.3.5: (challenging)

Take the pendulum, suppose the initial position is $\theta = 0$.

a. Find the expression for ω giving the trajectory with initial condition $(0, \omega_0)$. Hint: Figure out what C should be in terms of ω_0.
b. Find the crucial angular velocity ω_1, such that for any higher initial angular velocity, the pendulum will keep going around its axis, and for any lower initial angular velocity, the pendulum will simply swing back and forth. Hint: When the pendulum doesn't go over the top the expression for ω will be undefined for some θs.
c. What do you think happens if the initial condition is $(0, \omega_1)$, that is, the initial angle is 0, and the initial angular velocity is exactly ω_1.

Exercise 17.E.8.3.6

Take the damped nonlinear pendulum equation $\theta'' + \mu\theta' + (\frac{g}{L})\sin\theta = 0$ for some $\mu > 0$ (that is, there is friction). Suppose the friction is large, in particular $\mu^2 > 4(\frac{g}{L})$.

a. Find and classify the critical points.
b. Explain what your findings mean, and if it agrees with what you expect in reality.

Answer

a. Critical points are $\omega = 0$, $\theta = k\pi$ for any integer k. When k is odd, we have a saddle point. When k is even we get a sink.
b. The findings mean the pendulum will simply go to one of the sinks, for example $(0, 0)$ and it will not swing back and forth. The friction is too high for it to oscillate, just like an overdamped mass-spring system.

Exercise 17.E.8.3.7

Suppose we have the system predator-prey system where the foxes are also killed at a constant rate h (h foxes killed per unit time): $x' = (a - by)x,\ y' = (cx - d)y - h$.

a. Find the critical points and the Jacobin matrices of the system.
b. Put in the constants $a = 0.4$, $b = 0.01$, $c = 0.003$, $d = 0.3$, $h = 10$. Analyze the critical points. What do you think it says about the forest?

Answer

a. Solving for the critical points we get $(0, -\frac{h}{d})$ and $\left(\frac{bh+ad}{ac}, \frac{a}{b}\right)$. The Jacobian matrix at $(0, -\frac{h}{d})$ is $\begin{bmatrix} a + \frac{bh}{d} & 0 \\ -\frac{cd}{d} & -d \end{bmatrix}$ whose eigenvalues are $a + \frac{bh}{d}$ and $-d$. The eigenvalues are real of opposite signs and we get a saddle. (In the application, however, we are only looking at the positive quadrant so this critical point is irrelevant.) At $\left(\frac{bh+ad}{ac}, \frac{a}{b}\right)$ we get Jacobian matrix $\begin{bmatrix} 0 & -\frac{b(bh+ad)}{ac} \\ \frac{ac}{b} & \frac{bh+ad}{a} - d \end{bmatrix}$.
b. For the specific numbers given, the second critical point is $\left(\frac{550}{3}, 40\right)$ the matrix is $\begin{bmatrix} 0 & -\frac{11}{6} \\ \frac{3}{25} & \frac{1}{4} \end{bmatrix}$, which has eigenvalues $\frac{5 \pm i\sqrt{327}}{40}$. Therefore there is a spiral source; the solution spirals outwards. The solution eventually hits one of the axes, $x = 0$ or $y = 0$, so something will die out in the forest.

Exercise 17.E.8.3.8: (challenging)

Suppose the foxes never die. That is, we have the system $x' = (a - by)x,\ y' = cxy$. Find the critical points and notice they are not isolated. What will happen to the population in the forest if it starts at some positive numbers. Hint: Think of the constant of motion.

Answer

The critical points are on the line $x = 0$. In the positive quadrant the y' is always positive and so the fox population always grows. The constant of motion is $C = y^a e^{-cx-by}$, for any C this curve must hit the y-axis (why?), so the trajectory will simply approach a point on the y axis somewhere and the number of hares will go to zero.

17.E.4: 8.4: Limit cycles

Exercise 17.E.8.4.1

Show that the following systems have no closed trajectories.

a. $x' = x^3 + y, y' = y^3 + x^2$,
b. $x' = e^{x-y}, y' = e^{x+y}$,
c. $x' = x + 3y^2 - y^3, y' = y^3 + x^2$.

Exercise 17.E.8.4.2

Formulate a condition for a 2-by-2 linear system $\vec{x}\,' = A\vec{x}$ to not be a center using the Bendixson-Dulac theorem. That is, the theorem says something about certain elements of A.

Exercise 17.E.8.4.3

Explain why the Bendixson-Dulac Theorem does not apply for any conservative system $x'' + h(x) = 0$.

Exercise 17.E.8.4.4

A system such as $x' = x, y' = y$ has solutions that exist for all time t, yet there are no closed trajectories or other limit cycles. Explain why the Poincare-Bendixson Theorem does not apply.

Exercise 17.E.8.4.5

Differential equations can also be given in different coordinate systems. Suppose we have the system $r' = 1 - r^2$, $\theta' = 1$ given in polar coordinates. Find all the closed trajectories and check if they are limit cycles and if so, if they are asymptotically stable or not.

Exercise 17.E.8.4.6

Show that the following systems have no closed trajectories.

a. $x' = x + y^2$, $y' = y + x^2$,
b. $x' = -x\sin^2(y), y' = e^x$,
c. $x' = xy, y' = x + x^2$.

