MATH 26: INTRODUCTION TO LINEAR ALGEBRA



Kenn Huber
Irvine Valley College

LibreTexts™

Math 26: Introduction to Linear Algebra

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 an 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 05/06/2026

TABLE OF CONTENTS

Licensing

1: Systems of Linear Equations

- 1.1: Systems of Linear Equations
  - 1.1.1: Lines, Planes, and Space
  - 1.1.2: Parametric Descriptions
  - 1.1.3: Practice Set
- 1.2: Row Reduction
  - 1.2.1: Elimination and Augmented Matrices
  - 1.2.2: Echelon Forms
  - 1.2.3: Free Variables
  - 1.2.4: Practice Set
- 1.3: Vectors
  - 1.3.1: Vector Algebra and Geometry
  - 1.3.2: Vector Equations and Spans
  - 1.3.3: Practice Set
- 1.4: Matrix Equations
  - 1.4.1: Matrix Equations, Spans, and Consistency
  - 1.4.2: Solutions to Matrix Equations
  - 1.4.3: Practice Set
- 1.5: Linear Independence
  - 1.5.1: Defining Linear Independence
  - 1.5.2: Understanding Linear Independence
  - 1.5.3: Practice Set
- 1.6: Subspaces
  - 1.6.1: Defining Subspaces
  - 1.6.2: Common Subspaces
  - 1.6.3: Practice Set
- 1.7: Basis and Dimension
  - 1.7.1: Defining Basis
  - 1.7.2: Finding a Basis of a Subspace
  - 1.7.3: Rank and Nullity
  - 1.7.4: Practice Set

2: Linear Transformations and Matrix Algebra

- 2.1: Matrix Transformations
  - 2.1.1: Matrices as Functions
  - 2.1.2: Matrix Transformations
  - 2.1.3: Practice Set
- 2.2: Linear Transformations
  - 2.2.1: Linear Transformations
  - 2.2.2: Matrix Representations

- 2.2.3: Practice Set
- 2.3: Matrix Algebra
  - 2.3.1: Sums of Transformations and Matrices
  - 2.3.2: Transformation Composition and Matrix Multiplication
  - 2.3.3: Matrix Inverses
  - 2.3.4: Practice Set
- 2.4: Invertibility
  - 2.4.1: Invertibility of Linear Transformations
  - 2.4.2: Invertibility of Matrices
  - 2.4.3: Practice Set
- 2.5: Determinants- Definition
  - 2.5.1: Definition of Determinant
  - 2.5.2: Properties of the Determinant
  - 2.5.3: Practice Set
- 2.6: Cofactor Expansions
  - 2.6.1: Cofactor Expansion
  - 2.6.2: Cramer's Rule and Adjugates
  - 2.6.3: Practice Set
- 2.7: Determinants and Volumes
  - 2.7.1: Determinants for Area and Volume
  - 2.7.2: Practice Set

3: Eigenvalues and Eigenvectors

- 3.1: Eigenvalues and Eigenvectors
  - 3.1.1: Eigenvalues and Eigenvectors Definitions
  - 3.1.2: Eigenspaces
  - 3.1.3: Characteristic Polynomial
  - 3.1.4: Practice Set
- 3.2: Diagonalization and Geometry of Eigenvalues
  - 3.2.1: Coordinatization and Similar Matrices
  - 3.2.2: Diagonalization
  - 3.2.3: Practice Set
- 3.3: Geometry of Eigenvalues
  - 3.3.1: Geometry of Diagonalizable Matrices
  - 3.3.2: Complex Eigenvalues and Eigenvectors
  - 3.3.3: Geometry of Matrices with Complex Eigenvalues
  - 3.3.4: Practice Set
- 3.4: Dot Products and Orthogonal Complements
  - 3.4.1: Dot Products and Orthogonality
  - 3.4.2: Orthogonal Complements
  - 3.4.3: Practice Set
- 3.5: Orthogonal Projection and Least Squares
  - 3.5.1: Orthogonal Decomposition
  - 3.5.2: Least Squares Solutions
  - 3.5.3: Practice Set
- 3.6: Orthogonal Sets

- 3.6.1: Orthogonal Sets and Projection
- 3.6.2: Gram-Schmidt Process
- 3.6.3: Practice Set

4: Vector Spaces

- 4.1: General Vector Spaces and Subspaces
  - 4.1.1: General Vector Spaces
  - 4.1.2: Calculations in Vector Spaces
  - 4.1.3: Practice Set
- 4.2: Linear Transformations in General Vector Spaces
  - 4.2.1: Linear Transformations
  - 4.2.2: Coordinatization
  - 4.2.3: Practice Set
- 4.3: Examples and Applications
  - 4.3.1: Differential Equations
  - 4.3.2: Famous Transformations
  - 4.3.3: Inner Product Spaces
  - 4.3.4: Markov Chains
  - 4.3.5: Page Rank
  - 4.3.6: Practice Set

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 Linear Equations

1.1: Systems of Linear Equations
- 1.1: Lines, Planes, and Space
- 1.1: Parametric Descriptions
- 1.1: Practice Set

1.2: Row Reduction
- 1.2: Elimination and Augmented Matrices
- 1.2: Echelon Forms
- 1.2: Free Variables
- 1.2: Practice Set

1.3: Vectors
- 1.3: Vector Algebra and Geometry
- 1.3: Vector Equations and Spans
- 1.3: Practice Set

1.4: Matrix Equations
- 1.4: Matrix Equations, Spans, and Consistency
- 1.4: Solutions to Matrix Equations
- 1.4: Practice Set

1.5: Linear Independence
- 1.5: Defining Linear Independence
- 1.5: Understanding Linear Independence
- 1.5: Practice Set

1.6: Subspaces
- 1.6: Defining Subspaces
- 1.6: Common Subspaces
- 1.6: Practice Set

1.7: Basis and Dimension
- 1.7: Defining Basis
- 1.7: Finding a Basis of a Subspace
- 1.7: Rank and Nullity
- 1.7: Practice Set

This page titled 1: Systems of Linear Equations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

1.1: Systems of Linear Equations

Objectives

1. Understand the definition of $\mathbb{R}^n$, and what it means to use $\mathbb{R}^n$ to label points on a geometric object.
2. *Pictures:* solutions of systems of linear equations, parameterized solution sets.
3. *Vocabulary words:* ***consistent***, ***inconsistent***, ***solution set***.

During the first chapter of this textbook, we will be primarily concerned with understanding the solutions of systems of linear equations. We define these terms and attempt to visualize what they and their solutions look like.

This page titled 1.1: Systems of Linear Equations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

1.1: Lines, Planes, and Space

Definition: Linear

An equation in the unknowns $x, y, z, \ldots$ is called ***linear*** if both sides of the equation are a sum of (constant) multiples of $x, y, z, \ldots$, plus an optional constant.

For instance,

$$\begin{aligned} 3x + 4y &= 2z \\ -x - z &= 100 \end{aligned}$$

are linear equations, but

$$\begin{aligned} 3x + yz &= 3 \\ \sin(x) - \cos(y) &= 2 \end{aligned}$$

are not.

We will usually move the unknowns to the left side of the equation and move the constants to the right.

A ***system*** of linear equations is a collection of several linear equations, like

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases}$$

Definition: Solution Sets

- A ***solution*** of a system of equations is a list of numbers $x, y, z, \ldots$ that make all of the equations true simultaneously.
- The ***solution set*** of a system of equations is the collection of all solutions.
- ***Solving*** the system means finding all solutions with formulas involving some number of parameters.

A solution of a system of equations in n variables is a list of n numbers. For example, $(x, y, z) = (1, -2, 3)$ is a solution of the above system involving three variables. As we will be studying solutions of systems of equations throughout this text, now is a good time to fix our notions regarding lists of numbers.

A system of linear equations need not have a solution. For example, there do not exist numbers x and y making the following two equations true simultaneously:

$$\begin{cases} x + 2y = 3 \\ x + 2y = -3 \end{cases}$$

In this case, the solution set is *empty*. As this is a rather important property of a system of equations, it has its own name.

Definition: Inconsistent and Consistent

A system of equations is called ***inconsistent*** if it has no solutions. It is called ***consistent*** otherwise.

We use $\mathbb{R}$ to denote the set of all real numbers, i.e., the number line. This contains rational and irrational numbers like $0, \frac{3}{2}, -\pi, 104, \ldots$

Definition: $\mathbb{R}$

Let n be a positive whole number. We define

$$\mathbb{R}^n = \text{all ordered } n\text{-tuples of real numbers } (x_1, x_2, x_3, \ldots, x_n)$$

An n-tuple of real numbers is called a ***point*** of $\mathbb{R}^n$.

In other words, $\mathbb{R}^n$ is just the set of all (ordered) lists of n real numbers. We will draw pictures of $\mathbb{R}^n$ in a moment, but keep in mind that *this is the definition*. For example, $\left(0, \frac{3}{2}, -\pi\right)$ and $(1, -2, 3)$ are points of $\mathbb{R}^3$.

Illustration 1.1.1 - The number line

When $n = 1$, we just get $\mathbb{R}$ back: $\mathbb{R}^1 = \mathbb{R}$. Geometrically, this is the number line.

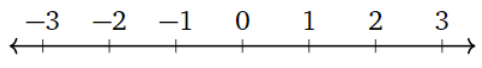


Figure 1.1.1

Illustration 1.1.2 - The Euclidean plane

When $n = 2$, we can think of $\mathbb{R}^2$ as the xy-plane. We can do so because every point on the plane can be represented by an ordered pair of real numbers, namely, its x- and y-coordinates.

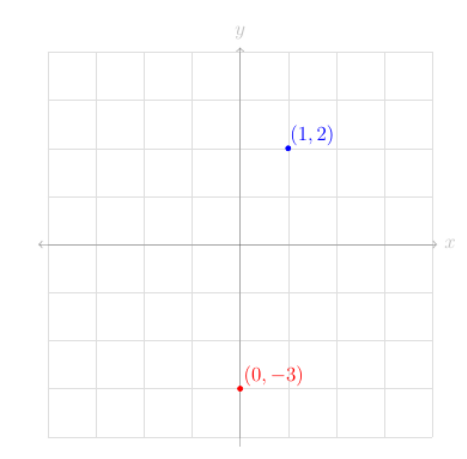


Figure 1.1.2

Illustration 1.1.3 - 3-Space

When $n = 3$, we can think of $\mathbb{R}^3$ as the *space* we (appear to) live in. We can do so because every point in space can be represented by an ordered triple of real numbers, namely, its x-, y-, and z-coordinates.

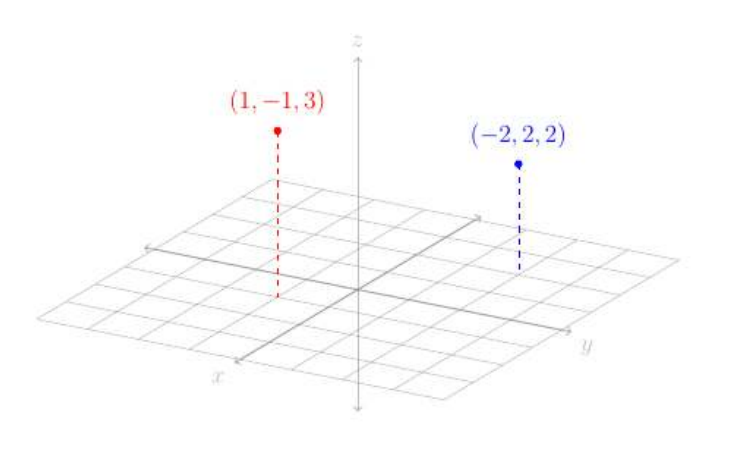


Figure 1.1.3

So what is $\mathbb{R}^4$? or $\mathbb{R}^5$? or $\mathbb{R}^n$? These are harder to visualize, so you have to go back to the definition: $\mathbb{R}^n$ is the set of all ordered n-tuples of real numbers $(x_1, x_2, x_3, \ldots, x_n)$.

They are still "geometric" spaces, in the sense that our intuition for $\mathbb{R}^2$ and $\mathbb{R}^3$ often extends to $\mathbb{R}^n$.

We will make definitions and state theorems that apply to any $\mathbb{R}^n$, but we will only draw pictures for $\mathbb{R}^2$ and $\mathbb{R}^3$.

The power of using these spaces is the ability to *label* various objects of interest, such as geometric objects and solutions of systems of equations, by the points of $\mathbb{R}^n$.

Illustration 1.1.4: Color Space

All colors you can see can be described by three quantities: the amount of red, green, and blue light in that color. (Humans are trichromatic.) Therefore, we can use the points of $\mathbb{R}^3$ to *label* all colors. For instance, the point $(.2, .4, .9)$ labels the color with 20% red, 40% green, and 90% blue intensity.

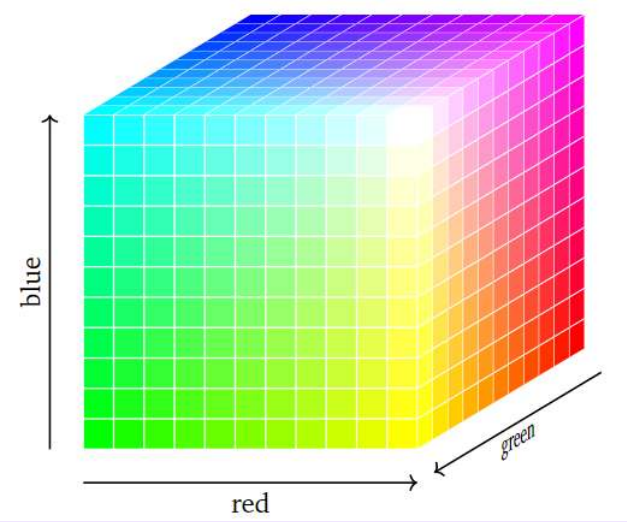


Figure 1.1.4

Illustration 1.1.5: Traffic Flow

In the Overview, we could have used $\mathbb{R}^4$ to *label* the amount of traffic (x, y, z, w) passing through four streets. In other words, if there are $10, 5, 3, 11$ cars per hour passing through roads x, y, z, w, respectively, then this can be recorded by the point $(10, 5, 3, 11)$ in $\mathbb{R}^4$. This is useful from a psychological standpoint: instead of having four numbers, we are now dealing with just *one* piece of data.

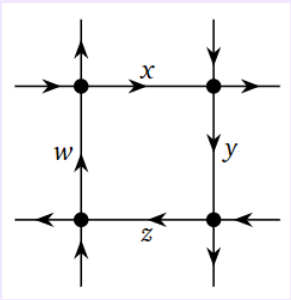


Figure 1.1.5

Illustration 1.1.6: QR Codes

A QR code is a method of storing data in a grid of black and white squares in a way that computers can easily read. A typical QR code is a 29×29 grid. Reading each line left-to-right and reading the lines top-to-bottom (like you read a book), we can think of such a QR code as a sequence of $29 \times 29 = 841$ digits, each digit being 1 (for white) or 0 (for black). In such a way, the entire QR code can be regarded as a point in $\mathbb{R}^{841}$. As in the previous examples, it is very useful from a psychological perspective to view a QR code as a *single* piece of data in this way.



Figure 1.1.6. The QR code for this textbook is a 29×29 array of black/white squares.

In the above examples, it was useful from a psychological perspective to replace a list of four numbers (representing traffic flow) or of 841 numbers (representing a QR code) by a single piece of data: a point in some $\mathbb{R}^n$. This is a powerful concept, and we will almost exclusively record solutions of systems of linear equations in this way.

1.1: Lines, Planes, and Space is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.1: Parametric Descriptions

Before discussing how to write the solution set of a system of linear equations below, it is helpful to see some pictures of what these solution sets look like geometrically.

Illustration 1.1.7: One Equation in Two Variables

Consider the linear equation $2x + y = 4$. We can rewrite this as $y = 4 - 2x$, which defines a line in the plane: the slope is -2, and the y-intercept is $(0, 4)$.



Figure 1.1.7

In this context, we call $2x + y = 4$ an implicit equation of the line.

Since $y = 4 - 2x$, if we let $x = t$, we can write the same line in parametric form as follows:

$$(x, y) = (t, 4 - 2t) \quad \text{for any} \quad t \in \mathbb{R}$$

This means that every point on the line has the form $(t, 4 - 2t)$ for some real number t. In this case, we call t a parameter, as it parameterizes the points on the line.

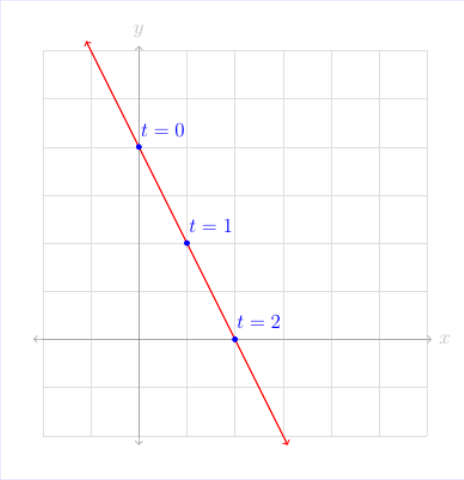


Figure 1.1.8

Notice that we could have solved for x instead, and gotten $x = 2 - \frac{1}{2}y$.

So if we wanted to let $y = s$, we would instead get a parametric form that looks like:

$$(x, y) = \left(2 - \frac{1}{2}s, s\right) \quad \text{for any} \quad s \in \mathbb{R}$$

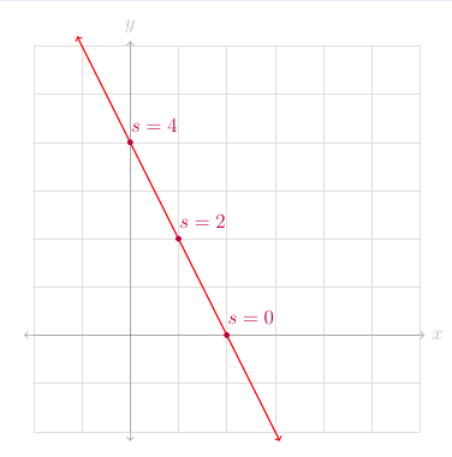


Figure 1.1.9

These are clearly the same line, and all are acceptable ways to describe it.

Definition: Lines

For our purposes, a ***line*** is a ray that is *straight* and *infinite* in both directions.

Illustration 1.1.8 - One Equation in Three Variables

Consider the linear equation $x + y + z = 1$. This is the ***implicit equation*** for a plane in space.

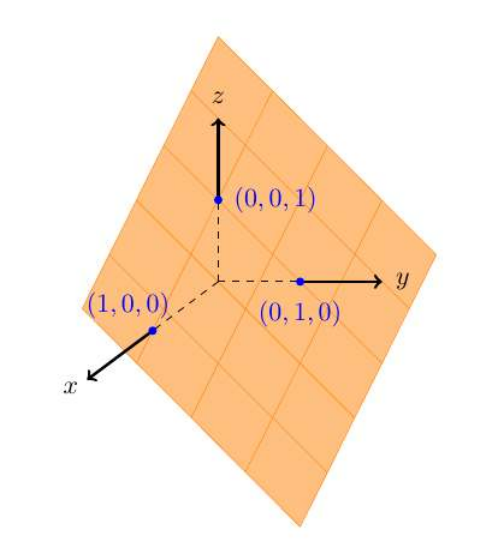


Figure 1.1.10

If we solve for x in this equation, we get $x = 1 - y - z$. Since there are no more restrictions, we can essentially pick any value for the variables y and z, and there will always be one x-value that makes the equation happy. So, let $y = t$ and $z = w$ and we see one ***parametric form***:

$$(x, y, z) = (1 - t - w, t, w) \quad \text{for any} \quad t, w \in \mathbb{R}$$

In this case, we need two ***parameters***, t and w, to describe all points on the plane.

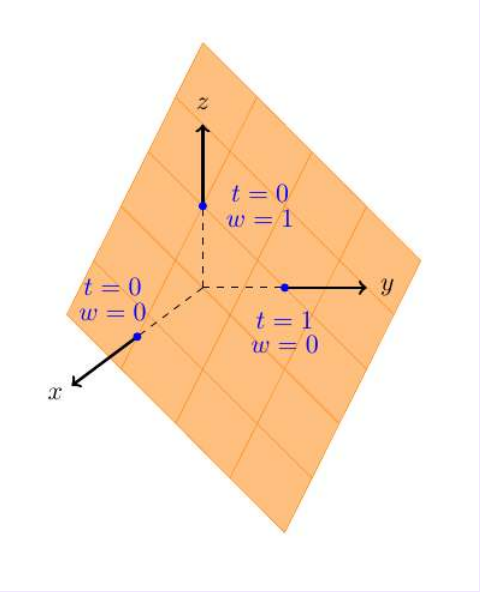


Figure 1.1.11

Note that the parameters t, w allow us to use $\mathbb{R}^2$ to label the points on the plane. However, this plane is *not the same* as the plane $\mathbb{R}^2$: indeed, every point on this plane has three coordinates, like the point $(0, 0, 1)$.

Definition: Planes

A *plane* is a flat sheet that is infinite in all directions.

Remark

The equation $x + y + z + w = 1$ defines a "3-plane" in 4-space, and more generally, a single linear equation in n variables defines an "$(n-1)$-plane" in n-space. We will make these statements precise in future chapters.

Solutions to systems of equations can come in many shapes.

Illustration 1.1.9

Consider the system of two linear equations

$$\begin{cases} x - 3y = -3 \\ 2x + y = 8 \end{cases}$$

Each equation individually defines a line in the plane, pictured below.

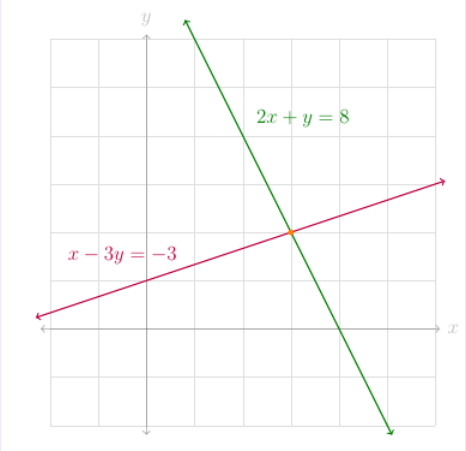


Figure 1.1.12

A solution to the *system* of both equations is a pair of numbers (x, y) that makes both equations true at once. In other words, it is a point that lies on both lines simultaneously. We can see in the picture above that there is only one point where the lines intersect: therefore, this system has exactly one solution. (This solution is $(3, 2)$, as the reader can verify.)

Usually, two lines in the plane will intersect in one point, but of course this is not always the case.

Illustration 1.1.10

Next, consider the system of two linear equations

$$\begin{cases} x - 3y = -3 \\ x - 3y = 3 \end{cases}$$

These define *parallel* lines in the plane.

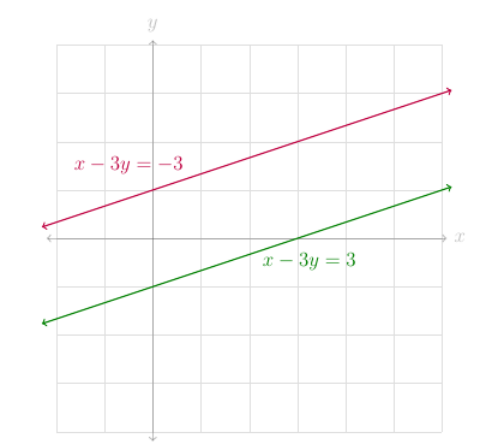


Figure 1.1.13

The fact that the lines do not intersect means that the system of equations has no solution. Of course, this is easy to see algebraically: if $x - 3y = -3$, then it cannot also be the case that $x - 3y = 3$.