Answer

Use Bendixson–Dulac Theorem.

a. $f_x + g_y = 1 + 1 > 0$, so no closed trajectories.
b. $f_x + g_y = -\sin^2(y) + 0 < 0$ for all x, y except the lines given by $y = k\pi$ (where we get zero), so no closed trajectories.

c. $f_x + g_y = y + 0 > 0$ for all x, y except the line given by $y = 0$ (where we get zero), so no closed trajectories.

Exercise 17.E.8.4.7

Suppose an autonomous system in the plane has a solution $x = \cos(t) + e^{-t}$, $y = \sin(t) + e^{-t}$. What can you say about the system (in particular about limit cycles and periodic solutions)?

Answer

Using Poincaré–Bendixson Theorem, the system has a limit cycle, which is the unit circle centered at the origin, as $x = \cos(t) + e^{-t}$, $y = \sin(t) + e^{-t}$ gets closer and closer to the unit circle. Thus $x = \cos(t)$, $y = \sin(t)$ is the periodic solution.

Exercise 17.E.8.4.8

Show that the limit cycle of the Van der Pol oscillator (for $\mu > 0$) must not lie completely in the set where $-\sqrt{\frac{1+\mu}{\mu}} < x < \sqrt{\frac{1+\mu}{\mu}}$. Compare with Figure 8.4.1.

Answer

$f(x, y) = y,\ g(x, y) = \mu(1 - x^2)y - x$. So $f_x + g_y = \mu(1 - x^2)$. The Bendixson–Dulac Theorem says there is no closed trajectory lying entirely in the set $x^2 < 1$.

Exercise 17.E.8.4.9

Suppose we have the system $r' = \sin(r)$, $\theta' = 1$ given in polar coordinates. Find all the closed trajectories.

Answer

The closed trajectories are those where $\sin(r) = 0$, therefore, all the circles centered at the origin with radius that is a multiple of π are closed trajectories.

17.E.5: 8.5: Chaos

Exercise 17.E.8.5.1

For the non-chaotic equation $x'' + 2px' + \omega_0^2 x = \frac{F_0}{m}\cos(\omega t)$, suppose we strobe with frequency ω as we mentioned above. Use the known steady periodic solution to find precisely the point which is the attractor for the Poincarè section.

Exercise 17.E.8.5.2: (project)

A simple fractal attractor can be drawn via the following chaos game. Draw three points of a triangle (just the vertices) and number them, say p_1, p_2 and p_3. Start with some random point p (does not have to be one of the three points above) and draw it. Roll a die, and use it to pick of the p_1, p_2, or p_3 randomly (for example 1 and 4 mean p_1, 2 and 5 mean p_2, and 3 and 6 mean p_3). Suppose we picked p_2, then let p_{new} be the point exactly halfway between p and p_2. Draw this point and let p now refer to this new point p_{new}. Rinse, repeat. Try to be precise and draw as many iterations as possible. Your points should be attracted to the so-called *Sierpinski triangle*. A computer was used to run the game for 10,000 iterations to obtain the picture in Figure 17.E.1.

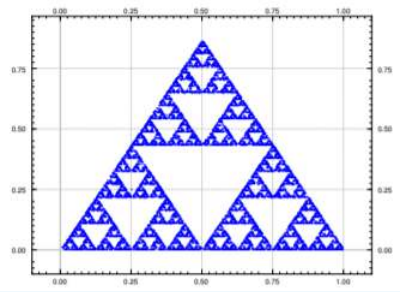


Figure $17.E.1$: $10,000$ iterations of the chaos game producing the Sierpinski triangle.

Exercise $17.E.8.5.3$: (project)

Construct the double pendulum described in the text with a string and two nuts (or heavy beads). Play around with the position of the middle nut, and perhaps use different weight nuts. Describe what you find.

Exercise $17.E.8.5.4$: (computer project)

Use a computer software (such as Matlab, Octave, or perhaps even a spreadsheet), plot the solution of the given forced Duffing equation with Euler's method. Plotting the solution for t from 0 to 100 with several different (small) step sizes. Discuss.

Exercise $17.E.8.5.5$

Find critical points of the Lorenz system and the associated linearizations.

Answer

Critical points: $(0,0,0)$, $(3\sqrt{8},3\sqrt{8},27)$, $(-3\sqrt{8},-3\sqrt{8},27)$. Linearization at $(0,0,0)$ using $u=x$, $v=y$, $w=z$ is $u'=-10u+10v$, $v'=28u-v$, $w'=-\left(\frac{8}{3}\right)w$. Linearization at $(3\sqrt{8},3\sqrt{8},27)$ using $u=x-3\sqrt{8}$, $v=y-3\sqrt{8}$, $w=z-27$ is $u'=-10u+10v$, $v'=u-v-3\sqrt{8}w$, $w'=3\sqrt{8}u+3\sqrt{8}v-\left(\frac{8}{3}\right)w$. Linearization at $(-3\sqrt{8},-3\sqrt{8},27)$ using $u=x+3\sqrt{8}$, $v=y+3\sqrt{8}$, $w=z-27$ is $u'=-10u+10v$, $v'=u-v+3\sqrt{8}w$, $w'=-3\sqrt{8}u-3\sqrt{8}v-\left(\frac{8}{3}\right)w$.

This page titled 17.E: Nonlinear Equations (Exercises) is shared under a CC BY-SA 4.0 license and was authored, remixed, and/or curated by Jiří Lebl.

- **8.E: Nonlinear Equations (Exercises)** has no license indicated.

Index

A

Airy's equation
16.2: Series Solutions of Linear Second Order ODEs
Almost Linear System
17.2: Stability and Classification of Isolated Critical Points
asymptotically stable solution
17.4: Limit cycles
attractor
17.5: Chaos
augmented matrix
1.3: Gaussian Elimination
2.3: Gaussian Elimination

B

basic variables
1.4: Uniqueness of the Reduced Row-Echelon Form
2.4: Uniqueness of the Reduced Row-Echelon Form
Bessel functions
16.3: Singular Points and the Method of Frobenius
boundary value problems
15.1: Boundary value problems
box product
4.9: The Cross Product

C

chaos
17.5: Chaos
cofactor
3.1: Basic Techniques
cofactor matrix
3.4: Applications of the Determinant
complex conjugate
6.1: Complex Numbers
complex numbers
6.1: Complex Numbers
compound interest
13.1: Growth and Decay
Conservative equations
17.2: Stability and Classification of Isolated Critical Points
constant matrix
1.3: Gaussian Elimination
2.3: Gaussian Elimination
convergence
15.3: More on the Fourier Series
Coordinate Isomorphism
9.9: The Matrix of a Linear Transformation
critical point
17.2: Stability and Classification of Isolated Critical Points
cross product
4.9: The Cross Product
cylindrical coordinates
8.2: Spherical and Cylindrical Coordinates

D

D'Alembert solution
15.8: D'Alembert Solution of The Wave Equation
De Moivre's Theorem
6.3: Roots of Complex Numbers
determinant
3.1: Basic Techniques
determinants
3.2: Properties of Determinants
direction fields
12.2: Slope fields
Dirichlet problem
15.9: Steady State Temperature and the Laplacian
distance formula
4.4: Length of a Vector
dot product
4.7: The Dot Product
double pendulum
17.5: Chaos
Duffing equation
17.5: Chaos

E

eigenvalue problems
15.1: Boundary value problems
Elementary operations
1.2: Systems of Equations, Algebraic Procedures
2.2: Systems of Equations, Algebraic Procedures
elementary row operations
1.3: Gaussian Elimination
2.3: Gaussian Elimination
elements
10.1: Sets and Set Notation
endpoint problems
15.1: Boundary value problems
equivalent matrices
1.4: Uniqueness of the Reduced Row-Echelon Form
2.4: Uniqueness of the Reduced Row-Echelon Form
escape velocity
13.3: Elementary Mechanics
Euler's formula
14.4: Constant coefficient second order linear ODEs

F

forced oscillations
14.10: Forced Oscillations and Resonance
Fourier expansion
4.11: Orthogonality
free variables
1.4: Uniqueness of the Reduced Row-Echelon Form
2.4: Uniqueness of the Reduced Row-Echelon Form
Fundamental Theorem of Algebra
6.4: The Quadratic Formula

G

gamma function
16.3: Singular Points and the Method of Frobenius
gaussian elimination
1.2: Systems of Equations, Algebraic Procedures
1.3: Gaussian Elimination
2.2: Systems of Equations, Algebraic Procedures
2.3: Gaussian Elimination
general solution
5.9: The General Solution of a Linear System
gravity
13.3: Elementary Mechanics

H

heat equation
15.9: Steady State Temperature and the Laplacian
Hermite's equation
16.2: Series Solutions of Linear Second Order ODEs

I

indicial equation
16.3: Singular Points and the Method of Frobenius
induction
10.2: Well Ordering and Induction
inner product
4.7: The Dot Product
15.2: The Trigonometric Series
interpolating polynomial
3.4: Applications of the Determinant
Isolated Critical Point
17.2: Stability and Classification of Isolated Critical Points
isomorphism
5.6: Isomorphisms
9.7: Isomorphisms

K

kernel
5.7: The Kernel and Image of A Linear Map
5.9: The General Solution of a Linear System
9.8: The Kernel and Image of a Linear Map
Kronecker delta
4.11: Orthogonality

L

Laplace equation
15.9: Steady State Temperature and the Laplacian
Laplace Expansion
3.1: Basic Techniques
Laplacian
15.9: Steady State Temperature and the Laplacian
limit cycles
17.4: Limit cycles
linear independence
4.11: Orthogonality
9.3: Linear Independence
linear map
5.6: Isomorphisms
linear transformation
5.6: Isomorphisms
5.9: The General Solution of a Linear System

M

mechanics
13.3: Elementary Mechanics
method of Frobenius
16.3: Singular Points and the Method of Frobenius
method of undetermined coefficients
14.2: The Method of Undetermined Coefficients I
14.3: The Method of Undetermined Coefficients II
minor of a matrix
3.1: Basic Techniques
minors or a matrix
3.1: Basic Techniques
modulus (complex number)
6.2: Polar Form

N

natural frequency
14.10: Forced Oscillations and Resonance
Newton's Law of Cooling
13.2: Cooling and Mixing
normal vector
4.8: Planes in R^n

null space
5.9: The General Solution of a Linear System

O

onto transformation
9.7: Isomorphisms
open rectangle
12.5: Existence and Uniqueness of Solutions of Nonlinear Equations
ordinary point
16: Power series methods
16.2: Series Solutions of Linear Second Order ODEs
orthogonal
4.7: The Dot Product
orthogonal matrix
4.11: Orthogonality
orthogonality
4.11: Orthogonality
Overdamping
14.8: Mechanical Vibrations