Illustration 1.1.11

There is one more possibility. Consider the system of equations

$$\begin{cases} x - 3y = -3 \\ 2x - 6y = -6 \end{cases}$$

The second equation is a multiple of the first, so these equations define the *same* line in the plane.

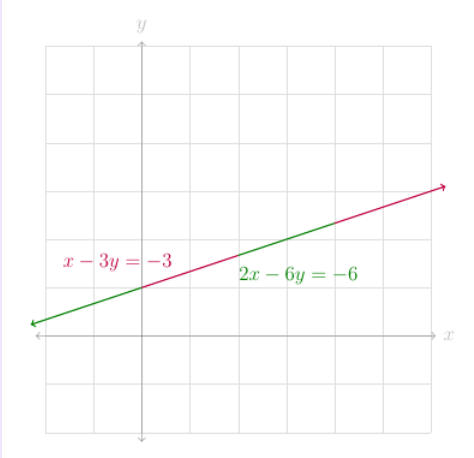


Figure 1.1.14

In this case, there are infinitely many solutions of the system of equations.

Illustration 1.1.12 - Two Equations in Three Variables

Consider the system of two linear equations

$$\begin{cases} x + y + z = 1 \\ x \quad\quad - z = 0 \end{cases}$$

Each equation individually defines a plane in space. The solutions of the system of both equations are the points that lie on both planes. We can see in the picture below that the planes intersect in a line. In particular, this system has infinitely many solutions.

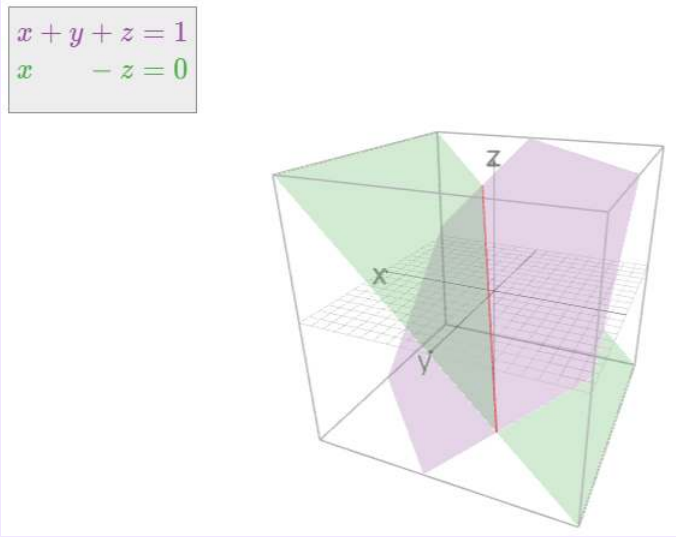


Figure 1.1.15

In this case, since we know the solution must be a line, we can find a ***parametric form*** with one ***parameter***, t.

One such way to do so is to set $x = t$, and the green plane ($x - z = 0$) forces z to satisfy $z = t$.

Plugging these each into the purple plane ($x + y + z = 1$), we see that $y = 1 - 2t$ must hold. Therefore, our line can be written:

$$(x, y, z) = (t, 1 - 2t, t) \quad \text{for any} \quad t \in \mathbb{R}$$

Note that in any case, the parameter t allows us to use $\mathbb{R}$ to *label* the points on the line. However, neither line is the same as the number line $\mathbb{R}$: indeed, every point on the line in Example 1.1.7 line has two coordinates, like the point $(0,4)$, and every point on the line here has three coordinates, like $(0,1,0)$.

Remark

In general, the solutions of a system of equations in n variables is the intersection of "$(n-1)$-planes" in n-space. This is always some kind of linear space, as we will discuss in the future.

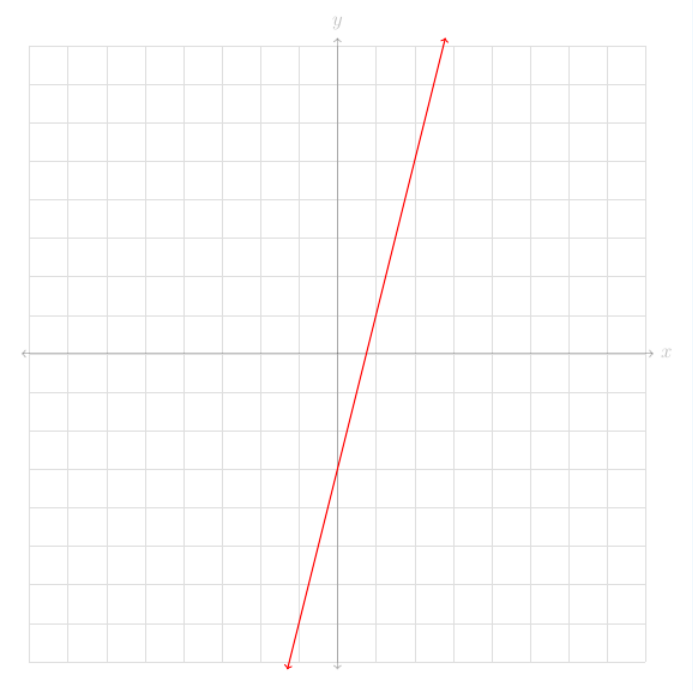
1.1: Parametric Descriptions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.1: Practice Set

Exercise 1.1.1*a*

Graph $y = 4x - 3$ in the xy-plane.

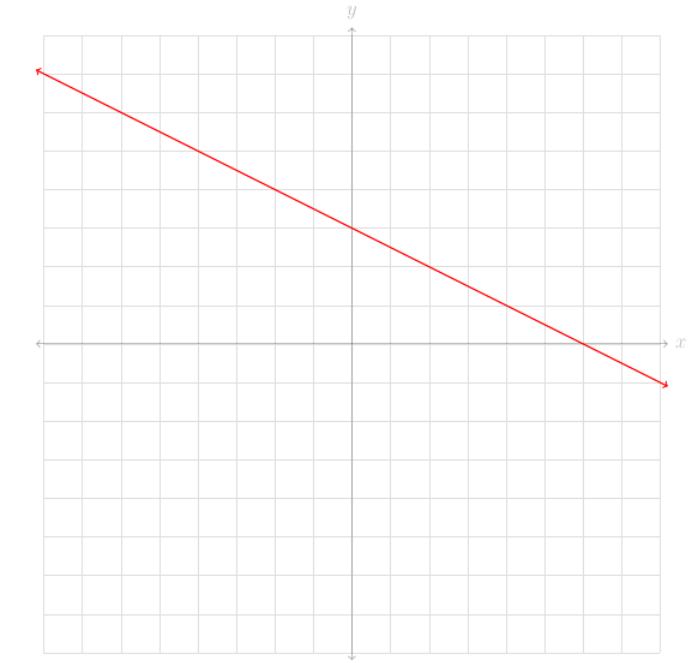
Answer



Exercise 1.1.1*b*

Graph $x + 2y = 6$ in the xy-plane.

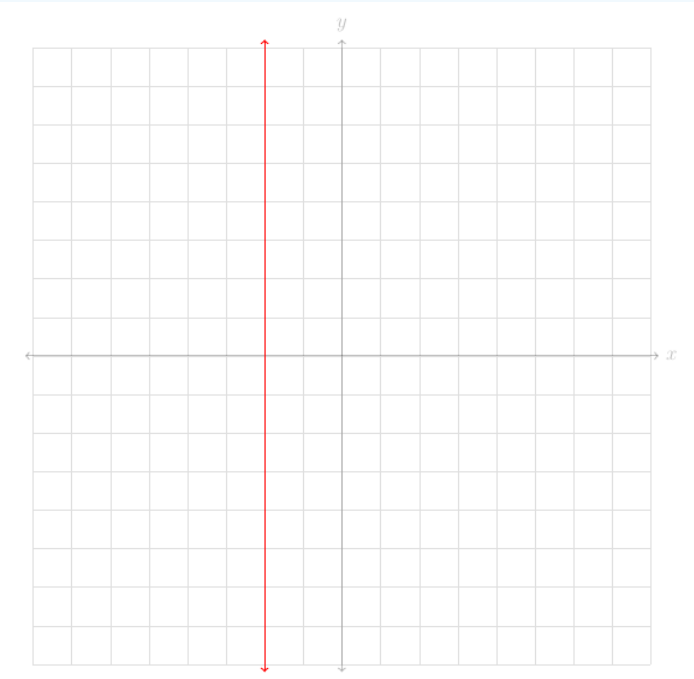
Answer



Exercise 1.1.1c

Graph $x = -2$ in the xy-plane.

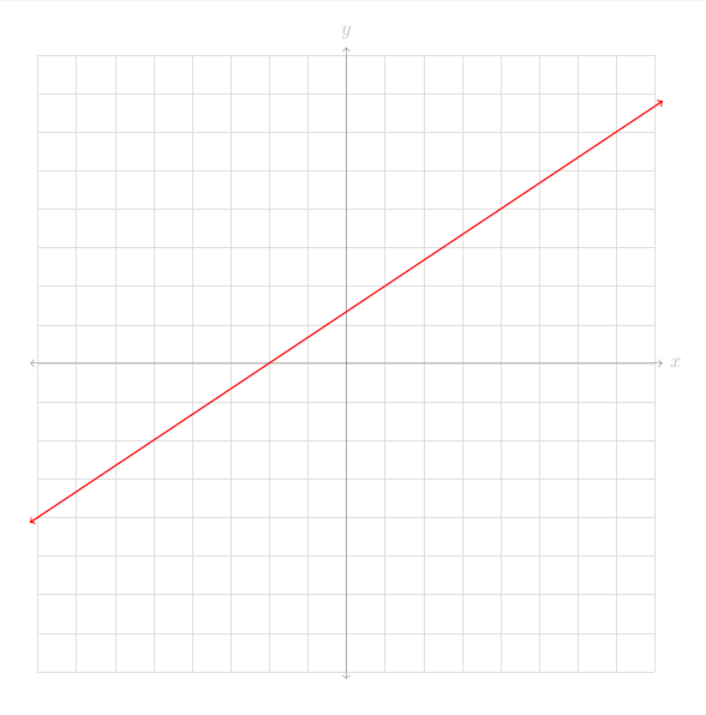
Answer



Exercise 1.1.1d

Graph $y - 2 = \frac{2}{3}(x - 1)$ in the xy-plane.

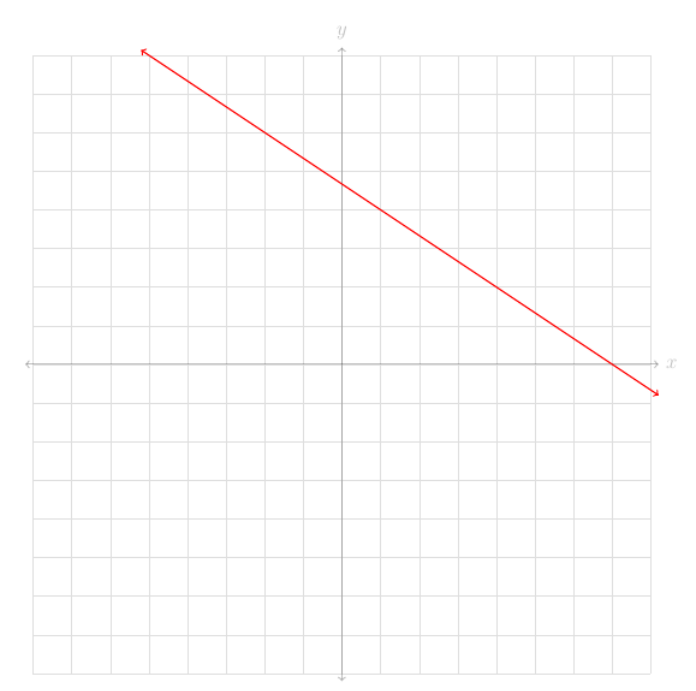
Answer



Exercise 1.1.1e

Graph $x(t) = 3t + 1, y(t) = -2t + 4$ in the xy-plane.

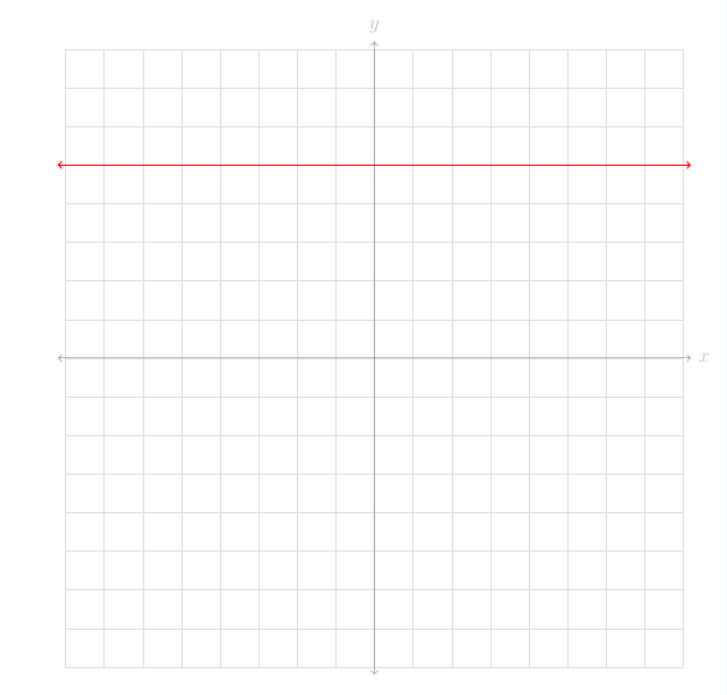
Answer



Exercise 1.1.1f

Graph $x(t) = 9t + 8, y(t) = 5$ in the xy-plane.

Answer

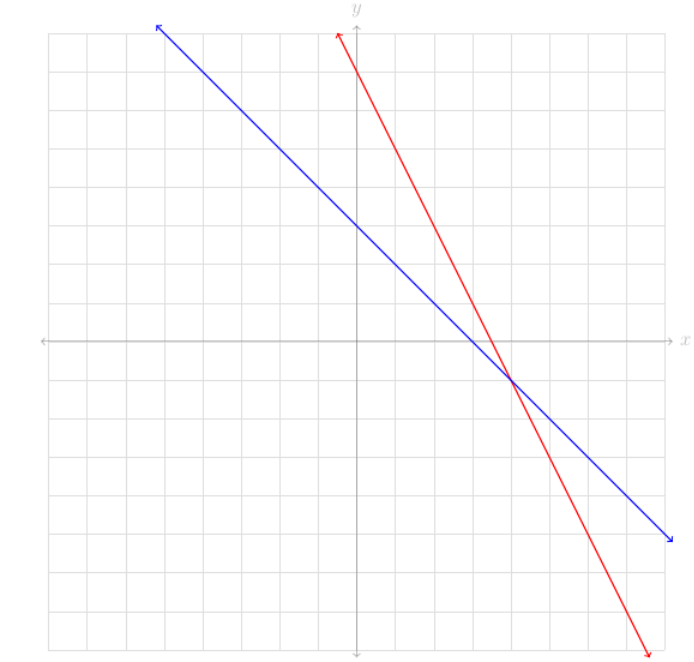


Exercise 1.1.2a

Describe the shape of the solution set to $\begin{cases} 2x + y = 7 \\ x + y = 3 \end{cases}$

Answer

These two lines are not parallel, so the solution set is a single point: $(4, -1)$

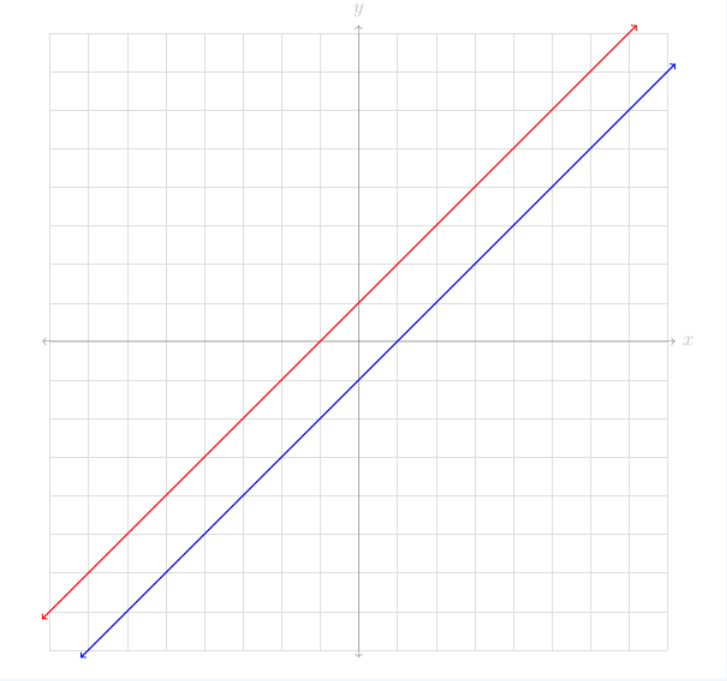


Exercise 1.1.2b

Describe the shape of the solution set to $\begin{cases} -x+y=1 \\ 4x-4y=4 \end{cases}$

Answer

These two lines are parallel, so the solution set is empty.

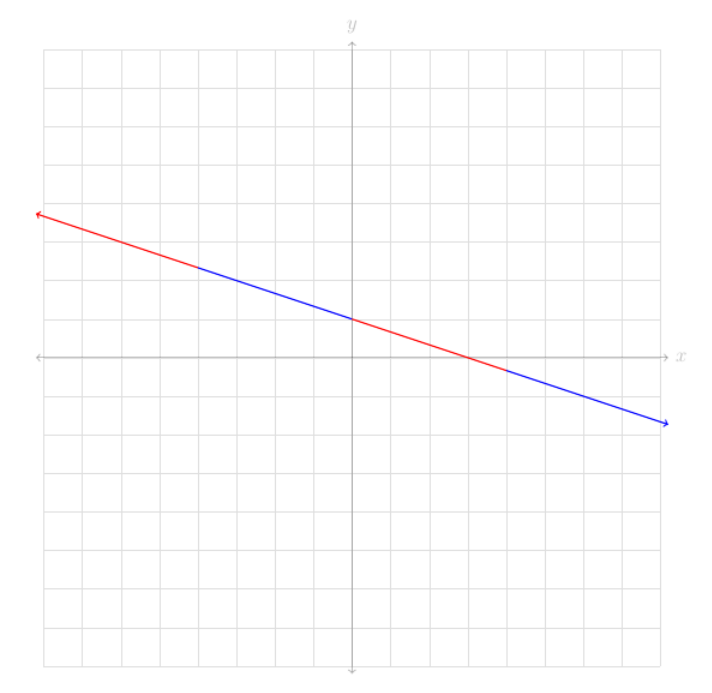


Exercise 1.1.2c

Describe the shape of the solution set to $\begin{cases} x+3y=3 \\ 2x+6y=6 \end{cases}$

Answer

These two lines are the same, so the solution set is the line $y=-\frac{1}{3}x+1$



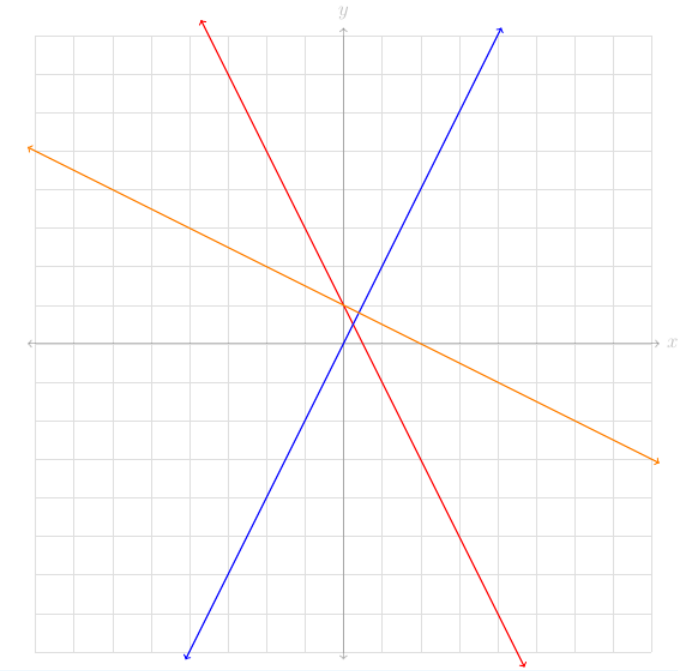
Exercise 1.1.2d

Describe the shape of the solution set to $\begin{cases} 2x+y=1 \\ 2x-y=0 \\ x+2y=2 \end{cases}$

Answer

These three lines do not intersect at a common point, so the solution set is empty.

https://math.libretexts.org/@go/page/186280



Exercise 1.1.2e

Describe the shape of the solution set to $\begin{cases} x+y+z=0 \\ x-y-z=0 \end{cases}$

Answer

These are two non-parallel planes, so their intersection must be a line.

Substituting the bottom equation $x=y+z$ into the top equation gives $(y+z)+(y+z)=0$

This simplifies to the linear relationship $y=-z$. Substituting this back gives $x=-z+z=0$.

Therefore our solution set is the line containing points of the form $(0,-z,z)$

Exercise 1.1.2f

Describe the shape of the solution set to $\begin{cases} 4x+3y=2 \\ x-z=1 \\ 3y+4z=-2 \end{cases}$

Answer

With three planes, we rely on algebra to determine our solution set.

The middle equation gives $x=1+z$ which we substitute into the top equation to get $4(1+z)+3y=2$

This simplifies into $3y+4z=-2$, which is exactly the third plane equation. Thus, we have $y=-\frac{4}{3}z-\frac{2}{3}$

Our solution set is the line containing points of the form $(1+z,-\frac{4}{3}z-\frac{2}{3},z)$

Exercise 1.1.2g

Describe the shape of the solution set to $\begin{cases} 5x-15y+10z=30 \\ \frac{1}{6}x-\frac{1}{2}y+\frac{1}{3}z=1 \end{cases}$

Answer

These two are both multiples of the same plane $x-3y+2z=6$, so our solution set is that entire plane.

Exercise 1.1.2h

Describe the shape of the solution set to $\begin{cases} x+2y-2z=4 \\ x-y+3z=4 \\ 3x+4z=0 \end{cases}$

Answer

We use the bottom equation to write $x=-\frac{4}{3}z$ and substitute into the other two:

$\left(-\frac{4}{3}z\right)+2y-2z=4$ simplifies to $y=\frac{5}{3}z+2$

$\left(-\frac{4}{3}z\right)-y+3z=4$ simplifies to $y=\frac{5}{3}z-4$

These are two parallel lines, so our solution set is empty.

Exercise 1.1.3

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) The intersection of two lines is a point.

(b) The intersection of three lines is a point.

(c) The intersection of two planes is a point.

(d) The intersection of three planes is a point.

(e) If two lines share a common point, then their intersection is exactly one point.

(f) If two lines share a common point, then their intersection is a line.

(g) If two planes share a common point, then their intersection is exactly one point.

(h) If two planes share a common point, then their intersection is a line.

(i) If two planes share a common point, then their intersection is a plane.

(j) If three planes share a common point, then their intersection is exactly one point.

(k) If three planes share a common point, then their intersection is a line.

(l) If three planes share a common point, then their intersection is a plane.

Answer

(a) **Sometimes True Sometimes False**

When the lines are not parallel they intersect at a point, but when they are parallel they do not intersect at any point, and when they are the same line they intersect at infinitely many points.

(b) **Sometimes True Sometimes False**

The three lines $y=x, y=2x, y=0$ all intersect at the origin, but the three lines $y=x, y=-x, y=1$ do not intersect.

(c) **Always False**

Two non-parallel planes always intersect to be a line or a plane.

(d) **Sometimes True Sometimes False**

The planes $x=0, y=0, z=0$ all intersect at the origin, but in Exercise 2h we saw three planes with no intersection.

(e) **Sometimes True Sometimes False**

If they are not the same line, they will share exactly one point. If they are the same line they will share infinitely many points.

(f) **Sometimes True Sometimes False**

As we saw in Exercise 2c, two copies of the same line intersect at infinitely many points. However, if they are not parallel lines or the same line, then the intersection is a point.

(g) **Always False**

Two non-parallel planes always intersect to be a line or a plane.

(h) **Sometimes True Sometimes False**

We saw in Exercise 2e that this can be true, and in Exercise 2g that this can be false.

(i) **Sometimes True Sometimes False**

We saw in Exercise 2g that this can be true, and in Exercise 2e that this can be false.

(j) **Sometimes True Sometimes False**

This is true for examples like $x = 0, y = 0, z = 0$ which are all planes that intersect at the origin, but Exercise 2f shows that this can also be false.

(k) **Sometimes True Sometimes False**

We saw in Exercise 2f that this can be true. However, if they are all three copies of the same plane, they can intersect at more than just a line.

(l) **Sometimes True Sometimes False**

This is obviously true if they are all the same plane, but Exercise 2f shows that this can also be false.

1.1: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.2: Row Reduction

Objectives

1. Learn to replace a system of linear equations by an augmented matrix.
2. Learn how the elimination method corresponds to performing row operations on an augmented matrix.
3. Understand when a matrix is in (reduced) row echelon form.
4. Learn which row reduced matrices come from inconsistent linear systems.
5. *Recipe:* the row-reduction algorithm.
6. *Vocabulary words:* ***row operation***, ***row equivalence***, ***matrix***, ***augmented matrix***, ***pivot***, ***(reduced) row echelon form, free variable***

In this section, we will present an algorithm for "solving" a system of linear equations.

Instead of relying on graphing (which can be hard in the xy-plane, and very hard in the xyz-space, and impossible for higher dimensions) or substitution, we begin with elimination and introduce matrices to help keep things organized.

This page titled 1.2: Row Reduction is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

1.2.1: Elimination and Augmented Matrices

We will solve systems of linear equations algebraically using the ***elimination*** method. In other words, we will combine the equations in various ways to try to eliminate as many variables as possible from each equation. There are three valid operations we can perform on our system of equations:

- ***Scaling****:* we can multiply both sides of an equation by a nonzero number.

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases} \xrightarrow{\text{multiply 1st by } -3} \begin{cases} -3x - 6y - 9z = -18 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases}$$

- ***Replacement****:* we can add a multiple of one equation to another, replacing the second equation with the result.

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases} \xrightarrow{\text{2nd} = \text{2nd} - 2\times\text{1st}} \begin{cases} x + 2y + 3z = 6 \\ -7y - 4z = 2 \\ 3x + y - z = -2 \end{cases}$$

- ***Swap****:* we can swap two equations.

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases} \xrightarrow{\text{3rd} \leftrightarrow \text{1st}} \begin{cases} 3x + y - z = -2 \\ 2x - 3y + 2z = 14 \\ x + 2y + 3z = 6 \end{cases}$$

Example 1.2.1.1

Use elimination to solve

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases}$$

Solution

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases} \xrightarrow{\text{2nd} = \text{2nd} - 2\times\text{1st}} \begin{cases} x + 2y + 3z = 6 \\ -7y - 4z = 2 \\ 3x + y - z = -2 \end{cases}$$

$$\xrightarrow{\text{3rd} = \text{3rd} - 3\times\text{1st}} \begin{cases} x + 2y + 3z = 6 \\ -7y - 4z = 2 \\ -5y - 10z = -20 \end{cases}$$

$$\xrightarrow{\text{2nd} \leftrightarrow \text{3rd}} \begin{cases} x + 2y + 3z = 6 \\ -5y - 10z = -20 \\ -7y - 4z = 2 \end{cases}$$

$$\xrightarrow{\text{divide 2nd by } -5} \begin{cases} x + 2y + 3z = 6 \\ y + 2z = 4 \\ -7y - 4z = 2 \end{cases}$$

$$\xrightarrow{\text{3rd} = \text{3rd} + 7\times\text{2nd}} \begin{cases} x + 2y + 3z = 6 \\ y + 2z = 4 \\ 10z = 30 \end{cases}$$

At this point we've eliminated both x and y from the third equation, and we can solve $10 \cdot z = 30$ to get $z = 3$. Substituting for z in the second equation gives $y + 2 \cdot 3 = 4$, or $y = -2$. Substituting for y and z in the first equation gives $x + 2 \cdot (-2) + 3 \cdot 3 = 6$, or $x = 1$. Thus, the only solution is $\boxed{(x, y, z) = (1, -2, 3)}$

We can check that our solution is correct by substituting $(x, y, z) = (1, -2, 3)$ into the original equation:

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases} \xrightarrow{\text{substitute}} \begin{cases} 1 + 2\cdot(-2) + 3\cdot 3 = 6 \\ 2\cdot 1 - 3\cdot(-2) + 2\cdot 3 = 14 \\ 3\cdot 1 + (-2) - 3 = -2 \end{cases}$$

Solving equations by elimination requires writing the variables x, y, z and the equals sign $=$ over and over again, merely as placeholders: all that is changing in the equations is the coefficient *numbers*. We can make our life easier by extracting only the numbers and putting them in a box:

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases} \xrightarrow{\text{becomes}} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right]$$

This is called an ***augmented matrix***. The word "augmented" refers to the vertical line, which we draw to remind ourselves where the equals sign belongs; a ***matrix*** is a grid of numbers without the vertical line. In this notation, our three valid ways of manipulating our equations become ***row operations***:

- ***Scaling****:* multiply all entries in a row by a nonzero number.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right] \xrightarrow{R_1 = R_1 \times -3} \left[\begin{array}{ccc|c} -3 & -6 & -9 & -18 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right]$$

  Here the notation R_1 simply means "the first row", and likewise for R_2, R_3, etc.
- ***Replacement****:* add a multiple of one row to another, replacing the second row with the result.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right] \xrightarrow{R_2 = R_2 - 2\times R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 3 & 1 & -1 & -2 \end{array}\right]$$

- ***Swap****:* interchange two rows.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{ccc|c} 3 & 1 & -1 & -2 \\ 2 & -3 & 2 & 14 \\ 1 & 2 & 3 & 6 \end{array}\right]$$

Remark

When we wrote our row operations above, we used expressions like $R_2 = R_2 - 2 \times R_1$. Of course this does not mean that the second row is equal to the second row minus twice the first row. Instead it means that we are *replacing* the second row with the second row minus twice the first row. This kind of syntax is used frequently in computer programming when we want to change the value of a variable.

Example 1.2.1.2

Use row operations to solve

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases}$$

Solution

We start by forming an augmented matrix:

$$\begin{cases} x + 2y + 3z = 6 \\ 2x - 3y + 2z = 14 \\ 3x + y - z = -2 \end{cases} \xrightarrow{\text{becomes}} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right]$$

Eliminating a variable from an equation means producing a zero to the left of the line in an augmented matrix. First, we produce zeros in the first column (i.e. we eliminate x) by subtracting multiples of the first row.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right] \xrightarrow{R_2=R_2-2R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 3 & 1 & -1 & -2 \end{array}\right]$$
$$\xrightarrow{R_3=R_3-3R_1} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & -5 & -10 & -20 \end{array}\right]$$

This was made much easier by the fact that the top-left entry is equal to 1, so we can simply multiply the first row by the number below and subtract. In order to eliminate y in the same way, we would like to produce a 1 in the second column. We could divide the second row by -7, but this would produce fractions; instead, let's divide the third by -5.

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & -5 & -10 & -20 \end{array}\right] \xrightarrow{R_3=R_3\cdot\left(-\frac{1}{5}\right)} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & 1 & 2 & 4 \end{array}\right]$$
$$\xrightarrow{R_2\leftrightarrow R_3} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & -7 & -4 & 2 \end{array}\right]$$
$$\xrightarrow{R_3=R_3+7R_2} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 10 & 30 \end{array}\right]$$
$$\xrightarrow{R_3=R_3\div 10} \left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

We swapped the second and third rows just to keep things orderly. Now we translate this augmented matrix back into a system of equations:

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right] \xrightarrow{\text{becomes}} \begin{cases} x + 2y + 3z = 6 \\ y + 2z = 4 \\ z = 3 \end{cases}$$

Hence, $z = 3$; back-substituting gives $\boxed{(x, y, z) = (1, -2, 3)}$

The process of doing row operations to a matrix does not change the solution set of the corresponding linear equations!

Indeed, the whole point of doing these operations is to solve the equations using the elimination method.

Definition: Row Equivalent

Two matrices are called ***row equivalent*** if one can be obtained from the other by doing some number of row operations.

So the linear equations of row-equivalent matrices have the *same solution set*.

Example 1.2.1.3: An Inconsistent System

Solve the following system of equations using row operations:

$$\begin{cases} x + y = 2 \\ 3x + 4y = 5 \\ 4x + 5y = 9 \end{cases}$$

Solution

First, we put our system of equations into an augmented matrix.

$$\begin{cases} x + y = 2 \\ 3x + 4y = 5 \\ 4x + 5y = 9 \end{cases} \xrightarrow{\text{augmented matrix}} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 3 & 4 & 5 \\ 4 & 5 & 9 \end{array}\right]$$

We clear the entries below the top-left using row replacement.

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 3 & 4 & 5 \\ 4 & 5 & 9 \end{array}\right] \xrightarrow{R_2=R_2-3R_1} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 4 & 5 & 9 \end{array}\right] \xrightarrow{R_3=R_3-4R_1} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{array}\right]$$

Now we clear the second entry from the last row.

$$\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{array}\right] \xrightarrow{R_3=R_3-R_2} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{array}\right]$$

This translates back into the system of equations

$$\begin{cases} x + y = 2 \\ y = -1 \\ 0 = 2 \end{cases}$$

Our original system has the same solution set as this system. But this system has no solutions: there are no values of x, y making the third equation true! We conclude that our original equation was **inconsistent**

1.2.1: Elimination and Augmented Matrices is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.2.2: Echelon Forms

We saw how to translate a system of linear equations into an augmented matrix; now we want to find an *algorithm* for "solving" such an augmented matrix. First. we must decide what it means for an augmented matrix to be "solved".

Definition: Row-Echelon Form

A matrix is in ***row-echelon*** *form* if:

1. All zero rows are at the bottom.
2. The first nonzero entry of a row is to the *right* of the first nonzero entry of the row above.
3. Below the first nonzero entry of a row, all entries are zero.

Here is a picture of a matrix in row-echelon form:

$$\begin{bmatrix} \boxed{\star} & \star & \star & \star & \star \\ 0 & \boxed{\star} & \star & \star & \star \\ 0 & 0 & 0 & \boxed{\star} & \star \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} \star &= \text{any number} \\ \boxed{\star} &= \text{any nonzero number} \end{aligned}$$

Definition: Pivot

A ***pivot*** is the first nonzero entry of a row of a matrix in row-echelon form.

A matrix in row-echelon form is generally easy to solve using back-substitution. For example,

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 10 & 30 \end{array}\right] \xrightarrow{\text{becomes}} \begin{cases} x + 2y + 3z = 6 \\ y + 2z = 4 \\ 10z = 30. \end{cases}$$

We immediately see that $z = 3$, which implies $y = 4 - 2 \cdot 3 = -2$ and $x = 6 - 2(-2) - 3 \cdot 3 = 1$. See Example 1.2.2.3.

Definition: Reduced Row-Echelon Form

A matrix is in ***reduced row-echelon form*** if it is in row-echelon form, and in addition:

4. Each pivot is equal to 1.
5. Each pivot is the only nonzero entry in its column.

Here is a picture of a matrix in reduced row-echelon form:

$$\begin{bmatrix} 1 & 0 & \star & 0 & \star \\ 0 & 1 & \star & 0 & \star \\ 0 & 0 & 0 & 1 & \star \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{aligned} \star &= \text{any number} \\ 1 &= \text{pivot} \end{aligned}$$

A matrix in reduced row-echelon form is in some sense completely solved. For example,

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right] \xrightarrow{\text{becomes}} \begin{cases} x = 1 \\ y = -2 \\ z = 3. \end{cases}$$

Illustration 1.2.2.1

The following matrices are in reduced row-echelon form:

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 8 & 0 \end{bmatrix} \quad \left[\begin{array}{cc|c} 1 & 17 & 0 \\ 0 & 0 & 1 \end{array}\right] \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The following matrices are in row-echelon form but not reduced row-echelon form:

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \quad \left[\begin{array}{ccc|c} 2 & 7 & 1 & 4 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{array}\right] \quad \begin{bmatrix} 1 & 17 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

The following matrices are not in echelon form:

$$\left[\begin{array}{ccc|c} 2 & 7 & 1 & 4 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right] \quad \left[\begin{array}{cc|c} 0 & 17 & 0 \\ 0 & 2 & 1 \end{array}\right] \quad \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

When deciding if an augmented matrix is in (reduced) row-echelon form, there is nothing special about the augmented column(s). Just ignore the vertical line.

If an augmented matrix is in reduced row-echelon form, the corresponding linear system is viewed as *solved*. We will see below why this is the case, and we will show that any matrix can be put into reduced row-echelon form using only row operations.

Theorem 1.2.2.1

Every matrix is row-equivalent to one and only one matrix in reduced row-echelon form.

We will give an algorithm, called ***row reduction*** or ***Gaussian elimination***, which demonstrates that every matrix is row-equivalent to *at least one* matrix in reduced row-echelon form.

The uniqueness statement is interesting—it means that, no matter *how* you row reduce, you *always* get the same matrix in reduced row-echelon form.

This assumes, of course, that you only do the three legal row operations, and you don't make any arithmetic errors.

We will not prove uniqueness, but maybe you can!

Algorithm: Row Reduction

- *Step 1a:* Swap the 1st row with a lower one so a leftmost nonzero entry is in the 1st row (if necessary).
- *Step 1b:* Scale the 1st row so that its first nonzero entry is equal to 1.
- *Step 1c:* Use row replacement so all entries below this 1 are 0.
- *Step 2a:* Swap the 2nd row with a lower one so that the leftmost nonzero entry is in the 2nd row.
- *Step 2b:* Scale the 2nd row so that its first nonzero entry is equal to 1.
- *Step 2c:* Use row replacement so all entries below this 1 are 0.
- *Step 3a:* Swap the 3rd row with a lower one so that the leftmost nonzero entry is in the 3rd row.
- etc.
- *Last Step:* Use row replacement to clear all entries above the pivots, starting with the last pivot.

Example 1.2.2.4

Row reduce this matrix:

$$\left[\begin{array}{ccc|c} 0 & -7 & -4 & 2 \\ 2 & 4 & 6 & 12 \\ 3 & 1 & -1 & -2 \end{array}\right]$$

Solution

$$\begin{pmatrix} 0 & -7 & -4 & 2 \\ 2 & 4 & 6 & 12 \\ 3 & 1 & -1 & -2 \end{pmatrix}$$

Step 1a: Row swap to make this nonzero.

$$R_1 \longleftrightarrow R_2 \quad \begin{pmatrix} 2 & 4 & 6 & 12 \\ 0 & -7 & -4 & 2 \\ 3 & 1 & -1 & -2 \end{pmatrix}$$

Step 1b: Scale to make this 1.

$$R_1 = R_1 \div 2 \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 3 & 1 & -1 & -2 \end{pmatrix}$$

Step 1c: Subtract a multiple of the first row to clear this.

Optional: swap rows 2 and 3 to make Step 2b easier next.

$$R_3 = R_3 - 3R_1 \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & -5 & -10 & -20 \end{pmatrix}$$

$$R_2 \longleftrightarrow R_3 \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & -5 & -10 & -20 \\ 0 & -7 & -4 & 2 \end{pmatrix}$$

Step 2a: This is already nonzero.
Step 2b: Scale to make this 1.

Note how Step 2b doesn't create fractions.

$$R_2 = R_2 \div -5 \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & -7 & -4 & 2 \end{pmatrix}$$

Step 2c: Add 7 times the second row to clear this.

$$R_3 = R_3 + 7R_2 \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 10 & 30 \end{pmatrix}$$

Step 3a: This is already nonzero.
Step 3b: Scale to make this 1.

$$R_3 = R_3 \div 10 \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Last step: add multiples of the third row to clear these.

$$R_2 = R_2 - 2R_3 \quad \begin{pmatrix} 1 & 2 & 3 & 6 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

$$R_1 = R_1 - 3R_3 \quad \begin{pmatrix} 1 & 2 & 0 & -3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Last step: add -2 times the third row to clear this.

$$R_1 = R_1 - 2R_2 \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

The reduced row-echelon form of the matrix is

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right] \xrightarrow{\text{translates to}} \begin{cases} x & & & = & 1 \\ & y & & = & -2 \\ & & z & = & 3 \end{cases}$$

The reduced row-echelon form of the matrix tells us that the only solution is $(x, y, z) = (1, -2, 3)$

Here is the row-reduction algorithm, summarized in pictures.

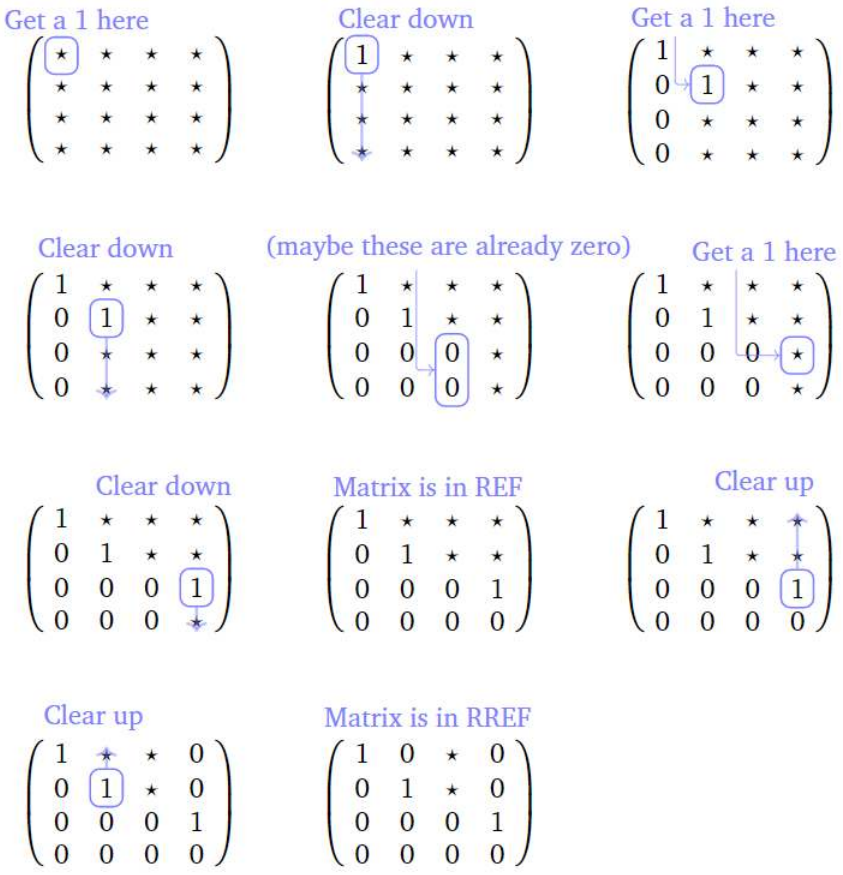


Figure 1.2.2.3

It will be very important to know where the pivots of a matrix after row reducing are; this is the reason for the following piece of terminology.

Definition: Pivot Position and Pivot Column

A ***pivot position*** of a matrix is an entry that is a pivot of a row-echelon form of that matrix.

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

Example 1.2.2.5: Pivot Positions

Find the pivot positions and pivot columns of this matrix

$$A = \left[\begin{array}{ccc|c} 0 & -7 & -4 & 2 \\ 2 & 4 & 6 & 12 \\ 3 & 1 & -1 & -2 \end{array}\right]$$

Solution

We saw in Example 1.2.2.5 that a row-echelon form of the matrix is

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 10 & 30 \end{array}\right]$$

The pivot positions of $\boldsymbol{A}$ are the entries that become pivots in a row-echelon form; they are marked in red below:

$$\left[\begin{array}{ccc|c} \color{red}{0} & -7 & -4 & 2 \\ 2 & \color{red}{4} & 6 & 12 \\ 3 & 1 & \color{red}{-1} & -2 \end{array}\right]$$

The first, second, and third columns are pivot columns.

Illustration 1.2.2.2: An Inconsistent System

If we attempt to solve the linear system

$$\begin{cases} 2x + 10y = -1 \\ 3x + 15y = 2 \end{cases}$$

using row reduction, we see:

$$\left[\begin{array}{cc|c} 2 & 10 & -1 \\ 3 & 15 & 2 \end{array}\right] \xrightarrow{R_1 = R_1 \div 2} \left[\begin{array}{cc|c} \color{red}{1} & 5 & -\frac{1}{2} \\ 3 & 15 & 2 \end{array}\right] \quad \text{(Step 1b)}$$

$$\xrightarrow{R_2 = R_2 - 3R_1} \left[\begin{array}{cc|c} 1 & 5 & -\frac{1}{2} \\ \color{red}{0} & 0 & \frac{7}{2} \end{array}\right] \quad \text{(Step 1c)}$$

$$\xrightarrow{R_2 = R_2 \times \frac{2}{7}} \left[\begin{array}{cc|c} 1 & 5 & -\frac{1}{2} \\ 0 & 0 & \color{red}{1} \end{array}\right] \quad \text{(Step 2b)}$$

$$\xrightarrow{R_1 = R_1 + \frac{1}{2}R_2} \left[\begin{array}{cc|c} 1 & 5 & \color{red}{0} \\ 0 & 0 & 1 \end{array}\right] \quad \text{(Step 2c)}$$

This row-reduced matrix corresponds to the *inconsistent* system

$$\begin{cases} x + 5y = 0 \\ 0 = 1 \end{cases}$$

In the above example, we saw how to recognize the reduced row-echelon form of an inconsistent system.

Theorem 1.2.2.2: The Row-Echelon Form of an Inconsistent System

An augmented matrix corresponds to an inconsistent system of equations if and only if the last column (i.e., the augmented column) is a *pivot column*.