P

parallelepiped
4.9: The Cross Product
Parametric Lines
4.6: Parametric Lines
particular solution
5.9: The General Solution of a Linear System
polar coordinates
8.1: Polar Coordinates and Polar Graphs
Polar Graphs
8.1: Polar Coordinates and Polar Graphs
Position Vector
4.1: Vectors in R^n
projection
4.7: The Dot Product

Q

quadratic formula
6.4: The Quadratic Formula

R

ratio test
16.1: Power Series
reduction of orders
14.6: Reduction of Order
regular singular point
16.3: Singular Points and the Method of Frobenius
relaxation oscillation
17.4: Limit cycles
resonance
14.10: Forced Oscillations and Resonance
right hand system
4.9: The Cross Product
row operations
3.2: Properties of Determinants

S

scalar equation of the plane
4.8: Planes in R^n
scalar product
4.7: The Dot Product
set
10.1: Sets and Set Notation
set notation
10.1: Sets and Set Notation
singular point
16.2: Series Solutions of Linear Second Order ODEs
skew lines
1.1: Systems of Equations, Geometry
2.1: Systems of Equations, Geometry
slope fields
12.2: Slope fields
span
9.2: Spanning Sets
spanning sets
9.2: Spanning Sets
spherical coordinates
8.2: Spherical and Cylindrical Coordinates
strange attractors
17.5: Chaos
subsets
9.2: Spanning Sets
subspace
4.11: Orthogonality
superposition
14.5: Higher order linear ODEs
system of linear equations
1.2: Systems of Equations, Algebraic Procedures
2.2: Systems of Equations, Algebraic Procedures

T

temperature decay constant
13.2: Cooling and Mixing
terminal velocity
13.3: Elementary Mechanics
Triangle Inequality
4.7: The Dot Product
6.1: Complex Numbers
triangular matrix
3.1: Basic Techniques

U

union
10.1: Sets and Set Notation

V

van der Pol oscillator
17.4: Limit cycles
variation of parameters
14.7: Variation of Parameters
vector addition
4.3: Geometric Meaning of Vector Addition
vector equation of the plane
4.8: Planes in R^n
vector space
9.1: Algebraic Considerations

W

Well Ordering Principle
10.2: Well Ordering and Induction

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Math C285: Linear Algebra and Diffrential Equations (Tran)

Webpages: 159

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY 4.0: 49.1% (78 pages)
- CC BY-SA 4.0: 31.4% (50 pages)
- CC BY-NC-SA 3.0: 10.7% (17 pages)
- Undeclared: 8.8% (14 pages)

By Page

- Math C285: Linear Algebra and Diffrential Equations (Tran) - *Undeclared*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Systems of Equations - *CC BY 4.0*
    - 1.1: Systems of Equations, Geometry - *CC BY 4.0*
    - 1.2: Systems of Equations, Algebraic Procedures - *CC BY 4.0*
    - 1.3: Gaussian Elimination - *CC BY 4.0*
    - 1.4: Uniqueness of the Reduced Row-Echelon Form - *CC BY 4.0*
    - 1.5: Rank and Homogeneous Systems - *CC BY 4.0*
    - 1.6: Balancing Chemical Reactions - *CC BY 4.0*
    - 1.7: Dimensionless Variables - *CC BY 4.0*
    - 1.8: An Application to Resistor Networks - *CC BY 4.0*
    - 1.9: Exercises - *CC BY 4.0*
  - 2: Systems of Equations - *CC BY 4.0*
    - 2.1: Systems of Equations, Geometry - *CC BY 4.0*
    - 2.2: Systems of Equations, Algebraic Procedures - *CC BY 4.0*
    - 2.3: Gaussian Elimination - *CC BY 4.0*
    - 2.4: Uniqueness of the Reduced Row-Echelon Form - *CC BY 4.0*
    - 2.5: Rank and Homogeneous Systems - *CC BY 4.0*
    - 2.6: Balancing Chemical Reactions - *CC BY 4.0*
    - 2.7: Dimensionless Variables - *CC BY 4.0*
    - 2.8: An Application to Resistor Networks - *CC BY 4.0*
    - 2.9: Exercises - *CC BY 4.0*
  - 3: Determinants - *CC BY 4.0*
    - 3.1: Basic Techniques - *CC BY 4.0*
    - 3.2: Properties of Determinants - *CC BY 4.0*
    - 3.3: Finding Determinants using Row Operations - *CC BY 4.0*
    - 3.4: Applications of the Determinant - *CC BY 4.0*
    - 3.E: Exercises - *CC BY 4.0*
  - 4: R^n - *CC BY 4.0*
    - 4.1: Vectors in R^n - *CC BY 4.0*
    - 4.2: Vector Algebra - *CC BY 4.0*
    - 4.3: Geometric Meaning of Vector Addition - *CC BY 4.0*
    - 4.4: Length of a Vector - *CC BY 4.0*
    - 4.5: Geometric Meaning of Scalar Multiplication - *CC BY 4.0*
    - 4.6: Parametric Lines - *CC BY 4.0*
    - 4.7: The Dot Product - *CC BY 4.0*
    - 4.8: Planes in R^n - *CC BY 4.0*
    - 4.9: The Cross Product - *CC BY 4.0*
    - 4.10: Spanning, Linear Independence and Basis in R^n - *CC BY 4.0*
    - 4.11: Orthogonality - *CC BY 4.0*
    - 4.12: Applications - *CC BY 4.0*
    - 4.E: Exercises - *CC BY 4.0*
  - 5: Linear Transformations - *CC BY 4.0*
    - 5.1: Linear Transformations - *CC BY 4.0*
    - 5.2: The Matrix of a Linear Transformation I - *CC BY 4.0*
    - 5.3: Properties of Linear Transformations - *CC BY 4.0*
    - 5.4: Special Linear Transformations in R^2 - *CC BY 4.0*
    - 5.5: One-to-One and Onto Transformations - *CC BY 4.0*
    - 5.6: Isomorphisms - *CC BY 4.0*
    - 5.7: The Kernel and Image of A Linear Map - *CC BY 4.0*
    - 5.8: The Matrix of a Linear Transformation II - *CC BY 4.0*