In other words, the row-reduced matrix of an inconsistent system looks like this:

$$\left[\begin{array}{cccc|c} 1 & 0 & \star & \star & \color{red}{0} \\ 0 & 1 & \star & \star & \color{red}{0} \\ 0 & 0 & 0 & 0 & \color{red}{1} \end{array}\right]$$

We have discussed two classes of matrices so far:

1. When the reduced row-echelon form of a matrix has a pivot in every non-augmented column, then it corresponds to a system with a unique solution:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right] \xrightarrow{\text{translates to}} \begin{cases} x = 1 \\ y = -2 \\ z = 3 \end{cases}$$

2. When the reduced row-echelon form of a matrix has a pivot in the last (augmented) column, then it corresponds to a system with no solutions:

$$\left[\begin{array}{cc|c} 1 & 5 & 0 \\ 0 & 0 & 1 \end{array}\right] \xrightarrow{\text{translates to}} \begin{cases} x + 5y = 0 \\ 0 = 1 \end{cases}$$

1.2.2: Echelon Forms is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.2.3: Free Variables

What happens when one of the non-augmented columns lacks a pivot?

Illustration 1.2.3.3: A System with a Free Variable:

Consider the linear system

$$\begin{cases} 2x + y + 12z = 1 \\ x + 2y + 9z = -1 \end{cases}$$

We solve it using row-reduction:

$$\left[\begin{array}{ccc|c} 2 & 1 & 12 & 1 \\ 1 & 2 & 9 & -1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & 2 & 9 & -1 \\ 2 & 1 & 12 & 1 \end{array}\right] \quad \text{(Optional)}$$

$$\xrightarrow{R_2 = R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & 2 & 9 & -1 \\ 0 & -3 & -6 & 3 \end{array}\right] \quad \text{(Step 1c)}$$

$$\xrightarrow{R_2 = R_2 \cdot \left(-\frac{1}{3}\right)} \left[\begin{array}{ccc|c} 1 & 2 & 9 & -1 \\ 0 & 1 & 2 & -1 \end{array}\right] \quad \text{(Step 2b)}$$

$$\xrightarrow{R_1 = R_1 - 2R_2} \left[\begin{array}{ccc|c} 1 & 0 & 5 & 1 \\ 0 & 1 & 2 & -1 \end{array}\right] \quad \text{(Step 2c)}$$

This row-reduced matrix corresponds to the linear system

$$\begin{cases} x + 5z = 1 \\ y + 2z = -1 \end{cases}$$

In what sense is the system solved? We rewrite as

$$\begin{cases} x = 1 - 5z \\ y = -1 - 2z \end{cases}$$

For any value of z, there is exactly one value of x and y that make the equations true. But we are free to choose *any* value of z.

We have found all solutions, they are the set of all values x, y and z which look like:

$$\begin{cases} x = 1 - 5z \\ y = -1 - 2z \\ z = z \end{cases} \qquad z \text{ is any real number}$$

This is the *parametric form* for the solution to the linear system. The variable z is called a *free variable*.

Graphically, this represents a line in the xyz-space, like we have seen in the previous section.

Note 1.2.3.1

Given the parametric form for the solution to a linear system, we can obtain specific solutions by replacing the free variables with any specific real numbers. For instance, setting $z = 0$ in the last example gives the solution $(x, y, z) = (1, -1, 0)$, and setting $z = 1$ gives the solution $(x, y, z) = (-4, -3, 1)$.

Definition: Free Variable

Consider a *consistent* system of equations in the variables $x_1, x_2, \ldots, x_n$. Let A be a row-echelon form of the augmented matrix for this system.

We say that x_i is a ***free variable*** if its corresponding column in A is *not* a pivot column.

In Illustration 1.2.3.3, the variable z was free because the reduced row-echelon form matrix was

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

In the matrix

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

the free variables are x_2 and x_4. (The augmented column is not free because it does not correspond to a variable.)

Recipe: Parametric form

The ***parametric form*** of the solution set of a consistent system of linear equations is obtained as follows.

1. Write the system as an augmented matrix.
2. Row reduce to reduced row-echelon form.
3. Write the corresponding (solved) system of linear equations.
4. *Move all free variables to the right hand side of the equations.*

Moving the free variables to the right hand side of the equations amounts to solving for the non-free variables (the ones that become pivot columns) in terms of the free variables. One can think of the free variables as being *independent* variables, and the non-free variables being *dependent*.

Note 1.2.3.2

You can choose *any value* for the free variables in a (consistent) linear system.

Free variables come from the *columns without pivots* in a matrix in row-echelon form.

Illustration 1.2.3.4

Suppose that the reduced row-echelon form of the matrix for a linear system in four variables x_1, x_2, x_3, x_4, is

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

The free variables are x_2 and x_4: they are the ones whose columns are *not* pivot columns.

This translates into the system of equations

$$\left\{\begin{array}{rcrcl} x_1 & & + 3x_4 & = & 2 \\ & x_3 & + 4x_4 & = & -1 \end{array}\right. \xrightarrow{\text{parametric form}} \left\{\begin{array}{rcrcl} x_1 & = & 2 & - & 3x_4 \\ x_3 & = & -1 & - & 4x_4 \end{array}\right.$$

What happened to x_2? It is a free variable, but no other variable depends on it. The general solution to the system is

$$(x_1, x_2, x_3, x_4) = (2 - 3x_4, x_2, -1 - 4x_4, x_4)$$

for any values of x_2 and x_4. For instance, $(2, 0, -1, 0)$ is a solution (with $x_2 = x_4 = 0$), and $(5, 1, 3, -1)$ is a solution (with $x_2 = 1, x_4 = -1$).

Illustration 1.2.3.5: A Parameterized Plane

The system of one linear equation

$$x + y + z = 1$$

comes from the matrix

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

which is already in reduced row-echelon form. The free variables are y and z. The parametric form for the general solution is

$$(x, y, z) = (1 - y - z, y, z)$$

for any values of y and z. This is the parametric equation for a plane in $\mathbb{R}^3$.

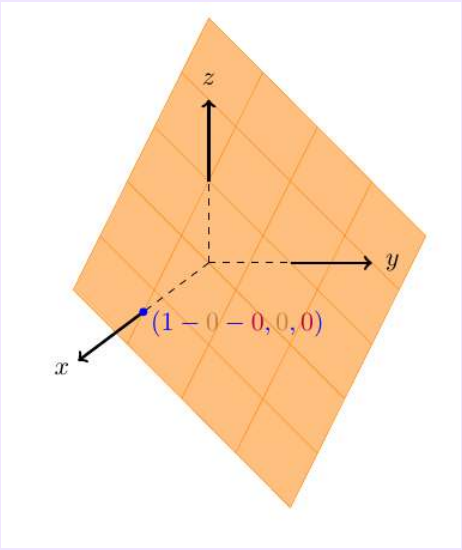


Figure **1.2.3.1**: A plane described by two parameters y and z. Any point on the plane is obtained by substituting suitable values for y and z.

There are *three possibilities* for the reduced row-echelon form of the augmented matrix of a linear system.

1. ***The last column (right of the augment line) is a pivot column.*** In this case, the system is *inconsistent*. There are zero solutions, i.e., the solution set is empty. For example, the matrix

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

   comes from a linear system with no solutions.
2. ***Every column except the last column is a pivot column.*** In this case, the system has a *unique* solution. For example, the matrix

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

   tells us that the unique solution is $(x, y, z) = (a, b, c)$.
3. ***The last column is not a pivot column, and some other column is not a pivot column either.*** In this case, the system has *infinitely many* solutions, corresponding to the infinitely many possible values of the free variable(s). For example, in the system corresponding to the matrix

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

   any values for x_2 and x_4 yield a solution to the system of equations.

1.2.3: Free Variables is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.2.4: Practice Set

Exercise 1.2.4.1

Determine if each matrix below is in Echelon Form, Reduced Row Echelon Form, or Neither:

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

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

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

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

(e) $\left[\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$

(f) $\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \end{array}\right]$

Answer

(a) Echelon Form (pivots not 1's)

(b) Reduced Row Echelon Form

(c) Reduced Row Echelon Form

(d) Neither (row of zeros not on bottom)

(e) Reduced Row Echelon Form

(f) Echelon Form (pivots not 1's)

Exercise 1.2.4.2a

Solve the system of equations $\begin{cases} x+y=9 \\ 2x-y=3 \end{cases}$ by setting up an augmented matrix and row reducing.

Answer

$$\left[\begin{array}{cc|c} 1 & 1 & 9 \\ 2 & -1 & 3 \end{array}\right] \xrightarrow{R_2=R_2-2R_1} \left[\begin{array}{cc|c} 1 & 1 & 9 \\ 0 & -3 & -15 \end{array}\right] \xrightarrow{R_2=-\frac{1}{3}R_2} \left[\begin{array}{cc|c} 1 & 1 & 9 \\ 0 & 1 & 5 \end{array}\right] \xrightarrow{R_1=R_1-R_2} \left[\begin{array}{cc|c} 1 & 0 & 4 \\ 0 & 1 & 5 \end{array}\right]$$

So the solution set is the point $(x,y)=(4,5)$

Exercise 1.2.4.2b

Solve the system of equations $\begin{cases} 4x-3y=2 \\ x+y=4 \\ 3x-y=4 \end{cases}$ by setting up an augmented matrix and row reducing.

Answer

$$\left[\begin{array}{cc|c} 4 & -3 & 2 \\ 1 & 1 & 4 \\ 3 & -1 & 4 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cc|c} 1 & 1 & 4 \\ 4 & -3 & 2 \\ 3 & -1 & 4 \end{array}\right] \xrightarrow{R_2 = R_2 - 4R_1} \left[\begin{array}{cc|c} 1 & 1 & 4 \\ 0 & -7 & -14 \\ 3 & -1 & 4 \end{array}\right] \xrightarrow{R_3 = R_3 - 3R_1} \left[\begin{array}{cc|c} 1 & 1 & 4 \\ 0 & -7 & -14 \\ 0 & -4 & -8 \end{array}\right]$$
$$\xrightarrow{R_2 = -\frac{1}{7}R_2} \left[\begin{array}{cc|c} 1 & 1 & 4 \\ 0 & 1 & 2 \\ 0 & -4 & -8 \end{array}\right] \xrightarrow{R_1 = R_1 - R_2} \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & -4 & -8 \end{array}\right] \xrightarrow{R_3 = R_3 + 4R_2} \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right]$$

So the solution set is the point $(x, y) = (2, 2)$

Exercise 1.2.4.2c

Solve the system of equations $\begin{cases} 3x + 2y = 7 \\ 4x - y = 11 \\ x - 3y = 3 \end{cases}$ by setting up an augmented matrix and row reducing.

Answer

$$\left[\begin{array}{cc|c} 3 & 2 & 7 \\ 4 & -1 & 11 \\ 1 & -3 & 3 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{cc|c} 1 & -3 & 3 \\ 4 & -1 & 11 \\ 3 & 2 & 7 \end{array}\right] \xrightarrow{R_2 = R_2 - 4R_1} \left[\begin{array}{cc|c} 1 & -3 & 3 \\ 0 & 11 & -1 \\ 3 & 2 & 7 \end{array}\right] \xrightarrow{R_3 = R_3 - 3R_1} \left[\begin{array}{cc|c} 1 & -3 & 3 \\ 0 & 11 & -1 \\ 0 & 11 & -2 \end{array}\right]$$
$$\xrightarrow{R_2 = \frac{1}{11}R_2} \left[\begin{array}{cc|c} 1 & -3 & 3 \\ 0 & 1 & -\frac{1}{11} \\ 0 & 11 & -2 \end{array}\right] \xrightarrow{R_1 = R_1 + 3R_2} \left[\begin{array}{cc|c} 1 & 0 & \frac{30}{11} \\ 0 & 1 & -\frac{1}{11} \\ 0 & 11 & -2 \end{array}\right] \xrightarrow{R_3 = R_3 - 11R_2} \left[\begin{array}{cc|c} 1 & 0 & \frac{30}{11} \\ 0 & 1 & -\frac{1}{11} \\ 0 & 0 & -1 \end{array}\right]$$

This system is inconsistent, so the solution set is empty.

Exercise 1.2.4.2d

Solve the system of equations $\begin{cases} x - y + z = 6 \\ -x + y + z = 4 \\ x + y - z = 0 \end{cases}$ by setting up an augmented matrix and row reducing.

Answer

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 6 \\ -1 & 1 & 1 & 4 \\ 1 & 1 & -1 & 0 \end{array}\right] \xrightarrow{R_2 = R_2 + R_1} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 6 \\ 0 & 0 & 2 & 10 \\ 1 & 1 & -1 & 0 \end{array}\right] \xrightarrow{R_3 = R_3 - R_1} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 6 \\ 0 & 0 & 2 & 10 \\ 0 & 2 & -2 & -6 \end{array}\right] \xrightarrow{R_2 \leftrightarrow R_3} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 6 \\ 0 & 2 & -2 & -6 \\ 0 & 0 & 2 & 10 \end{array}\right]$$
$$\xrightarrow{R_2 = \frac{1}{2}R_2} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 6 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 2 & 10 \end{array}\right] \xrightarrow{R_1 = R_1 + R_2} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 2 & 10 \end{array}\right] \xrightarrow{R_3 = \frac{1}{2}R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & -1 & -3 \\ 0 & 0 & 1 & 5 \end{array}\right] \xrightarrow{R_2 = R_2 + R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 5 \end{array}\right]$$

So the solution set is the point $(x, y, z) = (3, 2, 5)$

Exercise 1.2.4.2e

Solve the system of equations $\begin{cases} 2x - y = 4 \\ x + z = 2 \\ y + 2z = 0 \end{cases}$ by setting up an augmented matrix and row reducing.

Answer

$$\left[\begin{array}{ccc|c} 2 & -1 & 0 & 4 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 0 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 2 & -1 & 0 & 4 \\ 0 & 1 & 2 & 0 \end{array}\right] \xrightarrow{R_2 = R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & -1 & -2 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right]$$
$$\xrightarrow{R_2 = -R_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right] \xrightarrow{R_3 = R_3 - R_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Since $x+z=2 \rightarrow x=2-z$ and $y+2z=0 \rightarrow y=-2z$ and z is a free variable,
the solution set is a line containing points of the form $(x,y,z)=(2-z,-2z,z)$

Exercise 1.2.4.2f

Solve the system of equations $\begin{cases} x+\frac{1}{2}y-z=-5 \\ x+\frac{3}{2}y+z=7 \end{cases}$ by setting up an augmented matrix and row reducing.

Answer

$$\left[\begin{array}{cccc|c} 1 & \frac{1}{2} & -1 & -5 \\ 1 & \frac{3}{2} & 1 & 7 \end{array}\right] \xrightarrow{R_2=R_2-R_1} \left[\begin{array}{cccc|c} 1 & \frac{1}{2} & -1 & -5 \\ 0 & 1 & 2 & 12 \end{array}\right] \xrightarrow{R_1=R_1-\frac{1}{2}R_2} \left[\begin{array}{cccc|c} 1 & 0 & -2 & -11 \\ 0 & 1 & 2 & 12 \end{array}\right]$$

Since $x-2z=-11 \rightarrow x=-11+2z$ and $y+2z=12 \rightarrow y=12-2z$ and z is a free variable,
the solution set is a line containing points of the form $(x,y,z)=(-11+2z,12-2z,z)$

Exercise 1.2.4.2g

Solve the system of equations $\begin{cases} x_1+x_2+x_3+x_4=6 \\ x_2-x_3+x_4=8 \\ x_3-x_4=5 \\ x_1+x_2+x_3=-1 \end{cases}$ by setting up an augmented matrix and row reducing.

Answer

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 6 \\ 0 & 1 & -1 & 1 & 8 \\ 0 & 0 & 1 & -1 & 5 \\ 1 & 1 & 1 & 0 & -1 \end{array}\right] \xrightarrow{R_4=R_4-R_1} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 6 \\ 0 & 1 & -1 & 1 & 8 \\ 0 & 0 & 1 & -1 & 5 \\ 0 & 0 & 0 & -1 & -7 \end{array}\right] \xrightarrow{R_1=R_1-R_2} \left[\begin{array}{cccc|c} 1 & 0 & 2 & 0 & -2 \\ 0 & 1 & -1 & 1 & 8 \\ 0 & 0 & 1 & -1 & 5 \\ 0 & 0 & 0 & -1 & -7 \end{array}\right]$$

$$\xrightarrow{R_1=R_1-2R_3} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 2 & -12 \\ 0 & 1 & -1 & 1 & 8 \\ 0 & 0 & 1 & -1 & 5 \\ 0 & 0 & 0 & -1 & -7 \end{array}\right] \xrightarrow{R_2=R_2+R_3} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 2 & -12 \\ 0 & 1 & 0 & 0 & 13 \\ 0 & 0 & 1 & -1 & 5 \\ 0 & 0 & 0 & -1 & -7 \end{array}\right] \xrightarrow{R_4=-R_4} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 2 & -12 \\ 0 & 1 & 0 & 0 & 13 \\ 0 & 0 & 1 & -1 & 5 \\ 0 & 0 & 0 & 1 & 7 \end{array}\right]$$

$$\xrightarrow{R_1=R_1-2R_4} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -26 \\ 0 & 1 & 0 & 0 & 13 \\ 0 & 0 & 1 & -1 & 5 \\ 0 & 0 & 0 & 1 & 7 \end{array}\right] \xrightarrow{R_3=R_3+R_4} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -26 \\ 0 & 1 & 0 & 0 & 13 \\ 0 & 0 & 1 & 0 & 12 \\ 0 & 0 & 0 & 1 & 7 \end{array}\right]$$

So the solution set is the point $(x_1,x_2,x_3,x_4)=(-26,13,12,7)$

Exercise 1.2.4.2h

Solve the system of equations $\begin{cases} x_1+2x_2-x_4=5 \\ x_1+x_3+x_4=0 \\ 2x_2-x_3-2x_4=5 \\ 2x_1+6x_2-x_3-4x_4=15 \end{cases}$ by setting up an augmented matrix and row reducing.

Answer

$$\left[\begin{array}{cccc|c}1&2&0&-1&5\\1&0&1&1&0\\0&2&-1&-2&5\\2&6&-1&-4&15\end{array}\right]\xrightarrow{R_2=R_2-R_1}\left[\begin{array}{cccc|c}1&2&0&-1&5\\0&-2&1&2&-5\\0&2&-1&-2&5\\2&6&-1&-4&15\end{array}\right]\xrightarrow{R_4=R_4-2R_1}\left[\begin{array}{cccc|c}1&2&0&-1&5\\0&-2&1&2&-5\\0&2&-1&-2&5\\0&2&-1&-2&5\end{array}\right]$$

$$\xrightarrow{R_2=-\frac{1}{2}R_2}\left[\begin{array}{cccc|c}1&2&0&-1&5\\0&1&-\frac{1}{2}&-1&\frac{5}{2}\\0&2&-1&-2&5\\0&2&-1&-2&5\end{array}\right]\xrightarrow{R_1=R_1-2R_2}\left[\begin{array}{cccc|c}1&0&1&1&0\\0&1&-\frac{1}{2}&-1&\frac{5}{2}\\0&2&-1&-2&5\\0&2&-1&-2&5\end{array}\right]\xrightarrow{R_3=R_3-2R_2}\left[\begin{array}{cccc|c}1&0&1&1&0\\0&1&-\frac{1}{2}&-1&\frac{5}{2}\\0&0&0&0&0\\0&2&-1&-2&5\end{array}\right]$$

$$\xrightarrow{R_4=R_4-2R_2}\left[\begin{array}{cccc|c}1&0&1&1&0\\0&1&-\frac{1}{2}&-1&\frac{5}{2}\\0&0&0&0&0\\0&0&0&0&0\end{array}\right]$$

Since $x_1+x_3+x_4=0\rightarrow x_1=-x_3-x_4$ and $x_2-\frac{1}{2}x_3-x_4=\frac{5}{2}\rightarrow x_2=\frac{1}{2}x_3+x_4+\frac{5}{2}$ and x_3, x_4 are free variables, the solution set is the plane containing points of the form $(x_1,x_2,x_3,x_4)=\left(-x_3-x_4,\frac{1}{2}x_3+x_4+\frac{5}{2},x_3,x_4\right)$

? Exercise 1.2.4.3*a*

Determine the value of h which makes the system below have (if possible) (i) no solution, (ii) exactly one solution, (iii) infinitely many solutions.

$$\begin{cases}x+2y=1\\x+hy=2\end{cases}$$

Answer

$$\left[\begin{array}{cc|c}1&2&1\\1&h&2\end{array}\right]\xrightarrow{R_2=R_2-R_1}\left[\begin{array}{cc|c}1&2&1\\0&h-2&1\end{array}\right]$$

This is in echelon form.

(i) If $h-2=0$, then this system is inconsistent, so $h=2$.

(ii) If $h-2\neq 0$ then this system will be consistent with no free variables, so $h\neq 2$.

(iii) It is impossible for this system to have a free variable, so infinitely many solutions is IMPOSSIBLE regardless of the choice of h.

? Exercise 1.2.4.3*b*

Determine the value of h which makes the system below have (if possible) (i) no solution, (ii) exactly one solution, (iii) infinitely many solutions.

$$\begin{cases}3x-2y=1\\2x-3y=h\end{cases}$$

Answer

$$\left[\begin{array}{cc|c}3&-2&1\\2&-3&h\end{array}\right]\xrightarrow{R_2=R_2-\frac{2}{3}R_1}\left[\begin{array}{cc|c}3&-2&1\\0&-\frac{5}{3}&h-\frac{2}{3}\end{array}\right]$$

This is in echelon form.

(i) It is impossible for this system to be inconsistent, regardless of the choice of h.

(ii) This system will be consistent with no free variables for all possible values of h.

(iii) It is impossible for this system to have a free variable, so infinitely many solutions is IMPOSSIBLE regardless of the choice of h.

Exercise 1.2.4.3c

Determine the value of h which makes the system below have (if possible) (i) no solution, (ii) exactly one solution, (iii) infinitely many solutions.

$$\begin{cases} hx + hy = 0 \\ x + y = 1 \end{cases}$$

Answer

$$\left[\begin{array}{cc|c} h & h & 0 \\ 1 & 1 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ h & h & 0 \end{array}\right] \xrightarrow{R_2 = R_2 - hR_1} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & -h \end{array}\right]$$

This is in echelon form.

(i) If $h \neq 0$ this system will be inconsistent.

(ii) There will always be a free variable. It is IMPOSSIBLE to have exactly one solution regardless of the choice of h.

(iii) If $h = 0$, the system will be consistent and y will be a free variable, yielding infinitely many solutions.

Exercise 1.2.4.3d

Determine the value of h which makes the system below have (if possible) (i) no solution, (ii) exactly one solution, (iii) infinitely many solutions.

$$\begin{cases} hx + hy = 1 \\ x + y = 0 \end{cases}$$

Answer

$$\left[\begin{array}{cc|c} h & h & 1 \\ 1 & 1 & 0 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ h & h & 1 \end{array}\right] \xrightarrow{R_2 = R_2 - hR_1} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$$

This is in echelon form.

(i) This system will always be inconsistent regardless of the choice of h.

(ii) This system is never consistent. It is IMPOSSIBLE to have exactly one solution regardless of the choice of h.

(iii) This system is never consistent, so infinitely many solutions is IMPOSSIBLE regardless of the choice of h.

Exercise 1.2.4.3e

Determine the value of h which makes the system below have (if possible) (i) no solution, (ii) exactly one solution, (iii) infinitely many solutions.

$$\begin{cases} x - y - z = 1 \\ hx + 2y + 2z = 2 \end{cases}$$

Answer

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 1 \\ h & 2 & 2 & 2 \end{array}\right] \xrightarrow{R_2 = R_2 - hR_1} \left[\begin{array}{ccc|c} 1 & -1 & -1 & 1 \\ 0 & 2+h & 2+h & 2-h \end{array}\right]$$

This is in echelon form.

(i) If $2 + h = 0$ and $2 - h \neq 0$, this system will always be inconsistent. So no solution occurs when $h = -2$.

(ii) There will always be a free variable. It is IMPOSSIBLE to have exactly one solution regardless of the choice of h.

(iii) When $h \neq -2$ the system will be consistent and z will be a free variable, yielding infinitely many solutions.

1.2.4: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.3: Vectors

Objectives

1. Learn how to add and scale vectors in $\mathbb{R}^n$, both algebraically and geometrically.
2. Understand linear combinations geometrically.
3. Learn the definition of $\text{Span}\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ and how to draw pictures of spans.
4. *Recipe:* solve a vector equation using augmented matrices / decide if a vector is in a span.
5. *Pictures:* vector addition, vector subtraction, linear combinations, an inconsistent system of equations, a consistent system of equations, spans in $\mathbb{R}^2$ and $\mathbb{R}^3$.
6. *Vocabulary words:* ***vector***, ***linear combination, vector equation, span***

We have been drawing points in $\mathbb{R}^n$ as dots in the line, plane, space, etc. We can also draw them as *arrows*. Since we have two geometric interpretations in mind, we can now discuss the relationship between the two points of view.

This page titled 1.3: Vectors is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **2.1: Vectors** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.3.1: Vector Algebra and Geometry

Definition: Points and Vectors

Again, a ***point*** in $\mathbb{R}^n$ is drawn as a dot. A ***vector*** is a line segment between two points with direction (known as a ray) in $\mathbb{R}^n$, drawn as an arrow.

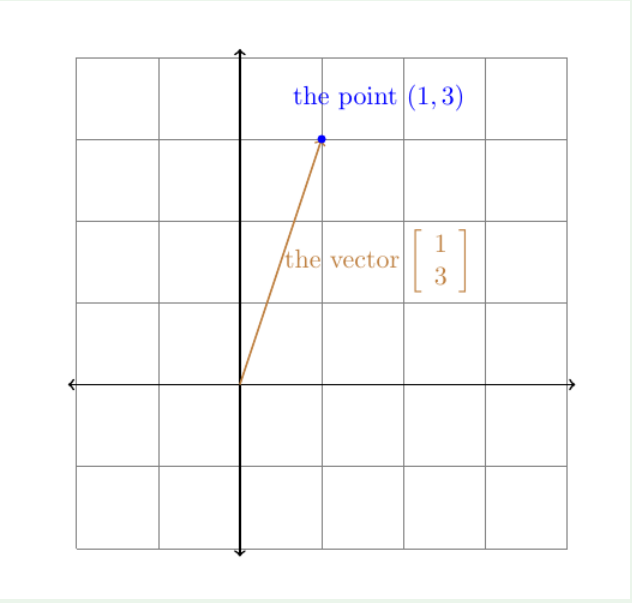


Figure 1.3.1.1

The difference is purely psychological: *points and vectors are both just lists of numbers*.

When we think of a point in $\mathbb{R}^n$ as a vector, we will usually write it vertically, like a matrix with one column:

$$v = \begin{bmatrix} a \\ b \end{bmatrix}$$

We will also write $\vec{0}$ for the zero vector, which is the vector with all entries equal to zero.

Why make the distinction between points and vectors? A vector need not start at the origin: *it can be located anywhere*! In other words, an arrow is determined by its length and its direction, not by its location. For instance, these arrows all represent the vector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

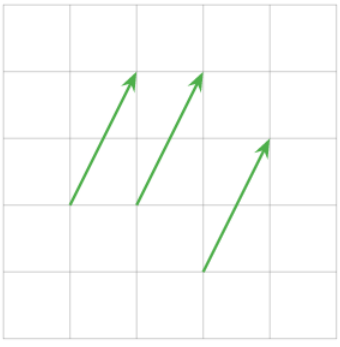


Figure 1.3.1.2

Note 1.3.1.1

Unless otherwise specified, we will assume that all vectors start at the origin.

Vectors make sense in the real world: many physical quantities, such as velocity, are represented as vectors. But it makes more sense to think of the velocity of a car as emanating from the current location of the car.

Remark

Some authors use boldface letters to represent vectors, as in " $\mathbf{v}$ ", while others use arrows, as in " $\vec{v}$ ". These are equivalent notations!

Note 1.3.1.2

Another way to think about a vector is as a *difference* between two points, or the line segment from one point to another.

For instance, $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is the arrow from $(1,1)$ to $(2,3)$.

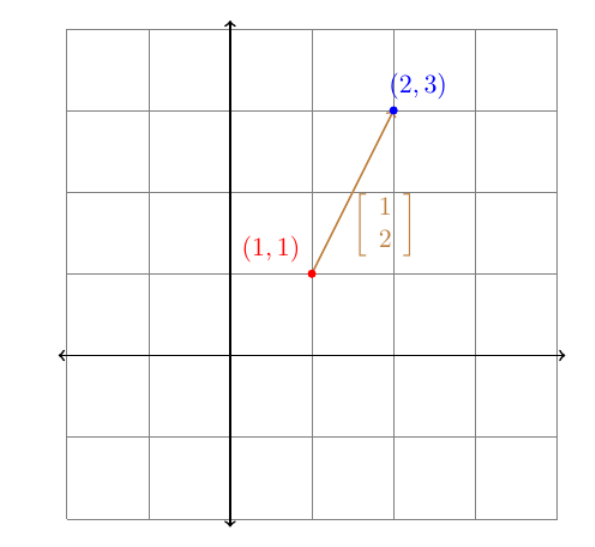


Figure 1.3.1.3

Here we learn how to add vectors together and how to multiply vectors by numbers, both algebraically and geometrically.

Definition: Vector addition and Scalar multiplication

- We can add two vectors together:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

- We can multiply, or ***scale***, a vector by a real number c:

$$c \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} c \cdot x_1 \\ c \cdot x_2 \\ \vdots \\ c \cdot x_n \end{bmatrix}$$

We call c a ***scalar*** to distinguish it from a vector. If $\vec{v}$ is a vector and c is a scalar, then $c\vec{v}$ is called a **scalar multiple** of $\vec{v}$.

Illustration 1.3.1.1

$$\begin{bmatrix}1\\2\\3\end{bmatrix}+\begin{bmatrix}4\\5\\6\end{bmatrix}=\begin{bmatrix}5\\7\\9\end{bmatrix} \quad \text{and} \quad -2\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}-2\\-4\\-6\end{bmatrix}$$

Geometrically, the sum of two vectors $\vec{v},\vec{w}$ is obtained as follows: place the tail of $\vec{w}$ at the head of $\vec{v}$. Then $\vec{v}+\vec{w}$ is the vector whose tail is the tail of $\vec{v}$ and whose head is the head of $\vec{w}$. Doing this both ways creates a parallelogram. For example,

$$\begin{bmatrix}1\\3\end{bmatrix}+\begin{bmatrix}4\\2\end{bmatrix}=\begin{bmatrix}5\\5\end{bmatrix}$$

Why? The width of $\vec{v}+\vec{w}$ is the sum of the widths, and likewise, the height of $\vec{v}+\vec{w}$ is the sum of the heights.

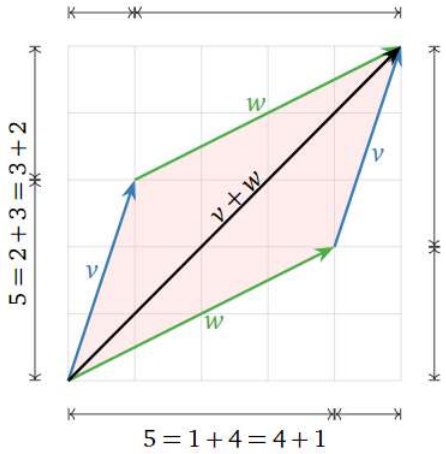


Figure 1.3.1.4

Geometrically, the difference of two vectors $\vec{v},\vec{w}$ is obtained as follows: place the tail of $\vec{v}$ and $\vec{w}$ at the same point. Then $\vec{v}-\vec{w}$ is the vector from the head of $\vec{w}$ to the head of $\vec{v}$. For example,

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

Why? To make subtraction of vectors act like subtraction of numbers! If you add $\vec{v}-\vec{w}$ to $\vec{w}$, you get $\vec{v}$.

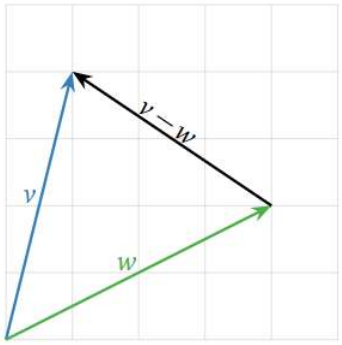


Figure 1.3.1.5

A scalar multiple of a vector $\vec{v}$ has the same (or opposite) direction, but a different length. For instance, $2\vec{v}$ is the vector in the direction of $\vec{v}$ but twice as long, and $-\frac{1}{2}\vec{v}$ is the vector in the opposite direction of $\vec{v}$, but half as long. Note that the set of all scalar multiples of a (nonzero) vector $\vec{v}$ is a *line*.

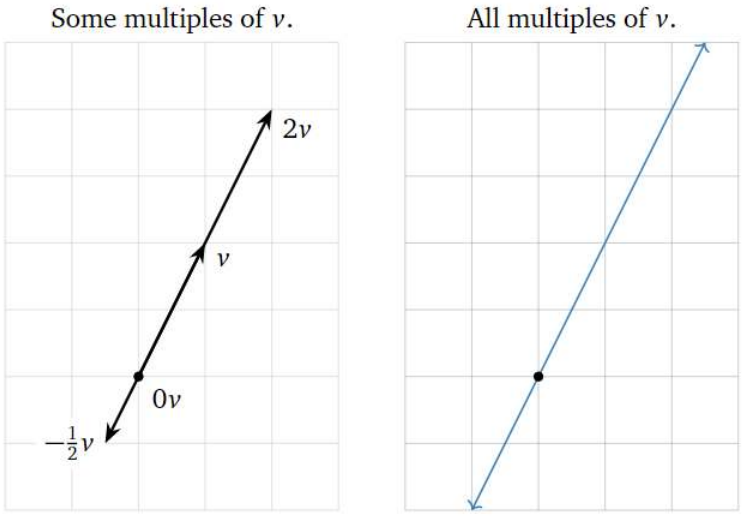


Figure **1.3.1.6**

We can add and scale vectors in the same equation.

Definition: Linear Combination

Let $c_1, c_2, \ldots, c_k$ be scalars, and let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ be vectors in $\mathbb{R}^n$. The vector in $\mathbb{R}^n$

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$$

is called a ***linear combination*** of the vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$, with ***weights*** or ***coefficients*** $c_1, c_2, \ldots, c_k$.

Geometrically, a linear combination is obtained by stretching / shrinking the vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ according to the coefficients, then adding them together using the parallelogram law.

Illustration 1.3.1.2

Let $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Here are some linear combinations of $\vec{v}_1$ and $\vec{v}_2$, drawn as points.

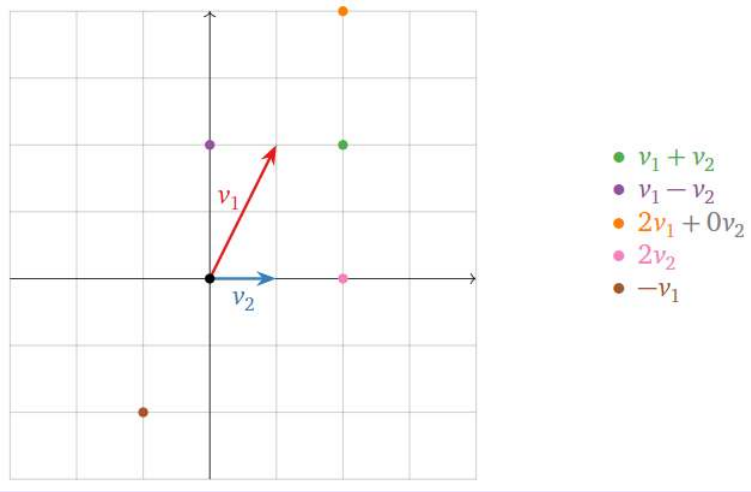


Figure **1.3.1.7**

The locations of these points are found using the parallelogram law for vector addition. We can see that any vector on the plane is a linear combination of $\vec{v}_1$ and $\vec{v}_2$, with suitable coefficients.

Illustration 1.3.1.3: Linear Combinations of a Single Vector

A linear combination of a single vector $v = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is just a scalar multiple of v. So some examples include

$$\vec{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \frac{3}{2}\vec{v} = \begin{bmatrix} \frac{3}{2} \\ 3 \end{bmatrix}, \quad -\frac{1}{2}\vec{v} = \begin{bmatrix} -\frac{1}{2} \\ -1 \end{bmatrix}, \quad \cdots$$

The set of all linear combinations is the *line through* $\vec{v}$. (Unless $\vec{v} = \vec{0}$, in which case any scalar multiple of $\vec{v}$ is again $\vec{0}$.)

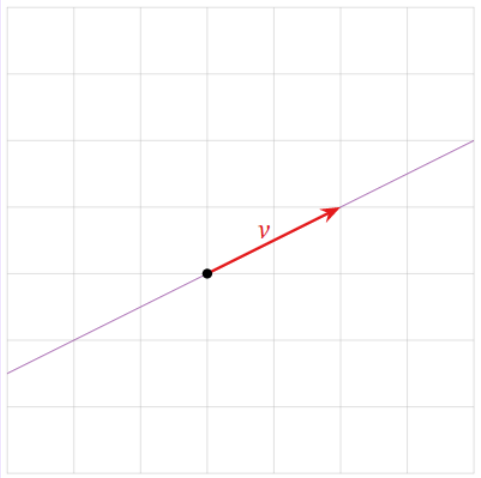


Figure 1.3.1.8

Illustration 1.3.1.4: Linear Combinations of Collinear Vectors

The set of all linear combinations of the vectors

$$\vec{v}_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

is the *line* containing both vectors.

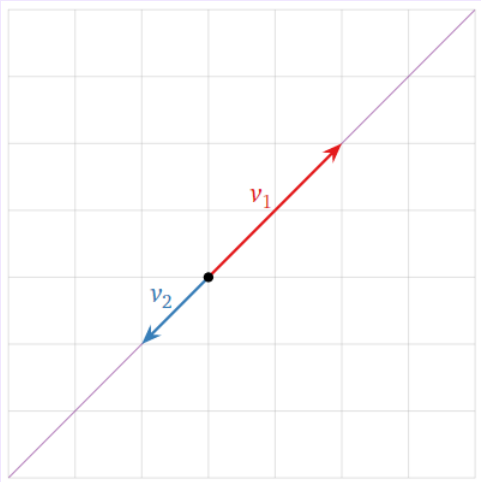


Figure 1.3.1.9

The difference between this and Example 1.3.1.2 is that both vectors lie on the same line. Hence any scalar multiples of $\vec{v}_1, \vec{v}_2$ lie on that line, as does their sum.

1.3.1: Vector Algebra and Geometry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.1: Vectors** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.3.2: Vector Equations and Spans

An equation involving vectors with n coordinates is the same as n equations involving only numbers. For example, the equation

$$x\begin{bmatrix}1\\2\\6\end{bmatrix}+y\begin{bmatrix}-1\\-2\\-1\end{bmatrix}=\begin{bmatrix}8\\16\\3\end{bmatrix}$$

simplifies to

$$\begin{bmatrix}x\\2x\\6x\end{bmatrix}+\begin{bmatrix}-y\\-2y\\-y\end{bmatrix}=\begin{bmatrix}8\\16\\3\end{bmatrix}\quad\text{or}\quad\begin{bmatrix}x-y\\2x-2y\\6x-y\end{bmatrix}=\begin{bmatrix}8\\16\\3\end{bmatrix}$$

For two vectors to be equal, all of their coordinates must be equal, so this is just the system of linear equations

$$\begin{cases}x-y=8\\2x-2y=16\\6x-y=3\end{cases}$$

Definition: Vector Equation

A ***vector equation*** is an equation involving a linear combination of vectors with possibly unknown coefficients.

Note 1.3.2.3

Asking whether or not a vector equation has a solution is the same as asking if a given vector is a linear combination of some other given vectors.

For example the vector equation above is asking if the vector $\begin{bmatrix}8\\16\\3\end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix}1\\2\\6\end{bmatrix}$ and $\begin{bmatrix}-1\\-2\\-1\end{bmatrix}$.

The thing we really care about is solving systems of linear equations, not solving vector equations. The whole point of vector equations is that they give us a different, and more geometric, way of viewing systems of linear equations.

In order to actually solve the vector equation

$$x\begin{bmatrix}1\\2\\6\end{bmatrix}+y\begin{bmatrix}-1\\-2\\-1\end{bmatrix}=\begin{bmatrix}8\\16\\3\end{bmatrix}$$

one has to solve the system of linear equations

$$\begin{cases}x-y=8\\2x-2y=16\\6x-y=3\end{cases}$$

This means forming the augmented matrix

$$\left[\begin{array}{cc|c}1&-1&8\\2&-2&16\\6&-1&3\end{array}\right]$$

and row reducing. Note that *the columns of the augmented matrix are the vectors from the original vector equation*, so it is not actually necessary to write the system of equations: one can go directly from the vector equation to the augmented matrix by "smooshing the vectors together".

Example 1.3.2.1

Is $\begin{bmatrix} 8 \\ 16 \\ 3 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix}$?

Solution

As discussed above, this question boils down to a row reduction:

$$\left[\begin{array}{cc|c} 1 & -1 & 8 \\ 2 & -2 & 16 \\ 6 & -1 & 3 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & -9 \\ 0 & 0 & 0 \end{array}\right]$$

From this we see that the equation is consistent, and the solution is $x = -1$ and $y = -9$. We conclude that $\begin{bmatrix} 8 \\ 16 \\ 3 \end{bmatrix}$ is indeed a linear combination of $\begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}$ and $\begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix}$, with coefficients -1 and -9:

$$\boxed{-\begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} - 9\begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} 8 \\ 16 \\ 3 \end{bmatrix}}$$

Recipe: Solving a Vector Equation

In general, the vector equation

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k = \vec{b}$$

where $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k, \vec{b}$ are vectors in $\mathbb{R}^n$ and $x_1, x_2, \ldots, x_k$ are unknown scalars, has the same solution set as the linear system with augmented matrix

$$\left[\begin{array}{ccc|c} | & | & | & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k & b \\ | & | & | & | \end{array}\right]$$

whose columns are the $\vec{v}_i$'s and the $\vec{b}$'s.

Now we have *three* equivalent ways of thinking about a linear system:

1. As a system of equations:

$$\begin{cases} 2x_1 + 3x_2 - 2x_3 = 7 \\ x_1 - x_2 - 3x_3 = 5 \end{cases}$$

2. As an augmented matrix:

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

3. As a vector equation ($x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n = \vec{b}$):

$$x_1\begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ -1 \end{bmatrix} + x_3\begin{bmatrix} -2 \\ -3 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$$

The third is geometric in nature: it lends itself to drawing pictures.

It will be important to know what are *all* linear combinations of a set of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ in $\mathbb{R}^n$. In other words, we would like to understand the set of all vectors $\vec{b}$ in $\mathbb{R}^n$ such that the vector equation (in the unknowns $x_1, x_2, \ldots, x_k$)

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k = \vec{b}$$

has a solution (i.e. is consistent).

Definition: Span

Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ be vectors in $\mathbb{R}^n$. The ***span*** of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ is the collection of all linear combinations of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$, and is denoted $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$. In symbols:

$$\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\} = \{x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k \,:\, x_1, x_2, \ldots, x_k \text{ in } \mathbb{R}\}$$

We also say that $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is the subset ***spanned by*** or ***generated by*** the vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$.

Note 1.3.2.4: Set Builder Notation

The notation

$$\{x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k \,:\, x_1, x_2, \ldots, x_k \text{ in } \mathbb{R}\}$$

reads as: "the set of all things of the form $x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k$ such that $x_1, x_2, \ldots, x_k$ are in $\mathbb{R}$." The colon is "such that"; everything to the left of it is "the set of all things of this form", and everything to the right is the condition that those things must satisfy to be in the set. Specifying a set in this way is called ***set builder notation***.

All mathematical notation is only shorthand: any sequence of symbols must translate into a usual sentence.

Note 1.3.2.5: Three characterizations of consistency

Now we have three equivalent ways of making the same statement:

1. A vector $\vec{b}$ is in the span of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$.
2. The vector equation

   $$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k = \vec{b}$$

   has a solution.
3. The linear system with augmented matrix

   $$\left[\begin{array}{cccc|c} | & | & & | & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k & \vec{b} \\ | & | & & | & | \end{array}\right]$$

   is consistent.

Equivalent means that, for any given list of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$, $\vec{b}$, either all three statements are true, or all three statements are false.

Drawing a picture of $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is the same as drawing a picture of all linear combinations of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$.

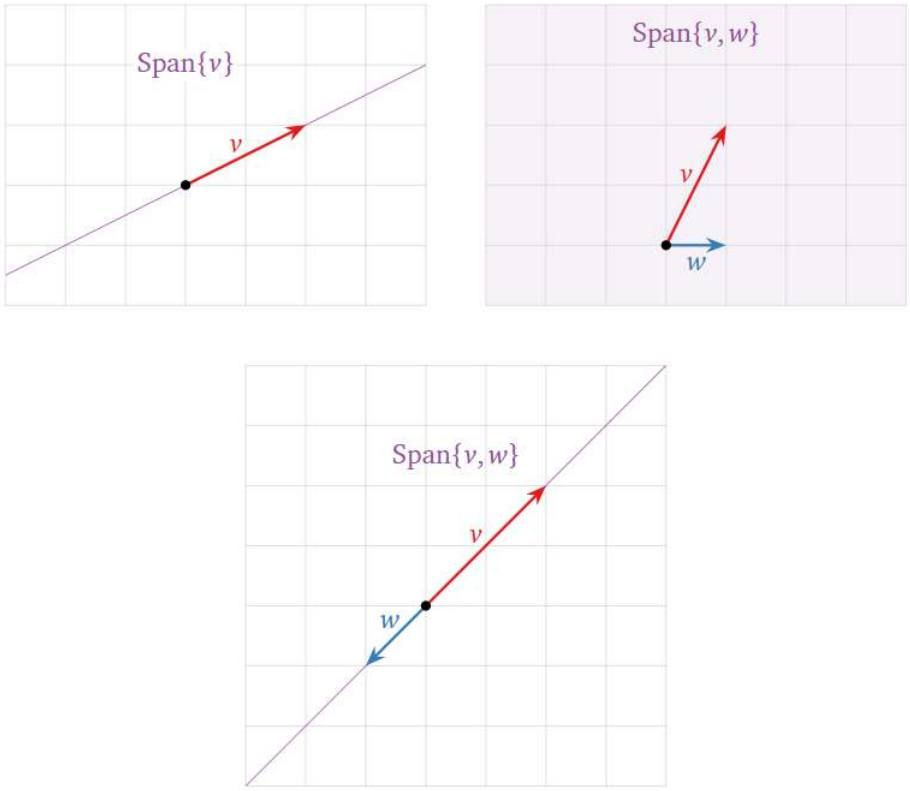


Figure 1.3.2.10: Pictures of spans in $\mathbb{R}^2$.