- 5.9: The General Solution of a Linear System - *CC BY 4.0*
- 5.E: Exercises - *Undeclared*

- 6: Complex Numbers - *CC BY 4.0*
  - 6.1: Complex Numbers - *CC BY 4.0*
  - 6.2: Polar Form - *CC BY 4.0*
  - 6.3: Roots of Complex Numbers - *CC BY 4.0*
  - 6.4: The Quadratic Formula - *CC BY 4.0*
  - 6.E: Exercises - *CC BY 4.0*
- 7: Spectral Theory - *CC BY 4.0*
  - 7.1: Eigenvalues and Eigenvectors of a Matrix - *CC BY 4.0*
  - 7.2: Diagonalization - *CC BY 4.0*
  - 7.3: Applications of Spectral Theory - *CC BY 4.0*
  - 7.4: Orthogonality - *CC BY 4.0*
  - 7.E: Exercises - *Undeclared*
- 8: Some Curvilinear Coordinate Systems - *CC BY 4.0*
  - 8.1: Polar Coordinates and Polar Graphs - *CC BY 4.0*
  - 8.2: Spherical and Cylindrical Coordinates - *CC BY 4.0*
  - 8.E: Exercises - *CC BY 4.0*
- 9: Vector Spaces - *CC BY 4.0*
  - 9.1: Algebraic Considerations - *CC BY 4.0*
  - 9.2: Spanning Sets - *CC BY 4.0*
  - 9.3: Linear Independence - *CC BY 4.0*
  - 9.4: Subspaces and Basis - *CC BY 4.0*
  - 9.5: Sums and Intersections - *CC BY 4.0*
  - 9.6: Linear Transformations - *CC BY 4.0*
  - 9.7: Isomorphisms - *CC BY 4.0*
  - 9.8: The Kernel and Image of a Linear Map - *CC BY 4.0*
  - 9.9: The Matrix of a Linear Transformation - *CC BY 4.0*
  - 9.E: Exercises - *Undeclared*
- 10: Some Prerequisite Topics - *CC BY 4.0*
  - 10.1: Sets and Set Notation - *CC BY 4.0*
  - 10.2: Well Ordering and Induction - *CC BY 4.0*
- 11: Introduction - *CC BY-SA 4.0*
  - 11.1: Notes About These Notes - *CC BY-SA 4.0*
  - 11.2: Introduction to Differential Equations - *CC BY-SA 4.0*
  - 11.3: Classification of Differential Equations - *CC BY-SA 4.0*
  - 11.4: First order ODEs - *CC BY-SA 4.0*
  - 11.E: Introduction (Exercises) - *CC BY-SA 4.0*
- 12: First order ODEs - *CC BY-SA 4.0*
  - 12.1: Integrals as solutions - *CC BY-SA 4.0*
  - 12.2: Slope fields - *CC BY-SA 4.0*
  - 12.3: Separable Equations - *CC BY-SA 4.0*
  - 12.4: Linear equations and the integrating factor - *CC BY-SA 4.0*
  - 12.5: Existence and Uniqueness of Solutions of Nonlinear Equations - *CC BY-NC-SA 3.0*
    - 12.5E: Existence and Uniqueness of Solutions of Nonlinear Equations (Exercises) - *CC BY-NC-SA 3.0*
  - 12.6: Substitution - *CC BY-SA 4.0*
  - 12.7: Autonomous equations - *CC BY-SA 4.0*
  - 12.8: Numerical methods- Euler's method - *CC BY-SA 4.0*
  - 12.9: Exact Equations - *CC BY-SA 4.0*
  - 12.10: First Order Linear PDE - *CC BY-SA 4.0*
  - 12.E: First order ODEs (Exercises) - *CC BY-SA 4.0*
- 13: Applications of First Order Equations - *CC BY-NC-SA 3.0*
  - 13.1: Growth and Decay - *CC BY-NC-SA 3.0*
    - 13.1E: Growth and Decay (Exercises) - *CC BY-NC-SA 3.0*
  - 13.2: Cooling and Mixing - *CC BY-NC-SA 3.0*
    - 13.2E: Cooling and Mixing (Exercises) - *CC BY-NC-SA 3.0*
  - 13.3: Elementary Mechanics - *CC BY-NC-SA 3.0*
    - 13.3E: Elementary Mechanics (Exercises) - *CC BY-NC-SA 3.0*
- 14: Higher order linear ODEs - *CC BY-SA 4.0*
  - 14.1: Second order linear ODEs - *CC BY-SA 4.0*
  - 14.2: The Method of Undetermined Coefficients I - *CC BY-NC-SA 3.0*
    - 14.2E: The Method of Undetermined Coefficients I (Exercises) - *CC BY-NC-SA 3.0*
  - 14.3: The Method of Undetermined Coefficients II - *CC BY-NC-SA 3.0*
    - 14.3E: The Method of Undetermined Coefficients II (Exercises) - *CC BY-NC-SA 3.0*
  - 14.4: Constant coefficient second order linear ODEs - *CC BY-SA 4.0*
  - 14.5: Higher order linear ODEs - *CC BY-SA 4.0*
  - 14.6: Reduction of Order - *CC BY-NC-SA 3.0*
    - 14.6E: Reduction of Order (Exercises) - *CC BY-NC-SA 3.0*
  - 14.7: Variation of Parameters - *CC BY-NC-SA 3.0*
    - 14.7E: Variation of Parameters (Exercises) - *CC BY-NC-SA 3.0*
  - 14.8: Mechanical Vibrations - *CC BY-SA 4.0*
  - 14.9: Nonhomogeneous Equations - *CC BY-SA 4.0*
  - 14.10: Forced Oscillations and Resonance - *CC BY-SA 4.0*