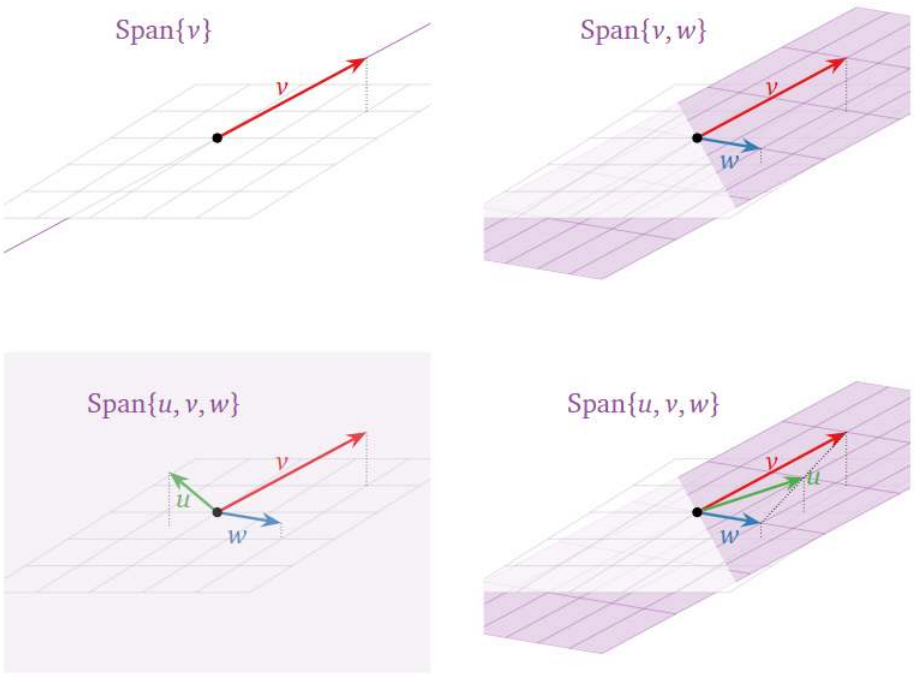


Figure 1.3.2.11: Pictures of spans in $\mathbb{R}^3$. The span of two noncollinear vectors is the plane containing the origin and the heads of the vectors. Note that three coplanar (but not collinear) vectors span a plane and not a 3-space, just as two collinear vectors span a line and not a plane.

1.3.2: Vector Equations and Spans is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.2: Vector Equations and Spans** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.3.3: Practice Set

Exercise 1.3.3.1

Let $\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ -1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 4 \end{bmatrix}$, and $\vec{c} = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 3 \end{bmatrix}$

(a) Find $10\vec{a}$

(b) Find $2\vec{a} - 3\vec{c}$

(c) Find $\vec{a} - \vec{b} + 4\vec{c}$

(d) Can $\vec{0}$ be written as a linear combination of $\vec{a}, \vec{b}, \vec{c}$?

(e) Can $\vec{0}$ be written as a non-trivial linear combination of $\vec{a}, \vec{b}, \vec{c}$?

(f) Is $\vec{d} = \begin{bmatrix} 3 \\ 3 \\ -3 \\ 12 \end{bmatrix} \in \text{Span}\{\vec{a}, \vec{b}, \vec{c}\}$?

(g) Find a vector which is *not* in $\text{Span}\{\vec{a}, \vec{b}, \vec{c}\}$

Answer

(a) $10\vec{a} = \begin{bmatrix} 10 \\ 20 \\ 10 \\ -10 \end{bmatrix}$

(b) $2\vec{a} - 3\vec{c} = \begin{bmatrix} -4 \\ -5 \\ -1 \\ -11 \end{bmatrix}$

(c) $\vec{a} - \vec{b} + 4\vec{c} = \begin{bmatrix} 8 \\ 13 \\ 6 \\ 7 \end{bmatrix}$

(d) Yes, $\vec{0} = 0\vec{a} + 0\vec{b} + 0\vec{c}$

(e) To find out we solve $\left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 \\ 2 & 1 & 3 & 0 \\ 1 & -1 & 1 & 0 \\ -1 & 4 & 3 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

Since there is only the trivial solution, the answer is NO.

(f) To find out we solve $\left[\begin{array}{cccc|c} 1 & 1 & 2 & 3 \\ 2 & 1 & 3 & 3 \\ 1 & -1 & 1 & -3 \\ -1 & 4 & 3 & 12 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

Since this is consistent, the answer is YES.

(g) To find one, we solve $\left[\begin{array}{cccc|c} 1 & 1 & 2 & \alpha \\ 2 & 1 & 3 & \beta \\ 1 & -1 & 1 & \gamma \\ -1 & 4 & 3 & \delta \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cccc|c} 1 & 0 & 0 & -\gamma-4\alpha+3\beta \\ 0 & 1 & 0 & -\gamma-\alpha+\beta \\ 0 & 0 & 1 & \gamma+3\alpha-2\beta \\ 0 & 0 & 0 & \delta-9\alpha+5\beta \end{array}\right]$

Any vector of the form $\begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}$ which satisfies $\delta-9\alpha+5\beta \neq 0$ makes this system inconsistent, and will not be in the span.

One such choice is $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

? Exercise 1.3.3.2

Let $\vec{p}=\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\vec{q}=\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$, and $\vec{r}=\begin{bmatrix} 0 \\ 2 \\ -2 \end{bmatrix}$

(a) Find $-\frac{1}{2}\vec{r}$

(b) Find $11\vec{q}+5\vec{p}$

(c) Find $\vec{p}+2\vec{q}-\vec{r}$

(d) Can $\vec{0}$ be written as a linear combination of $\vec{p},\vec{q},\vec{r}$?

(e) Can $\vec{0}$ be written as a non-trivial linear combination of $\vec{p},\vec{q},\vec{r}$?

(f) Is $\vec{t}=\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \in \text{Span}\{\vec{p},\vec{q},\vec{r}\}$?

(g) Find a vector which is *not* in $\text{Span}\{\vec{p},\vec{q},\vec{r}\}$

Answer

(a) $-\frac{1}{2}\vec{r}=\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

(b) $11\vec{q}+5\vec{p}=\begin{bmatrix} 16 \\ 22 \\ -6 \end{bmatrix}$

(c) $\vec{p}+2\vec{q}-\vec{r}=\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

(d) Yes, $\vec{0}=0\vec{p}+0\vec{q}+0\vec{r}$

(e) To find out we solve $\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 1 & -1 & -2 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$

Since there are free variables, the answer is YES. One such combination is $\vec{p}-\vec{q}+\vec{r}=\vec{0}$

(f) To find out we solve $\left[\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 2 & 2 & 1 \\ 1 & -1 & -2 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & -1 & \frac{1}{2} \\ 0 & 1 & 1 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$

Since this is consistent, the answer is YES.

(g) To find one, we solve $\left[\begin{array}{ccc|c} 1 & 1 & 0 & a \\ 0 & 2 & 2 & b \\ 1 & -1 & -2 & c \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & -1 & a-\frac{1}{2}b \\ 0 & 1 & 1 & \frac{1}{2}b \\ 0 & 0 & 0 & c-a+b \end{array}\right]$

Any vector of the form $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ which satisfies $c-a+b \neq 0$ makes this system inconsistent, and will not be in the span. One such choice is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

Exercise 1.3.3.3

Let $\vec{u} = \begin{bmatrix} 7 \\ 5 \\ -2 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} 5 \\ 5 \\ 5 \end{bmatrix}$

(a) Find $0\vec{u}$

(b) Find $-\frac{1}{2}\vec{v} + \frac{3}{10}\vec{w}$

(c) Find $\vec{v} - 3\vec{w} + 2\vec{u}$

(d) Can $\vec{0}$ be written as a linear combination of $\vec{u}, \vec{v}, \vec{w}$?

(e) Can $\vec{0}$ be written as a non-trivial linear combination of $\vec{u}, \vec{v}, \vec{w}$?

(f) Is $\vec{y} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \in \text{Span}\{\vec{u}, \vec{v}, \vec{w}\}$?

(g) Find a vector which is *not* in $\text{Span}\{\vec{u}, \vec{v}, \vec{w}\}$

Answer

(a) $0\vec{u} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

(b) $-\frac{1}{2}\vec{v} + \frac{3}{10}\vec{w} = \begin{bmatrix} 1 \\ \frac{1}{2} \\ 1 \end{bmatrix}$

(c) $\vec{v} - 3\vec{w} + 2\vec{u} = \begin{bmatrix} 0 \\ -3 \\ -18 \end{bmatrix}$

(d) Yes, $\vec{0} = 0\vec{u} + 0\vec{v} + 0\vec{w}$

(e) To find out we solve $\left[\begin{array}{ccc|c} 7 & 1 & 5 & 0 \\ 5 & 2 & 5 & 0 \\ -2 & 1 & 5 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$

Since there is only the trivial solution, the answer is NO.

(f) To find out we solve $\left[\begin{array}{ccc|c} 7 & 1 & 5 & 0 \\ 5 & 2 & 5 & 1 \\ -2 & 1 & 5 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -\frac{1}{63} \\ 0 & 1 & 0 & \frac{7}{9} \\ 0 & 0 & 1 & -\frac{9}{35} \end{array}\right]$

Since this is consistent, the answer is YES.

(we knew this without doing this calculation, since we saw the RREF in part (e) had a pivot in every row)

(g) To find one, we solve $\left[\begin{array}{ccc|c} 7 & 1 & 5 & a \\ 5 & 2 & 5 & b \\ -2 & 1 & 5 & c \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{9}a - \frac{1}{63}b - \frac{1}{63}c \\ 0 & 1 & 0 & \frac{7}{9}b - \frac{5}{9}a \\ 0 & 0 & 1 & \frac{2}{35}a - \frac{9}{35}b + \frac{1}{5}c \end{array}\right]$

This is always consistent, regardless of the choice of a, b, c. It is IMPOSSIBLE to pick a vector not in the span.

(we knew this without doing this calculation, since we saw the RREF in part (e) had a pivot in every row)

Exercise 1.3.3.4

Let $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ and $\vec{z} = \begin{bmatrix} 1 \\ h \\ 0 \end{bmatrix}$

(a) For what value(s) of h is $\vec{z} \in \text{Span}\{\vec{x}, \vec{y}\}$?

(b) For what value(s) of h is $\vec{0}$ a nontrivial combination of $\vec{x}, \vec{y}$, and $\vec{z}$?

(c) For what value(s) of h is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \in \text{Span}\{\vec{x}, \vec{y}, \vec{z}\}$?

(d) For what value(s) of h is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \notin \text{Span}\{\vec{x}, \vec{y}, \vec{z}\}$?

(e) For what value(s) of h is $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \in \text{Span}\{\vec{x}, \vec{y}, \vec{z}\}$?

(f) For what value(s) of h is $\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} \notin \text{Span}\{\vec{x}, \vec{y}, \vec{z}\}$?

Answer

(a) To find out, we solve $\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 1 & h & 0 \\ 3 & 0 & 0 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} 1 & 0 & -\frac{1}{3} + \frac{2}{3}h \\ 0 & 1 & \frac{2}{3} - \frac{1}{3}h \\ 0 & 0 & 1 - 2h \end{array}\right]$

This will be consistent only when $1 - 2h = 0$, so $\vec{z}$ is in this span when $h = \frac{1}{2}$.

(b) To find out, we solve $\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 2 & 1 & h & 0 \\ 3 & 0 & 0 & 0 \end{array}\right] \xrightarrow{\text{echelon form}} \left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{3} + \frac{2}{3}h & 0 \\ 0 & 1 & \frac{2}{3} - \frac{1}{3}h & 0 \\ 0 & 0 & 1 - 2h & 0 \end{array}\right]$

This will have a free variable only when $1 - 2h = 0$, so $\vec{0}$ is a nontrivial combination when $h = \frac{1}{2}$.

(c) To find out, we solve $\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 1 & h & 1 \\ 3 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\text{echelon form}} \left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{3} + \frac{2}{3}h & \frac{1}{3} \\ 0 & 1 & \frac{2}{3} - \frac{1}{3}h & \frac{1}{3} \\ 0 & 0 & 1 - 2h & 0 \end{array}\right]$

This is always consistent. So this vector is ALWAYS in the span, regardless of the value of h.

(d) Since we saw in part (c) that the corresponding system is always consistent. It is IMPOSSIBLE for this vector not to be in the span.

(e) To find out, we solve $\left[\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 1 & h & 1 \\ 3 & 0 & 0 & 2 \end{array}\right] \xrightarrow{\text{echelon form}} \left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{3} + \frac{2}{3}h & \frac{1}{3} \\ 0 & 1 & \frac{2}{3} - \frac{1}{3}h & \frac{1}{3} \\ 0 & 0 & 1 - 2h & 1 \end{array}\right]$

When $1 - 2h \neq 0$ this is consistent. So this vector is in the span when $h \neq \frac{1}{2}$.

(f) This vector will not be in the span when the system solved in part (e) is inconsistent, so when $h = \frac{1}{2}$.

Exercise 1.3.3.6

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If $\vec{u} \in \text{Span}\{\vec{v}\}$, then $\vec{v} \in \text{Span}\{\vec{u}\}$

(b) If $\vec{u} \in \text{Span}\{\vec{v}, \vec{w}\}$, then $\vec{u} \in \text{Span}\{\vec{v}\}$ or $\vec{u} \in \text{Span}\{\vec{w}\}$

(c) The span of one vector contains infinitely many different vectors.

(d) The span of two non-zero vectors contains infinitely many different vectors.

(e) The span of two non-zero vectors from $\mathbb{R}^2$ contains all possible vectors in $\mathbb{R}^2$.

(f) If there is more than one way for $\vec{0}$ to be written as a linear combination of the vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in \mathbb{R}^3$, then there is more than one way to write $\vec{v}_1$ as a linear combination of $\vec{v}_2$ and $\vec{v}_3$.

(g) If a vector $\vec{y}$ can be written as a linear combination of $\vec{v}_1, \vec{v}_2, \vec{v}_3$ in at least two distinct ways, then the vector $\vec{0}$ can be written as a linear combination of $\vec{v}_1, \vec{v}_2, \vec{v}_3$ in infinitely many different ways.

Answer

(a) **Sometimes True Sometimes False**

When $\vec{u} = \vec{v}$ this is true, but for $\vec{u} = \vec{0}$ and $\vec{v} \neq \vec{0}$ it is false.

(b) **Sometimes True Sometimes False**

When $\vec{u} = \vec{v}$ this is true, but for $\vec{u} = \vec{v} + \vec{w}$ and $\vec{v}, \vec{w} \neq \vec{0}$ it is false.

(c) **Sometimes True Sometimes False**

When the vector is nonzero this statement is always true. However, when the vector is $\vec{0}$, it is false.

(d) **Always True**

There are infinitely many different multiples of any non-zero vector.

(e) **Sometimes True Sometimes False**

This statement is true when the vectors are not parallel, and false when the two vectors are parallel to each other.

(f) **Sometimes True Sometimes False**

This is true when $\vec{v}_1 = \vec{v}_2$, but can be false for example when $\vec{v}_2 = \vec{v}_3 = \vec{0}$ and $\vec{v}_1 \neq \vec{0}$.

(g) **Always True**

If $\vec{y} = a\vec{v}_1 + b\vec{v}_2 + c\vec{v}_3 = \alpha\vec{v}_1 + \beta\vec{v}_2 + \gamma\vec{v}_3$, then $\vec{0} = (a-\alpha)\vec{v}_1 + (b-\beta)\vec{v}_2 + (c-\gamma)\vec{v}_3$.

1.3.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.4: Matrix Equations

Objectives

1. Understand the equivalence between a system of linear equations, an augmented matrix, a vector equation, and a matrix equation.
2. Characterize the vectors $\vec{b}$ such that $A\vec{x} = \vec{b}$ is consistent, in terms of the span of the columns of A.
3. Characterize matrices A such that $A\vec{x} = \vec{b}$ is consistent for all vectors $\vec{b}$.
4. Understand the relationship between the solution set of $A\vec{x} = 0$ and the solution set of $A\vec{x} = \vec{b}$.
5. Understand the difference between the solution set and the column span.
6. *Recipe:* multiply a vector by a matrix (two ways), solve a matrix equation in parametric form.
7. *Picture:* the set of all vectors $\vec{b}$ such that $A\vec{x} = \vec{b}$ is consistent.
8. *Vocabulary words:* ***matrix equation, homogeneous, inhomogeneous***, ***trivial solution***

In this section we introduce a very concise way of writing a system of linear equations: $A\vec{x} = \vec{b}$. Here A is a matrix and $\vec{x}, \vec{b}$ are vectors (generally of different sizes), so first we must explain how to multiply a matrix by a vector. Once this is formally defined, we relate this notation to our previous work in sections 1.1 and 1.2, and get a new perspective on the shapes of solution sets for systems of linear equations.

This page titled 1.4: Matrix Equations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber.

- **2.3: Matrix Equations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **2.4: Solution Sets** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.4.1: Matrix Equations, Spans, and Consistency

Note 1.4.1.1

When we say "A is an $m \times n$ matrix," we mean that A has m rows and n columns.

Remark

In this book, we do *not* reserve the letters m and n for the numbers of rows and columns of a matrix. If we write "A is an $n \times m$ matrix", then n is the number of rows of A and m is the number of columns.

Definition: Product

Let A be an $m \times n$ matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$:

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$$

The ***product*** of A with a vector $\vec{x}$ in $\mathbb{R}^n$ is the linear combination

$$A\vec{x} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n$$

This is a vector in $\mathbb{R}^m$.

Illustration 1.4.1.1

$$\begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1\begin{bmatrix} 4 \\ 7 \end{bmatrix} + 2\begin{bmatrix} 5 \\ 8 \end{bmatrix} + 3\begin{bmatrix} 6 \\ 9 \end{bmatrix} = \begin{bmatrix} 32 \\ 50 \end{bmatrix}$$

In order for $A\vec{x}$ to make sense, the number of entries of $\vec{x}$ has to be the same as the number of columns of A: we are using the entries of $\vec{x}$ as the coefficients of the columns of A in a linear combination. The resulting vector has the same number of entries as the number of *rows* of A, since each column of A has that number of entries.

Note 1.4.1.2

If A is an $m \times n$ matrix (m rows, n columns), then $A\vec{x}$ makes sense when $\vec{x}$ has n entries. The product $A\vec{x}$ has m entries.

Proposition 1.4.1.1: Properties of the Matrix-Vector Product

Let A be an $m \times n$ matrix, let $\vec{u}, \vec{v}$ be vectors in $\mathbb{R}^n$, and let c be a scalar. Then:

- $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$
- $A(c\vec{u}) = cA\vec{u}$

Definition: Matrix Equation

A ***matrix equation*** is an equation of the form $A\vec{x} = \vec{b}$, where A is an $m \times n$ matrix, $\vec{b}$ is a vector in $\mathbb{R}^m$, and $\vec{x}$ is a vector whose coefficients $x_1, x_2, \ldots, x_n$ are unknown.

In this book we will study two complementary questions about a matrix equation $A\vec{x} = \vec{b}$:

1. Given a specific choice of $\vec{b}$, what are all of the solutions to $A\vec{x} = \vec{b}$?
2. What are all of the choices of $\vec{b}$ so that $A\vec{x} = \vec{b}$ is consistent?

The first question is more like the questions you might be used to from your earlier courses in algebra; you have a lot of practice solving equations like $x^2 - 1 = 0$ for x. The second question is perhaps a new concept for you. Soon we will see that the two questions are intimately related.

Note 1.4.1.3: Matrix Equations and Vector Equations

Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ and $\vec{b}$ be vectors in $\mathbb{R}^m$. Consider the vector equation

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n = \vec{b}$$

This is equivalent to the matrix equation $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \quad \text{and} \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Conversely, if A is any $m \times n$ matrix, then $A\vec{x} = \vec{b}$ is equivalent to the vector equation

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n = \vec{b}$$

where $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ are the columns of A, and $x_1, x_2, \ldots, x_n$ are the entries of $\vec{x}$.

Example 1.4.1.1

Write the vector equation

$$2\vec{v}_1 + 3\vec{v}_2 - 4\vec{v}_3 = \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}$$

as a matrix equation, where $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are vectors in $\mathbb{R}^3$.

Solution

Let A be the matrix with columns $\vec{v}_1, \vec{v}_2, \vec{v}_3$, and let $\vec{x}$ be the vector with entries $2, 3, -4$. Then

$$A\vec{x} = \begin{bmatrix} | & | & | \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ -4 \end{bmatrix} = 2\vec{v}_1 + 3\vec{v}_2 - 4\vec{v}_3$$

So the vector equation is equivalent to the matrix equation $\boxed{A\vec{x} = \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}}$

Note 1.4.1.4: Four Ways of Writing a Linear System

We now have *four* equivalent ways of writing (and thinking about) a system of linear equations:

1. As a system of equations:

$$\begin{cases} 2x_1 + 3x_2 - 2x_3 = 7 \\ x_1 - x_2 - 3x_3 = 5 \end{cases}$$

2. As an augmented matrix:

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

3. As a vector equation ($x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n = \vec{b}$):

$$x_1\begin{bmatrix} 2 \\ 1 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ -1 \end{bmatrix} + x_3\begin{bmatrix} -2 \\ -3 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$$

4. As a matrix equation ($A\vec{x} = \vec{b}$):

$$\begin{bmatrix} 2 & 3 & -2 \\ 1 & -1 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \end{bmatrix}$$

In particular, *all four have the same solution set*.

Note 1.4.1.5

We will move back and forth freely between the four ways of writing a linear system, over and over again, for the rest of the book.

The above definition of Matrix-Vector Product is a useful way of defining the product of a matrix with a vector when it comes to understanding the relationship between matrix equations and vector equations. Here we give a definition that is better-adapted to computations by hand.

Definition: Row Vector

A ***row vector*** is a matrix with one row. The ***product*** of a row vector of length n and a (column) vector of length n is

$$\begin{bmatrix} a_1 & a_2 & \cdots a_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

This is a scalar in $\mathbb{R}$.

This type of product can be done between any two vectors with the same number of entries, and is written as $\vec{a} \cdot \vec{x}$

Recipe: The Row-Column Rule for Matrix-Vector Multiplication

If A is an $m \times n$ matrix with rows $\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m$, and $\vec{x}$ is a vector in $\mathbb{R}^n$, then

$$A\vec{x} = \begin{bmatrix} \text{—}\vec{r}_1\text{—} \\ \text{—}\vec{r}_2\text{—} \\ \vdots \\ \text{—}\vec{r}_m\text{—} \end{bmatrix}\vec{x} = \begin{bmatrix} \vec{r}_1 \cdot \vec{x} \\ \vec{r}_2 \cdot \vec{x} \\ \vdots \\ \vec{r}_m \cdot \vec{x} \end{bmatrix}$$

Illustration 1.4.1.2

Using this notation, we see that:

$$\begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \\ \begin{bmatrix} 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \end{bmatrix}$$
$$= \begin{bmatrix} 4 \cdot 1 + 5 \cdot 2 + 6 \cdot 3 \\ 7 \cdot 1 + 8 \cdot 2 + 9 \cdot 3 \end{bmatrix} = \begin{bmatrix} 32 \\ 50 \end{bmatrix}$$

This is the same answer as before:

$$\begin{bmatrix} 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1 \begin{bmatrix} 4 \\ 7 \end{bmatrix} + 2 \begin{bmatrix} 5 \\ 8 \end{bmatrix} + 3 \begin{bmatrix} 6 \\ 9 \end{bmatrix}$$
$$= \begin{bmatrix} 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 \\ 1 \cdot 7 + 2 \cdot 8 + 3 \cdot 9 \end{bmatrix} = \begin{bmatrix} 32 \\ 50 \end{bmatrix}$$

Let A be a matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$:

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$$

Then

$A\vec{x} = \vec{b}$ has a solution

$\iff$ there exist $x_1, x_2, \ldots, x_n$ such that $A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \vec{b}$

$\iff$ there exist $x_1, x_2, \ldots, x_n$ such that $x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n = \vec{b}$

$\iff$ $\vec{b}$ is a linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$

$\iff$ $\vec{b}$ is in the span of the columns of A.