- 14.E: Higher order linear ODEs (Exercises) - *CC BY-SA 4.0*

- 15: Fourier series and PDEs - *CC BY-SA 4.0*
  - 15.1: Boundary value problems - *CC BY-SA 4.0*
  - 15.2: The Trigonometric Series - *CC BY-SA 4.0*
  - 15.3: More on the Fourier Series - *CC BY-SA 4.0*
  - 15.4: Sine and Cosine Series - *CC BY-SA 4.0*
  - 15.5: Applications of Fourier Series - *CC BY-SA 4.0*
  - 15.6: PDEs, Separation of Variables, and The Heat Equation - *CC BY-SA 4.0*
  - 15.7: One Dimensional Wave Equation - *CC BY-SA 4.0*
  - 15.8: D'Alembert Solution of The Wave Equation - *CC BY-SA 4.0*
  - 15.9: Steady State Temperature and the Laplacian - *CC BY-SA 4.0*
  - 15.10: Dirichlet Problem in the Circle and the Poisson Kernel - *CC BY-SA 4.0*
  - 15.E: Fourier Series and PDEs (Exercises) - *CC BY-SA 4.0*
- 16: Power series methods - *CC BY-SA 4.0*
  - 16.1: Power Series - *CC BY-SA 4.0*
  - 16.2: Series Solutions of Linear Second Order ODEs - *CC BY-SA 4.0*
  - 16.3: Singular Points and the Method of Frobenius - *CC BY-SA 4.0*
  - 16.E: Power series methods (Exercises) - *CC BY-SA 4.0*
- 17: Nonlinear Systems - *CC BY-SA 4.0*
  - 17.1: Linearization, Critical Points, and Equilibria - *CC BY-SA 4.0*
  - 17.2: Stability and Classification of Isolated Critical Points - *CC BY-SA 4.0*
  - 17.3: Applications of Nonlinear Systems - *CC BY-SA 4.0*
  - 17.4: Limit cycles - *CC BY-SA 4.0*
  - 17.5: Chaos - *CC BY-SA 4.0*
  - 17.6: Appendix B- Table of Laplace Transforms - *CC BY-SA 4.0*
  - 17.E: Nonlinear Equations (Exercises) - *CC BY-SA 4.0*
- Back Matter - *Undeclared*
  - Index - *Undeclared*
  - Glossary - *Undeclared*
  - Detailed Licensing - *Undeclared*
  - Detailed Licensing - *Undeclared*

Detailed Licensing

Overview

Title: Math C285: Linear Algebra and Diffrential Equations (Tran)

Webpages: 159

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY 4.0: 49.1% (78 pages)
- CC BY-SA 4.0: 31.4% (50 pages)
- CC BY-NC-SA 3.0: 10.7% (17 pages)
- Undeclared: 8.8% (14 pages)