Note 1.4.1.6: Spans and Consistency

The matrix equation $A\vec{x} = \vec{b}$ has a solution if and only if $\vec{b}$ is in the span of the columns of A.

This gives an equivalence between an *algebraic* statement ($A\vec{x} = \vec{b}$ is consistent) and a *geometric* statement ($\vec{b}$ is in the span of the columns of A).

Example 1.4.1.2: An Inconsistent System

Let $A = \begin{bmatrix} 2 & -6 \\ -3 & 9 \end{bmatrix}$. Does the equation $A\vec{x} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$ have a solution?

Solution

First, we answer the question geometrically. The columns of A are

$$\vec{v}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} -6 \\ 9 \end{bmatrix}$$

and the target vector (on the right-hand side of the equation) is $\vec{w} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$. The equation $A\vec{x} = \vec{w}$ is consistent if and only if $\vec{w}$ is contained in the span of the columns of A. So we draw a picture:

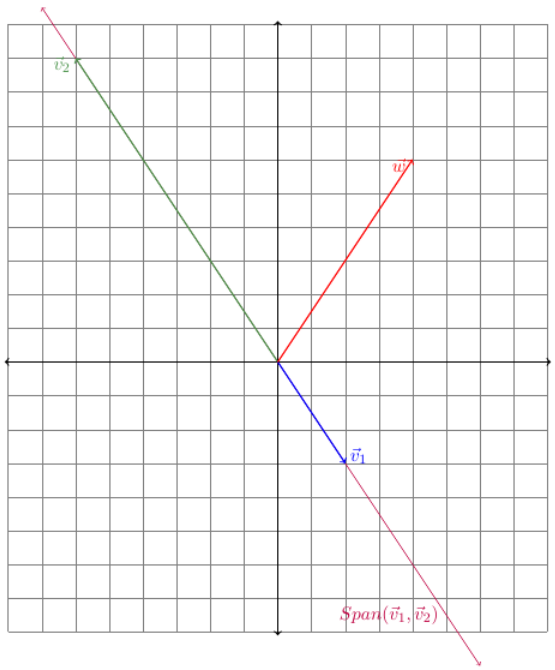


Figure 1.4.1.1

It does not appear that $\vec{w}$ lies in $\text{Span}\{\vec{v}_1, \vec{v}_2\}$, so the equation is inconsistent.

Let us check our geometric answer by solving the matrix equation using row-reduction. We put the system into an augmented matrix and row reduce:

$$\left[\begin{array}{cc|c} 2 & -6 & 4 \\ -3 & 9 & 6 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & -3 & 2 \\ 0 & 0 & 1 \end{array}\right]$$

The last equation is $0 = 1$, so the system is indeed inconsistent, and the matrix equation

$$\begin{bmatrix} 2 & -6 \\ -3 & 9 \end{bmatrix} \vec{x} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

has **no solution**

Example 1.4.1.3: A Consistent System

Let $A = \begin{bmatrix} 2 & -6 \\ -3 & 9 \end{bmatrix}$. Does the equation $A\vec{x} = \begin{bmatrix} 4 \\ -6 \end{bmatrix}$ have a solution?

Solution

First, we answer the question geometrically. The columns of A are

$$\vec{v}_1 = \begin{bmatrix} 2 \\ -3 \end{bmatrix} \quad \text{and} \quad v_2 = \begin{bmatrix} -6 \\ 9 \end{bmatrix}$$

and the target vector (on the right-hand side of the equation) is $\vec{w} = \begin{bmatrix} 4 \\ -6 \end{bmatrix}$. The equation $A\vec{x} = \vec{w}$ is consistent if and only if $\vec{w}$ is contained in the span of the columns of A. So we draw a picture:

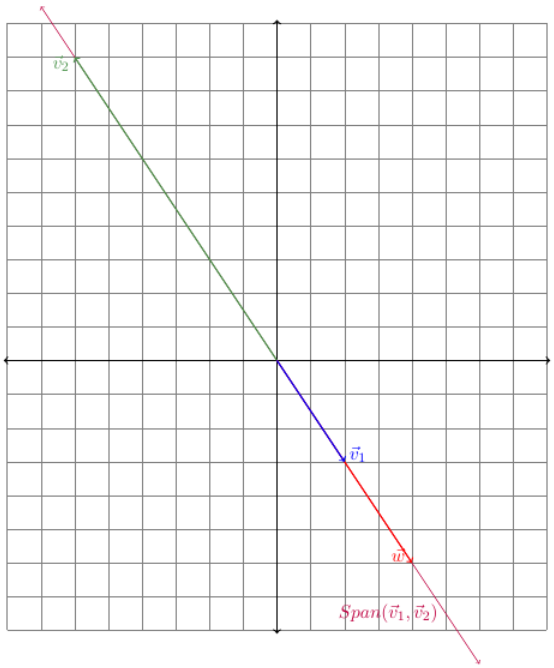


Figure 1.4.1.2

It appears that $\vec{w}$ is indeed contained in the span of the columns of A; in fact, we can see one possible solution:

$$\vec{w} = 2\vec{v}_1 + 0\vec{v}_2 \implies \vec{x} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

Let us check our geometric answer by solving the matrix equation using row-reduction. We put the system into an augmented matrix and row reduce:

$$\left[\begin{array}{cc|c} 2 & -6 & 4 \\ -3 & 9 & -6 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & -3 & 2 \\ 0 & 0 & 0 \end{array}\right]$$

This gives us $x_1 - 3x_2 = 2$ and x_2 a free variable. There are actually **infinitely many solutions**

This is consistent with the picture, and if we pick $x_2 = t$ and $x_1 = 2 + 3t$ we can see:

$$(2+3t)\begin{bmatrix} 2 \\ -3 \end{bmatrix} + t\begin{bmatrix} -6 \\ 9 \end{bmatrix} = \begin{bmatrix} 4 \\ -6 \end{bmatrix} \quad \text{or} \quad A\begin{bmatrix} 3t+2 \\ t \end{bmatrix} = \begin{bmatrix} 4 \\ -6 \end{bmatrix}$$

Building on the Note 1.4.1.6: Spans and Consistency, we have the following criterion for when $A\vec{x} = \vec{b}$ is consistent for *every* choice of $\vec{b}$.

Theorem 1.4.1.1

Let A be an $m \times n$ (non-augmented) matrix. The following are equivalent:

1. $A\vec{x} = \vec{b}$ has a solution for all $\vec{b}$ in $\mathbb{R}^m$.
2. The span of the columns of A is all of $\mathbb{R}^m$.
3. A has a pivot position in every row.

Proof

The equivalence of 1 and 2 is established by the Note 1.4.1.6: Spans and Consistency as applied to every $\vec{b}$ in $\mathbb{R}^m$.

Now we show that 1 and 3 are equivalent. (Since we know 1 and 2 are equivalent, this implies 2 and 3 are equivalent as well.) If A has a pivot in every row, then its reduced row echelon form looks like this:

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

and therefore, $\left[\begin{array}{c|c} A & \vec{b} \end{array}\right]$ reduces to this:

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

There is no $\vec{b}$ that makes it inconsistent, so there is always a solution. Conversely, if A does not have a pivot in each row, then its reduced row echelon form looks like this:

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

which can give rise to an inconsistent system after augmenting with $\vec{b}$:

$$\left[\begin{array}{ccccc|c} 1 & 0 & \star & 0 & \star & 0 \\ 0 & 1 & \star & 0 & \star & 0 \\ 0 & 0 & 0 & 0 & 0 & 16 \end{array}\right]$$

Recall that ***equivalent*** means that, for any given matrix A, either *all* of the conditions of the above theorem are true, or they are all false.

Note 1.4.1.7

Be careful when reading the statement of the above theorem. The first two conditions look very much like this Note **1.4.1.6**: Spans and Consistency, but they are logically quite different because of the quantifier "*for all* $\vec{b}$".

Illustration 1.4.1.3: The criteria of the theorem are not satisfied

We saw that the matrix $A = \begin{bmatrix} 2 & -6 \\ -3 & 9 \end{bmatrix}$ from our previous examples has RREF $\begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}$

Since this does not have a pivot in each row, we were able to find at least one vector $\vec{w} = \begin{bmatrix} 4 \\ -6 \end{bmatrix}$ which made $A\vec{x} = \vec{w}$ inconsistent.

This is why we saw geometrically that the span of the columns of this matrix formed a line instead of filling the entirety of $\mathbb{R}^2$.

Illustration 1.4.1.4: The criteria of the theorem are satisfied

Consider the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ which has RREF $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

This has a pivot in each row, so regardless of what $\vec{b}$ we choose in $\mathbb{R}^2$, the equation $A\vec{x} = \vec{b}$ will be consistent.

$$\left[\begin{array}{cc|c} 1 & 2 & b_1 \\ 3 & 5 & b_2 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & -5b_1 + 2b_2 \\ 0 & 1 & 3b_1 - b_2 \end{array}\right]$$

This means that $(-5b_1 + 2b_2)\begin{bmatrix} 1 \\ 3 \end{bmatrix} + (3b_1 - b_2)\begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$

Geometrically, we can see that since these two column vectors are not parallel, their span will cover all of $\mathbb{R}^2$.

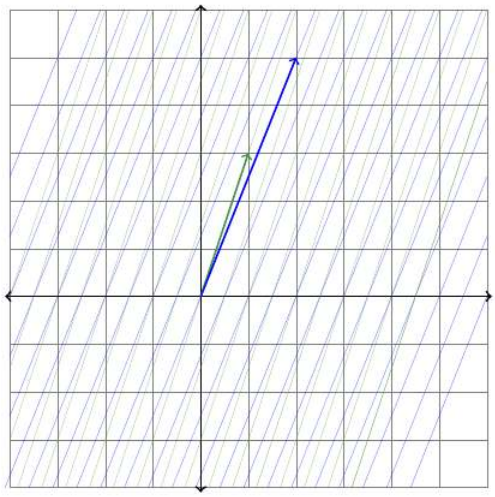


Figure 1.4.1.3

1.4.1: Matrix Equations, Spans, and Consistency is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.3: Matrix Equations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.4: Solutions to Matrix Equations

The equation $A\vec{x} = \vec{b}$ is easier to solve when $\vec{b} = \vec{0}$, so we start with this case.

Definition: Homogeneous & Inhomogeneous

A system of linear equations of the form $A\vec{x} = \vec{0}$ is called ***homogeneous***.

A system of linear equations of the form $A\vec{x} = \vec{b}$ for $\vec{b} \neq \vec{0}$ is called ***nonhomogeneous***.

A homogeneous system is just a system of linear equations where all constants on the right side of the equals sign are zero.

Note 1.4.8

A homogeneous system always has the solution $\vec{x} = \vec{0}$. This is called the ***trivial solution***. Any nonzero solution is called ***nontrivial***.

Observation 1.4.1

The equation $A\vec{x} = \vec{0}$ has a nontrivial solution $\iff$ there is a free variable $\iff$ A has a column without a pivot position.

Example 1.4.4: No nontrivial solutions

Find the solution set of $A\vec{x} = \vec{0}$, where

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

Solution

We form an augmented matrix and row reduce:

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

The only solution is the trivial solution $\boxed{\vec{x} = \vec{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}$

Observation 1.4.2

When we row reduce the augmented matrix for a homogeneous system of linear equations, the last column will be zero throughout the row-reduction process. We saw this in the last Example 1.4.8:

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

So it is not really necessary to write augmented matrices when solving homogeneous systems.

When the homogeneous equation $A\vec{x} = \vec{0}$ does have nontrivial solutions, it turns out that the solution set can be conveniently expressed as a span.

Illustration 1.4.5: Parametric Vector Form (homogeneous case)

Consider the following matrix in reduced row echelon form:

$$A = \begin{bmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The matrix equation $A\vec{x} = \vec{0}$ corresponds to the system of equations

$$\begin{cases} x_1 & - & 8x_3 & - & 7x_4 & = & 0 \\ x_2 & + & 4x_3 & + & 3x_4 & = & 0 \end{cases}$$

We can write the parametric form as follows:

$$\begin{cases} x_1 = 8x_3 + 7x_4 \\ x_2 = -4x_3 - 3x_4 \\ x_3 = x_3 \\ x_4 = x_4 \end{cases}$$

We wrote the redundant equations $x_3 = x_3$ and $x_4 = x_4$ in order to turn the above system into a *vector equation*:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 8 \\ -4 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 7 \\ -3 \\ 0 \\ 1 \end{bmatrix}$$

This vector equation is called the ***parametric vector form*** of the solution set. Since x_3 and x_4 are allowed to be anything, this says that the solution set is the set of all linear combinations of $\begin{bmatrix} 8 \\ -4 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 7 \\ -3 \\ 0 \\ 1 \end{bmatrix}$. In other words, the solution set is

$$\boxed{\text{Span}\left\{\begin{bmatrix} 8 \\ -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ -3 \\ 0 \\ 1 \end{bmatrix}\right\}}$$

Note that each of these vectors inside of our span represent a different linear combination of the columns of A:

$$A\begin{bmatrix} 8 \\ -4 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 8 \\ -4 \\ 1 \\ 0 \end{bmatrix}$$

$$= 8\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 4\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} -8 \\ 4 \\ 0 \end{bmatrix} + 0\begin{bmatrix} -7 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$A\begin{bmatrix} 7 \\ -3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 7 \\ -3 \\ 0 \\ 1 \end{bmatrix}$$

$$= 7\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 3\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} -8 \\ 4 \\ 0 \end{bmatrix} + 1\begin{bmatrix} -7 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Any solution to the homogeneous equation is a vector representing some way that the columns of A "cancel out".

Here is the general procedure.

Recipe: Parametric Vector Form (Homogeneous Case)

Let A be an $m \times n$ matrix. Suppose that the free variables in the homogeneous equation $A\vec{x} = \vec{0}$ are, for example, x_3, x_6, and x_8.

1. Find the reduced row echelon form of A.
2. Write the parametric form of the solution set, including the redundant equations $x_3 = x_3$, $x_6 = x_6$, $x_8 = x_8$. Put equations for all of the x_i in order.
3. Make a single vector equation from these equations by making the coefficients of x_3, x_6, and x_8 into vectors v_3, v_6, and v_8, respectively.

The solutions to $A\vec{x} = \vec{0}$ will then be expressed in the form

$$\vec{x} = x_3\vec{v}_3 + x_6\vec{v}_6 + x_8\vec{v}_8$$

for some vectors $\vec{v}_3, \vec{v}_6, \vec{v}_8$ in $\mathbb{R}^n$, and any scalars x_3, x_6, x_8. This is called the ***parametric vector form*** of the solution.

In this case, the solution set can be written as $\text{Span}\{\vec{v}_3,\ \vec{v}_6,\ \vec{v}_8\}$.

We emphasize the following fact in particular.

Note 1.4.9

The set of solutions to a homogeneous equation $A\vec{x} = \vec{0}$ is a span.

Example 1.4.5: The solution set is a line

Compute the parametric vector form of the solution set of $A\vec{x} = \vec{0}$, where

$$A = \begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix}$$

Solution

We row reduce (without augmenting), as suggested in the above observation:

$$\begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}$$

This corresponds to the single equation $x_1 - 3x_2 = 0$. We write the parametric form including the redundant equation $x_2 = x_2$:

$$\begin{cases} x_1 &= 3x_2 \\ x_2 &= x_2 \end{cases}$$

We turn these into a single vector equation:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

This is the parametric vector form of the solution set. Since x_2 is allowed to be anything, this says that the solution set is the set of all scalar multiples of $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$, otherwise known as

$$\boxed{\text{Span}\left\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right\}}$$

We know how to draw the picture of a span of a vector: it is a line. Therefore, this is a picture of the solution set:

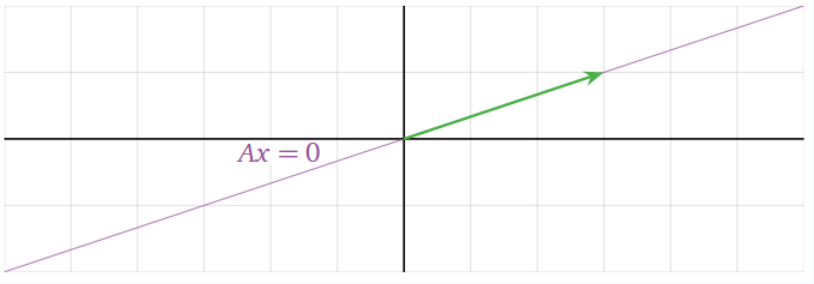


Figure 1.4.4

For any solution $\vec{x} = \begin{bmatrix} 3t \\ t \end{bmatrix}$, this gives:

$$A\vec{x} = \begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix} \begin{bmatrix} 3t \\ t \end{bmatrix} = 3t \begin{bmatrix} 1 \\ 2 \end{bmatrix} + t \begin{bmatrix} -3 \\ -6 \end{bmatrix} = t \left(3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} -3 \\ -6 \end{bmatrix} \right) = t \left(\vec{0} \right) = \vec{0}$$

Note 1.4.10

In order to actually *find* a nontrivial solution to $A\vec{x} = \vec{0}$ in the above Example 1.4.5, it suffices to substitute any nonzero value for the free variable x_2. For instance, taking $x_2 = 10$ gives the nontrivial solution $\vec{x} = 10 \cdot \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 30 \\ 10 \end{bmatrix}$.

Example 1.4.6: The solution set is a plane

Compute the parametric vector form of the solution set of $A\vec{x} = \vec{0}$, where

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{bmatrix}$$

Solution

We row reduce (without augmenting, as suggested in the above Observation 1.4.2):

$$\begin{bmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

This corresponds to the single equation $x_1 - x_2 + 2x_3 = 0$. We write the parametric form including the redundant equations $x_2 = x_2$ and $x_3 = x_3$:

$$\begin{cases} x_1 &= x_2 &- 2x_3 \\ x_2 &= x_2 & \\ x_3 &= & x_3 \end{cases}$$

We turn these into a single vector equation:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

This is the parametric vector form of the solution set. Since x_2 and x_3 are allowed to be anything, this says that the solution set is the set of all linear combinations of $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$. In other words, the solution set is

$$\boxed{\text{Span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}}$$

We know how to draw the span of two noncollinear vectors in $\mathbb{R}^3$: it is a plane. Therefore, this is a picture of the solution set:

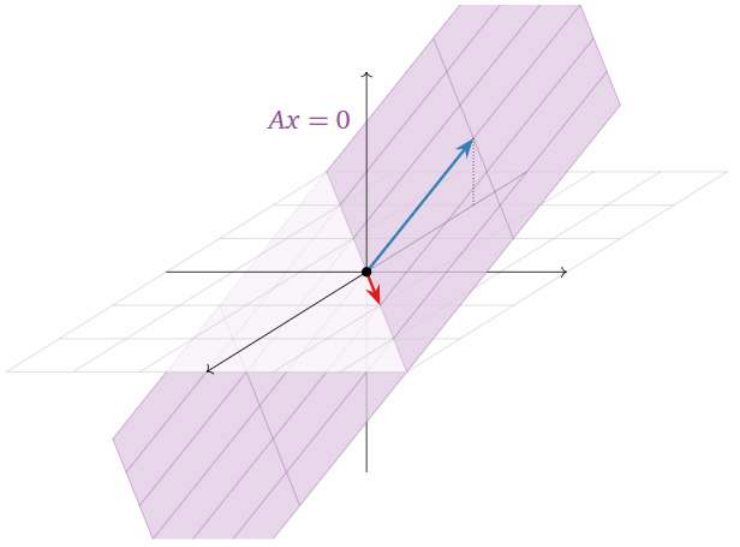


Figure 1.4.5

A natural question is: are the solution sets for nonhomogeneous equations also spans? As we will see shortly, they are never spans, but they are closely related to spans.

There is a natural relationship between the number of free variables and the "size" of the solution set, as follows.

Note 1.4.11: Dimension of the Solution Set

The above examples show us the following pattern: when there is one free variable in a consistent matrix equation, the solution set is a line, and when there are two free variables, the solution set is a plane, etc. The number of free variables is called the *dimension* of the solution set.

We will develop a rigorous definition of dimension later, but for now the dimension will simply mean the number of free variables.

Intuitively, the dimension of a solution set is the number of parameters you need to describe a point in the solution set. For a line only one parameter is needed, and for a plane two parameters are needed. This is similar to how the location of a building on Peachtree Street—which is like a line—is determined by one number and how a street corner in Manhattan—which is like a plane—is specified by two numbers.

Recall that a matrix equation $A\vec{x} = \vec{b}$ is called ***nonhomogeneous*** when $\vec{b} \neq \vec{0}$.

Example 1.4.7: The solution set is a line

What is the solution set of $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} -3 \\ -6 \end{bmatrix}$$

(Compare to this Example 1.4.5, where we solved the corresponding homogeneous equation.)

Solution

We row reduce the associated augmented matrix:

$$\left[\begin{array}{cc|c} 1 & -3 & -3 \\ 2 & -6 & -6 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & -3 & -3 \\ 0 & 0 & 0 \end{array}\right]$$

This corresponds to the single equation $x_1 - 3x_2 = -3$. We can write the parametric form as follows:

$$\begin{cases} x_1 &= 3x_2 - 3 \\ x_2 &= x_2 + 0 \end{cases}$$

We turn the above system into a *vector equation*:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \begin{bmatrix} -3 \\ 0 \end{bmatrix}$$

This vector equation is called the ***parametric vector form*** of the solution set. We write the solution set as

$$\boxed{\text{Span}\left\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right\} + \begin{bmatrix} -3 \\ 0 \end{bmatrix}}$$

Here is a picture of the solution set:

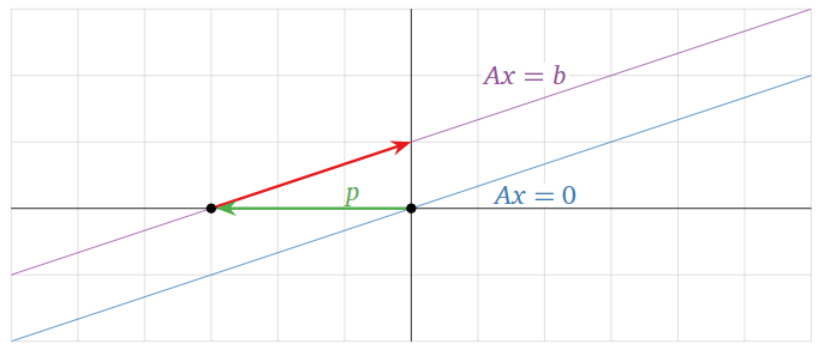


Figure 1.4.6

Why does the solution have this shape? Let's multiply each of these vectors by A to see:

$$A\vec{p} = \begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix} \begin{bmatrix} -3 \\ 0 \end{bmatrix} = -3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 0 \begin{bmatrix} -3 \\ -6 \end{bmatrix} = \begin{bmatrix} -3 \\ -6 \end{bmatrix}$$

Notice that this vector gives you one particular linear combination of the columns of A that add up to $\vec{b}$. We call $\vec{p}$ a ***particular solution*** of $A\vec{x} = \vec{b}$

$$A\vec{h} = \begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1 \begin{bmatrix} -3 \\ -6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Since $\vec{h}$ satisfies $A\vec{h} = \vec{0}$, we see that $A(\vec{p} + t\vec{h}) = A\vec{p} + tA\vec{h} = \vec{b} + t\vec{0} = \vec{b}$

In a consistent system $A\vec{x} = \vec{b}$, the solutions will always be of this form: $\vec{x} = \vec{p} + \vec{h}$ where $A\vec{p} = \vec{b}$ and $A\vec{h} = \vec{0}$.