By Page

- Math C285: Linear Algebra and Diffrential Equations (Tran) - *Undeclared*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Systems of Equations - *CC BY 4.0*
    - 1.1: Systems of Equations, Geometry - *CC BY 4.0*
    - 1.2: Systems of Equations, Algebraic Procedures - *CC BY 4.0*
    - 1.3: Gaussian Elimination - *CC BY 4.0*
    - 1.4: Uniqueness of the Reduced Row-Echelon Form - *CC BY 4.0*
    - 1.5: Rank and Homogeneous Systems - *CC BY 4.0*
    - 1.6: Balancing Chemical Reactions - *CC BY 4.0*
    - 1.7: Dimensionless Variables - *CC BY 4.0*
    - 1.8: An Application to Resistor Networks - *CC BY 4.0*
    - 1.9: Exercises - *CC BY 4.0*
  - 2: Systems of Equations - *CC BY 4.0*
    - 2.1: Systems of Equations, Geometry - *CC BY 4.0*
    - 2.2: Systems of Equations, Algebraic Procedures - *CC BY 4.0*
    - 2.3: Gaussian Elimination - *CC BY 4.0*
    - 2.4: Uniqueness of the Reduced Row-Echelon Form - *CC BY 4.0*
    - 2.5: Rank and Homogeneous Systems - *CC BY 4.0*
    - 2.6: Balancing Chemical Reactions - *CC BY 4.0*
    - 2.7: Dimensionless Variables - *CC BY 4.0*
    - 2.8: An Application to Resistor Networks - *CC BY 4.0*
    - 2.9: Exercises - *CC BY 4.0*
  - 3: Determinants - *CC BY 4.0*
    - 3.1: Basic Techniques - *CC BY 4.0*
    - 3.2: Properties of Determinants - *CC BY 4.0*
    - 3.3: Finding Determinants using Row Operations - *CC BY 4.0*
    - 3.4: Applications of the Determinant - *CC BY 4.0*
    - 3.E: Exercises - *CC BY 4.0*
  - 4: R^n - *CC BY 4.0*
    - 4.1: Vectors in R^n - *CC BY 4.0*
    - 4.2: Vector Algebra - *CC BY 4.0*
    - 4.3: Geometric Meaning of Vector Addition - *CC BY 4.0*
    - 4.4: Length of a Vector - *CC BY 4.0*
    - 4.5: Geometric Meaning of Scalar Multiplication - *CC BY 4.0*
    - 4.6: Parametric Lines - *CC BY 4.0*
    - 4.7: The Dot Product - *CC BY 4.0*
    - 4.8: Planes in R^n - *CC BY 4.0*
    - 4.9: The Cross Product - *CC BY 4.0*
    - 4.10: Spanning, Linear Independence and Basis in R^n - *CC BY 4.0*
    - 4.11: Orthogonality - *CC BY 4.0*
    - 4.12: Applications - *CC BY 4.0*
    - 4.E: Exercises - *CC BY 4.0*
  - 5: Linear Transformations - *CC BY 4.0*
    - 5.1: Linear Transformations - *CC BY 4.0*
    - 5.2: The Matrix of a Linear Transformation I - *CC BY 4.0*
    - 5.3: Properties of Linear Transformations - *CC BY 4.0*
    - 5.4: Special Linear Transformations in R^2 - *CC BY 4.0*
    - 5.5: One-to-One and Onto Transformations - *CC BY 4.0*
    - 5.6: Isomorphisms - *CC BY 4.0*
    - 5.7: The Kernel and Image of A Linear Map - *CC BY 4.0*
    - 5.8: The Matrix of a Linear Transformation II - *CC BY 4.0*

- 5.9: The General Solution of a Linear System - *CC BY 4.0*
- 5.E: Exercises - *Undeclared*

- 6: Complex Numbers - *CC BY 4.0*
  - 6.1: Complex Numbers - *CC BY 4.0*
  - 6.2: Polar Form - *CC BY 4.0*
  - 6.3: Roots of Complex Numbers - *CC BY 4.0*
  - 6.4: The Quadratic Formula - *CC BY 4.0*
  - 6.E: Exercises - *CC BY 4.0*
- 7: Spectral Theory - *CC BY 4.0*
  - 7.1: Eigenvalues and Eigenvectors of a Matrix - *CC BY 4.0*
  - 7.2: Diagonalization - *CC BY 4.0*
  - 7.3: Applications of Spectral Theory - *CC BY 4.0*
  - 7.4: Orthogonality - *CC BY 4.0*
  - 7.E: Exercises - *Undeclared*
- 8: Some Curvilinear Coordinate Systems - *CC BY 4.0*
  - 8.1: Polar Coordinates and Polar Graphs - *CC BY 4.0*
  - 8.2: Spherical and Cylindrical Coordinates - *CC BY 4.0*
  - 8.E: Exercises - *CC BY 4.0*
- 9: Vector Spaces - *CC BY 4.0*
  - 9.1: Algebraic Considerations - *CC BY 4.0*
  - 9.2: Spanning Sets - *CC BY 4.0*
  - 9.3: Linear Independence - *CC BY 4.0*
  - 9.4: Subspaces and Basis - *CC BY 4.0*
  - 9.5: Sums and Intersections - *CC BY 4.0*
  - 9.6: Linear Transformations - *CC BY 4.0*
  - 9.7: Isomorphisms - *CC BY 4.0*
  - 9.8: The Kernel and Image of a Linear Map - *CC BY 4.0*
  - 9.9: The Matrix of a Linear Transformation - *CC BY 4.0*
  - 9.E: Exercises - *Undeclared*
- 10: Some Prerequisite Topics - *CC BY 4.0*
  - 10.1: Sets and Set Notation - *CC BY 4.0*
  - 10.2: Well Ordering and Induction - *CC BY 4.0*
- 11: Introduction - *CC BY-SA 4.0*
  - 11.1: Notes About These Notes - *CC BY-SA 4.0*
  - 11.2: Introduction to Differential Equations - *CC BY-SA 4.0*
  - 11.3: Classification of Differential Equations - *CC BY-SA 4.0*
  - 11.4: First order ODEs - *CC BY-SA 4.0*
  - 11.E: Introduction (Exercises) - *CC BY-SA 4.0*
- 12: First order ODEs - *CC BY-SA 4.0*
  - 12.1: Integrals as solutions - *CC BY-SA 4.0*
  - 12.2: Slope fields - *CC BY-SA 4.0*
  - 12.3: Separable Equations - *CC BY-SA 4.0*
  - 12.4: Linear equations and the integrating factor - *CC BY-SA 4.0*
  - 12.5: Existence and Uniqueness of Solutions of Nonlinear Equations - *CC BY-NC-SA 3.0*
    - 12.5E: Existence and Uniqueness of Solutions of Nonlinear Equations (Exercises) - *CC BY-NC-SA 3.0*
  - 12.6: Substitution - *CC BY-SA 4.0*
  - 12.7: Autonomous equations - *CC BY-SA 4.0*
  - 12.8: Numerical methods- Euler's method - *CC BY-SA 4.0*
  - 12.9: Exact Equations - *CC BY-SA 4.0*
  - 12.10: First Order Linear PDE - *CC BY-SA 4.0*
  - 12.E: First order ODEs (Exercises) - *CC BY-SA 4.0*
- 13: Applications of First Order Equations - *CC BY-NC-SA 3.0*
  - 13.1: Growth and Decay - *CC BY-NC-SA 3.0*
    - 13.1E: Growth and Decay (Exercises) - *CC BY-NC-SA 3.0*
  - 13.2: Cooling and Mixing - *CC BY-NC-SA 3.0*
    - 13.2E: Cooling and Mixing (Exercises) - *CC BY-NC-SA 3.0*
  - 13.3: Elementary Mechanics - *CC BY-NC-SA 3.0*
    - 13.3E: Elementary Mechanics (Exercises) - *CC BY-NC-SA 3.0*
- 14: Higher order linear ODEs - *CC BY-SA 4.0*
  - 14.1: Second order linear ODEs - *CC BY-SA 4.0*
  - 14.2: The Method of Undetermined Coefficients I - *CC BY-NC-SA 3.0*
    - 14.2E: The Method of Undetermined Coefficients I (Exercises) - *CC BY-NC-SA 3.0*
  - 14.3: The Method of Undetermined Coefficients II - *CC BY-NC-SA 3.0*
    - 14.3E: The Method of Undetermined Coefficients II (Exercises) - *CC BY-NC-SA 3.0*
  - 14.4: Constant coefficient second order linear ODEs - *CC BY-SA 4.0*
  - 14.5: Higher order linear ODEs - *CC BY-SA 4.0*
  - 14.6: Reduction of Order - *CC BY-NC-SA 3.0*
    - 14.6E: Reduction of Order (Exercises) - *CC BY-NC-SA 3.0*
  - 14.7: Variation of Parameters - *CC BY-NC-SA 3.0*
    - 14.7E: Variation of Parameters (Exercises) - *CC BY-NC-SA 3.0*
  - 14.8: Mechanical Vibrations - *CC BY-SA 4.0*
  - 14.9: Nonhomogeneous Equations - *CC BY-SA 4.0*
  - 14.10: Forced Oscillations and Resonance - *CC BY-SA 4.0*

- 14.E: Higher order linear ODEs (Exercises) - *CC BY-SA 4.0*

- 15: Fourier series and PDEs - *CC BY-SA 4.0*
  - 15.1: Boundary value problems - *CC BY-SA 4.0*
  - 15.2: The Trigonometric Series - *CC BY-SA 4.0*
  - 15.3: More on the Fourier Series - *CC BY-SA 4.0*
  - 15.4: Sine and Cosine Series - *CC BY-SA 4.0*
  - 15.5: Applications of Fourier Series - *CC BY-SA 4.0*
  - 15.6: PDEs, Separation of Variables, and The Heat Equation - *CC BY-SA 4.0*
  - 15.7: One Dimensional Wave Equation - *CC BY-SA 4.0*
  - 15.8: D'Alembert Solution of The Wave Equation - *CC BY-SA 4.0*
  - 15.9: Steady State Temperature and the Laplacian - *CC BY-SA 4.0*
  - 15.10: Dirichlet Problem in the Circle and the Poisson Kernel - *CC BY-SA 4.0*
  - 15.E: Fourier Series and PDEs (Exercises) - *CC BY-SA 4.0*
- 16: Power series methods - *CC BY-SA 4.0*
  - 16.1: Power Series - *CC BY-SA 4.0*
  - 16.2: Series Solutions of Linear Second Order ODEs - *CC BY-SA 4.0*
  - 16.3: Singular Points and the Method of Frobenius - *CC BY-SA 4.0*
  - 16.E: Power series methods (Exercises) - *CC BY-SA 4.0*
- 17: Nonlinear Systems - *CC BY-SA 4.0*
  - 17.1: Linearization, Critical Points, and Equilibria - *CC BY-SA 4.0*
  - 17.2: Stability and Classification of Isolated Critical Points - *CC BY-SA 4.0*
  - 17.3: Applications of Nonlinear Systems - *CC BY-SA 4.0*
  - 17.4: Limit cycles - *CC BY-SA 4.0*
  - 17.5: Chaos - *CC BY-SA 4.0*
  - 17.6: Appendix B- Table of Laplace Transforms - *CC BY-SA 4.0*
  - 17.E: Nonlinear Equations (Exercises) - *CC BY-SA 4.0*
- Back Matter - *Undeclared*
  - Index - *Undeclared*
  - Glossary - *Undeclared*
  - Detailed Licensing - *Undeclared*
  - Detailed Licensing - *Undeclared*