Example 1.4.8: The solution set is a plane

What is the solution set of $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

(Compare this Example 1.4.6.)

Solution

We row reduce the associated augmented matrix:

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

This corresponds to the single equation $x_1 - x_2 + 2x_3 = 1$. We can write the parametric form as follows:

$$\begin{cases} x_1 &= x_2 - 2x_3 + 1 \\ x_2 &= x_2 \quad\quad + 0 \\ x_3 &= \quad\quad x_3 + 0 \end{cases}$$

We turn the above system into a *vector equation*:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

This vector equation is called the ***parametric vector form*** of the solution set. Since x_2 and x_3 are allowed to be anything, this says that the solution set is the set of all linear combinations of $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$, *translated* by the particular solution $\vec{p} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. This is a plane which contains $\vec{p}$ and is parallel to $\text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}\right\}$: it is a *translate of a plane*. We write the solution set as

$$\boxed{\text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}\right\} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}$$

Here is a picture of the solution set:

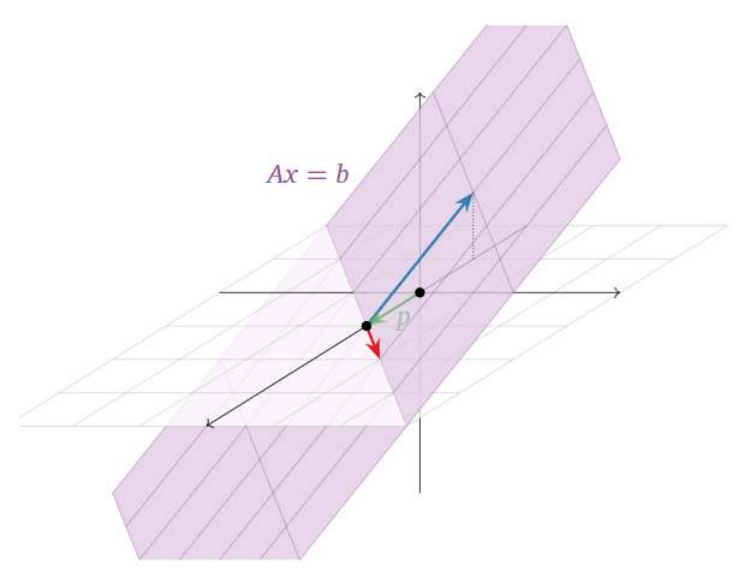


Figure 1.4.7

Key Observation 1.4.3

If $A\vec{x} = \vec{b}$ is consistent, the set of solutions is obtained by taking one ***particular solution*** $\vec{p}$ of $A\vec{x} = \vec{b}$, and adding all solutions of $A\vec{x} = \vec{0}$.

In particular, if $A\vec{x} = \vec{b}$ is consistent, the solution set is a ***translate of a span***.

The ***parametric vector form*** of the solutions of $A\vec{x} = \vec{b}$ is just the parametric vector form of the solutions of $A\vec{x} = \vec{0}$, plus a particular solution $\vec{p}$.

Remark

Row reducing to find the parametric vector form will give you one particular solution $\vec{p}$ of $A\vec{x} = \vec{b}$. But the Key Observation 1.4.3 is true for any solution $\vec{p}$. In other words, if we row reduce in a different way and find a different solution $\vec{q}$ to $A\vec{x} = \vec{b}$ then the solutions to $A\vec{x} = \vec{b}$ can be obtained from the solutions to $A\vec{x} = 0$ by either adding $\vec{p}$ or by adding $\vec{q}$.

Example 1.4.9: The solution set is a point

What is the solution set of $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 1 & 3 & 4 \\ 2 & -1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Solution

We form an augmented matrix and row reduce:

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

The only solution is $\vec{p} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$.

According to the Key Observation 1.4.3, this is supposed to be a translate of a span by $\vec{p}$. Indeed, we saw in the first Example 1.4.4 that the only solution of $A\vec{x} = \vec{0}$ is the trivial solution, i.e., that the solution set is the one-point set $\{\vec{0}\}$. The solution set of the inhomogeneous equation $A\vec{x} = \vec{b}$ is

$$\boxed{\{\vec{0}\} + \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}}$$

Note that $\{\vec{0}\} = \text{Span}\{\vec{0}\}$, so the homogeneous solution set is a span.

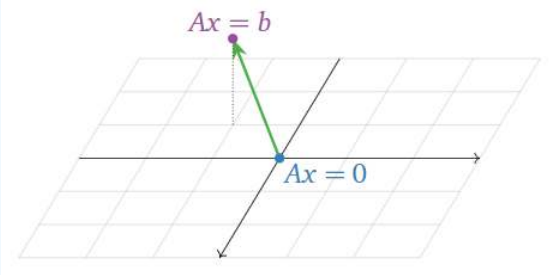


Figure 1.4.8

Note 1.4.11: Dimension of the Solution Set

As in this important note, when there is one free variable in a consistent matrix equation, the solution set is a line—this line does not pass through the origin when the system is inhomogeneous—when there are two free variables, the solution set is a plane (again not through the origin when the system is inhomogeneous), etc.

Note 1.4.12

To every $m \times n$ matrix A, we have now associated two completely different geometric objects, both described using spans.

- The ***solution set***: for fixed $\vec{b}$, this is the set of all $\vec{x}$ such that $A\vec{x} = \vec{b}$.

- This is a span if $\vec{b} = \vec{0}$, and it is a translate of a span if $\vec{b} \neq \vec{0}$ (and $A\vec{x} = \vec{b}$ is consistent).
- It is a subset of $\mathbb{R}^n$.
- It is computed by solving a system of equations: usually by row reducing and finding the parametric vector form.

- The ***span of the columns*** *of* A: this is the set of all $\vec{b}$ such that $A\vec{x} = \vec{b}$ is consistent.
  - This is always a span.
  - It is a subset of $\mathbb{R}^m$.
  - It is not computed by solving a system of equations: row-reduction plays no role.

Do not confuse these two geometric constructions! In the first the question is which $\vec{x}$'s work for a given $\vec{b}$ and in the second the question is which $\vec{b}$'s work for some $\vec{x}$.

1.4: Solutions to Matrix Equations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.4: Solution Sets** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.4: Practice Set

Exercise 1.4.1a

Find the solution set to $\begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 3 & 0 & 2 \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$ in parametric form (if it exists).

Answer

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

Since this system is inconsistent, the matrix equation has no solution.

Exercise 1.4.1b

Find the solution set to $\begin{bmatrix} 1 & -3 & 2 \\ 2 & 1 & -1 \\ 5 & -1 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix}$ in parametric form (if it exists).

Answer

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

This represents $\begin{cases} x_1 - \frac{1}{7}x_3 = \frac{6}{7} \\ x_2 - \frac{5}{7}x_3 = \frac{2}{7} \end{cases} \implies \begin{cases} x_1 = \frac{6}{7} + \frac{1}{7}x_3 \\ x_2 = \frac{2}{7} + \frac{5}{7}x_3 \\ x_3 \text{ free} \end{cases}$

So in parametric vector form we have $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{6}{7} \\ \frac{2}{7} \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \frac{1}{7} \\ \frac{5}{7} \\ 1 \end{bmatrix}$

Exercise 1.4.1c

Find the solution set to $\begin{bmatrix} 9 & 18 & -3 \\ 11 & 12 & 1 \\ 1 & -3 & -2 \end{bmatrix} \vec{x} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$ in parametric form (if it exists).

Answer

$$\left[\begin{array}{ccc|c} 9 & 18 & -3 & 3 \\ 11 & 12 & 1 & -1 \\ 1 & -3 & -2 & 2 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

This represents $\begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = -1 \end{cases}$

So in parametric vector form we have $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$

Exercise 1.4.1d

Find the solution set to $\begin{bmatrix} 1 & 4 & 0 & 2 \\ 2 & 1 & -7 & 3 \\ 2 & -1 & -9 & -2 \\ 2 & -1 & -9 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 6 \\ 4 \\ -3 \\ -1 \end{bmatrix}$ in parametric form (if it exists).

Answer

$$\left[\begin{array}{cccc|c} 1 & 4 & 0 & 2 & 6 \\ 2 & 1 & -7 & 3 & 4 \\ 2 & -1 & -9 & -2 & -3 \\ 2 & -1 & -9 & 0 & -1 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cccc|c} 1 & 0 & -4 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

This represents $\begin{cases} x_1 - 4x_3 = 0 \\ x_2 + x_3 = 1 \\ x_4 = 1 \end{cases} \implies \begin{cases} x_1 = 4x_3 \\ x_2 = 1 - x_3 \\ x_3 \text{ free} \\ x_4 = 1 \end{cases}$

So in parametric vector form we have $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ -1 \\ 1 \\ 0 \end{bmatrix}$

Exercise 1.4.1e

Find the solution set to $\begin{bmatrix} 0 & 1 & 4 & 1 \\ 1 & 3 & 5 & 7 \\ 0 & 2 & 8 & 2 \\ 1 & 2 & 1 & 6 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ -2 \end{bmatrix}$ in parametric form (if it exists).

Answer

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

Since this system is inconsistent, the matrix equation has no solution.

Exercise 1.4.1f

Find the solution set to $\begin{bmatrix} 0 & 1 & 4 & 1 & 0 \\ 1 & 3 & 5 & 7 & 0 \\ 0 & 2 & 8 & 2 & -1 \\ 1 & 2 & 1 & 6 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ -2 \end{bmatrix}$ in parametric form (if it exists).

Answer

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

This represents $\begin{cases} x_1 - 7x_3 + 4x_4 = -6 \\ x_2 + 4x_3 + x_4 = 2 \\ x_5 = 3 \end{cases} \implies \begin{cases} x_1 = -6 + 7x_3 - 4x_4 \\ x_2 = 2 - 4x_3 - x_4 \\ x_3 \text{ free} \\ x_4 \text{ free} \\ x_5 = 3 \end{cases}$

So in parametric vector form we have $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -6 \\ 2 \\ 0 \\ 0 \\ 3 \end{bmatrix} + x_3 \begin{bmatrix} 7 \\ -4 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -4 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$

? Exercise 1.4.2a

Use inspection to find all solutions to the matrix equation $\begin{bmatrix} a & 2a \\ b & 2b \\ c & 2c \\ d & 2d \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

Answer

Consider $\vec{v}_1 = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 2a \\ 2b \\ 2c \\ 2d \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

One particular solution of $x_1\vec{v}_1 + x_2\vec{v}_2 = \vec{b}$ is the trivial one: $0\vec{v}_1 + 0\vec{v}_2 = \vec{b}$

Therefore, we have $\vec{x}_p = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

We turn to any nontrivial homogeneous solutions of $x_1\vec{v}_1 + x_2\vec{v}_2 = \vec{0}$

Clearly, since $\vec{v}_2 = 2\vec{v}_1$, we can write $2\vec{v}_1 - 1\vec{v}_2 = \vec{0}$

This gives us the homogeneous solutions spanned by $\vec{x}_h = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

Altogether, the solution set is $\vec{x} = \vec{x}_p + t\vec{x}_h = \begin{bmatrix} 0+2t \\ 0-t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 2 \\ -1 \end{bmatrix}$

? Exercise 1.4.2b

Use inspection to find all solutions to the matrix equation $\begin{bmatrix} a & 2a \\ b & 2b \\ c & 2c \\ d & 2d \end{bmatrix} \vec{x} = \begin{bmatrix} 100a \\ 100b \\ 100c \\ 100d \end{bmatrix}$

Answer

Consider $\vec{v}_1 = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 2a \\ 2b \\ 2c \\ 2d \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 100a \\ 100b \\ 100c \\ 100d \end{bmatrix}$

One particular solution of $x_1\vec{v}_1 + x_2\vec{v}_2 = \vec{b}$ is this one: $100\vec{v}_1 + 0\vec{v}_2 = \vec{b}$

Therefore, we have $\vec{x}_p = \begin{bmatrix} 100 \\ 0 \end{bmatrix}$

We turn to any nontrivial homogeneous solutions of $x_1\vec{v}_1 + x_2\vec{v}_2 = \vec{0}$

Clearly, since $\vec{v}_2 = 2\vec{v}_1$, we can write $2\vec{v}_1 - 1\vec{v}_2 = \vec{0}$

This gives us the homogeneous solutions spanned by $\vec{x}_h = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

Altogether, the solution set is $\vec{x} = \vec{x}_p + t\vec{x}_h = \begin{bmatrix} 100+2t \\ 0-t \end{bmatrix} = \begin{bmatrix} 100 \\ 0 \end{bmatrix} + t\begin{bmatrix} 2 \\ -1 \end{bmatrix}$

Exercise 1.4.2c

Use inspection to find all solutions to the matrix equation $\begin{bmatrix} a & 0 & a \\ b & b & 2b \\ 0 & c & c \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Answer

Consider $\vec{v}_1 = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ b \\ c \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} a \\ 2b \\ c \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

One particular solution of $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{b}$ is the trivial one: $0\vec{v}_1 + 0\vec{v}_2 + 0\vec{v}_3 = \vec{b}$

Therefore, we have $\vec{x}_p = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

We turn to any nontrivial homogeneous solutions of $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{0}$

Clearly, since $\vec{v}_3 = \vec{v}_1 + \vec{v}_2$, we can write $1\vec{v}_1 + 1\vec{v}_2 - 1\vec{v}_3 = \vec{0}$

This gives us the homogeneous solutions spanned by $\vec{x}_h = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

Altogether, the solution set is $\vec{x} = \vec{x}_p + t\vec{x}_h = \begin{bmatrix} 0+t \\ 0+t \\ 0-t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

Exercise 1.4.2d

Use inspection to find all solutions to the matrix equation $\begin{bmatrix} a & 0 & a \\ b & b & 2b \\ 0 & c & c \end{bmatrix} \vec{x} = \begin{bmatrix} a \\ 0 \\ -c \end{bmatrix}$

Answer

Consider $\vec{v}_1 = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ b \\ c \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} a \\ 2b \\ c \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} a \\ 0 \\ -c \end{bmatrix}$

One particular solution of $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{b}$ is this one: $1\vec{v}_1 - 1\vec{v}_2 + 0\vec{v}_3 = \vec{b}$

Therefore, we have $\vec{x}_p = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

We turn to any nontrivial homogeneous solutions of $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{0}$

Clearly, since $\vec{v}_3 = \vec{v}_1 + \vec{v}_2$, we can write $1\vec{v}_1 + 1\vec{v}_2 - 1\vec{v}_3 = \vec{0}$

This gives us the homogeneous solutions spanned by $\vec{x}_h = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

Altogether, the solution set is $\vec{x} = \vec{x}_p + t\vec{x}_h = \begin{bmatrix} 1+t \\ -1+t \\ 0-t \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

Exercise 1.4.2e

Use inspection to find all solutions to the matrix equation $\begin{bmatrix} a & 0 & 0 & 3a \\ 0 & b & 0 & 4b \\ 0 & 0 & c & -5c \end{bmatrix} \vec{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Answer

Consider $\vec{v}_1 = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ b \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} 3a \\ 4b \\ -5c \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

One particular solution of $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 + x_4\vec{v}_4 = \vec{b}$ is the trivial one: $0\vec{v}_1 + 0\vec{v}_2 + 0\vec{v}_3 + 0\vec{v}_4 = \vec{b}$

Therefore, we have $\vec{x}_p = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

We turn to any nontrivial homogeneous solutions of $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 + x_4\vec{v}_4 = \vec{0}$

Clearly, since $\vec{v}_4 = 3\vec{v}_1 + 4\vec{v}_2 - 5\vec{v}_3$, we can write $3\vec{v}_1 + 4\vec{v}_2 - 5\vec{v}_3 - 1\vec{v}_4 = \vec{0}$

This gives us the homogeneous solutions spanned by $\vec{x}_h = \begin{bmatrix} 3 \\ 4 \\ -5 \\ -1 \end{bmatrix}$

Altogether, the solution set is $\vec{x} = \vec{x}_p + t\vec{x}_h = \begin{bmatrix} 0+3t \\ 0+4t \\ 0-5t \\ 0-t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ 4 \\ -5 \\ -1 \end{bmatrix}$

? Exercise 1.4.2f

Use inspection to find all solutions to the matrix equation $\begin{bmatrix} a & 0 & 0 & 3a \\ 0 & b & 0 & 4b \\ 0 & 0 & c & -5c \end{bmatrix}\vec{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

Answer

Consider $\vec{v}_1 = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 0 \\ b \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix}, \vec{v}_4 = \begin{bmatrix} 3a \\ 4b \\ -5c \end{bmatrix}$, and $\vec{b} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$

One particular solution of $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 + x_4\vec{v}_4 = \vec{b}$ is this one: $1\vec{v}_1 + 1\vec{v}_2 + 1\vec{v}_3 + 0\vec{v}_4 = \vec{b}$

Therefore, we have $\vec{x}_p = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$

We turn to any nontrivial homogeneous solutions of $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 + x_4\vec{v}_4 = \vec{0}$

Clearly, since $\vec{v}_4 = 3\vec{v}_1 + 4\vec{v}_2 - 5\vec{v}_3$, we can write $3\vec{v}_1 + 4\vec{v}_2 - 5\vec{v}_3 - 1\vec{v}_4 = \vec{0}$

This gives us the homogeneous solutions spanned by $\vec{x}_h = \begin{bmatrix} 3 \\ 4 \\ -5 \\ -1 \end{bmatrix}$

Altogether, the solution set is $\vec{x} = \vec{x}_p + t\vec{x}_h = \begin{bmatrix} 1+3t \\ 1+4t \\ 1-5t \\ 0-t \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ 4 \\ -5 \\ -1 \end{bmatrix}$

? Exercise 1.4.3

Let A be a matrix with m rows and n columns. Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If the reduced row echelon form of A has a pivot in every column, then $A\vec{x} = \vec{0}$ has infinitely many solutions.

(b) If the reduced row echelon form of A has a pivot in every row, then $A\vec{x} = \vec{0}$ has infinitely many solutions.

(c) If $A\vec{u} = A\vec{v}$, then $\vec{u} - \vec{v}$ is a solution to $A\vec{x} = \vec{0}$.

(d) If $A\vec{u} = A\vec{v}$, then $\vec{u} + \vec{v}$ is a solution to $A\vec{x} = \vec{0}$.

(e) If $A\vec{u} = \vec{0}$, then $c\vec{u}$ is a solution to $A\vec{x} = \vec{0}$ for all c.

(f) If $A\vec{u} = \vec{b}$ for some $\vec{b} \neq \vec{0}$, then $c\vec{u}$ is a solution to $A\vec{x} = \vec{b}$ for all c.

(g) If $\vec{b}$ is a linear combination of the columns of A, then $A\vec{x} = \vec{b}$ is consistent.

(h) If $A\vec{x} = \vec{b}$ has two distinct solutions, then $A\vec{x} = \vec{0}$ must have infinitely many solutions.

(i) If $A\vec{x} = \vec{b}$ is inconsistent for some $\vec{b} \neq \vec{0}$, then $A\vec{x} = \vec{0}$ is inconsistent.

(j) When $m = n$, there is at least one vector $\vec{x} \in \mathbb{R}^n$ satisfying $A\vec{x} = \vec{x}$.

(k) The solution set to $A\vec{x} = \vec{b}$ can be written as a span of vectors.

(l) The set of $\vec{b}$ such that $A\vec{x} = \vec{b}$ is consistent can be written as a span of vectors.

Answer

(a) **Always False**

A pivot in every column means no free variables, so the system will have exactly one or zero solutions.

(b) **Sometimes True Sometimes False**

This is true when there are more columns than rows, but false when the matrix is square.

(c) **Always True**

$A(\vec{u} - \vec{v}) = A\vec{u} - A\vec{v} = \vec{0}$ always holds.

(d) **Sometimes True Sometimes False**

When $\vec{u} = \vec{0}$ this is definitely true, but if $A\vec{u} \neq \vec{0}$ then $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v} = 2A\vec{u} \neq \vec{0}$ makes this false.

(e) **Always True**

If $A\vec{u} = \vec{0}$, then $A(c\vec{u}) = c(A\vec{u}) = c\vec{0} = \vec{0}$ occurs.

(f) **Always False**

If $A\vec{u} = \vec{b}$, then $A(c\vec{u}) = c(A\vec{u}) = c\vec{b} \neq \vec{b}$ whenever $c \neq 1$.

(g) **Always True**

If the columns of A are $\vec{v}_i$ and $\vec{b} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n$, then $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ satisfies $A\vec{x} = \vec{b}$.

(h) **Always True**

If $A\vec{u}_1 = A\vec{u}_2$, then $A(\vec{u}_1 - \vec{u}_2) = A\vec{u}_1 - A\vec{u}_2 = \vec{0}$ and since $\vec{u}_1 - \vec{u}_2 \neq \vec{0}$, there are infinitely many multiples of this vector in the homogeneous solution set.

(i) **Always False**

The homogeneous system $A\vec{x} = \vec{0}$ is always consistent, since it is satisfied by the trivial solution.

(j) **Always True**

The zero vector always satisfies $A\vec{0} = \vec{0}$.

(k) **Sometimes True Sometimes False**

This is true when $\vec{b} = \vec{0}$, but false when $\vec{b} \neq \vec{0}$.

(l) **Always True**

The span of the columns of A contains all linear combinations of those columns, which are precisely the vectors $\vec{b}$ that make $A\vec{x} = \vec{b}$ consistent.

1.4: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.5: Linear Independence

Objectives

1. Understand the concept of linear independence.
2. Learn two criteria for linear independence.
3. Understand the relationship between linear independence and pivot columns / free variables.
4. *Recipe:* test if a set of vectors is linearly independent / find an equation of linear dependence.
5. *Picture:* whether a set of vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$ is linearly independent or not.
6. *Vocabulary words:* ***linear dependence relation*** / ***equation of linear dependence***.
7. *Essential vocabulary words:* ***linearly independent***, ***linearly dependent***.

Sometimes the span of a set of vectors is "smaller" than you expect from the number of vectors, as in the picture below. This means that (at least) one of the vectors is redundant: it can be removed without affecting the span. In the present section, we formalize this idea in the notion of *linear independence*.

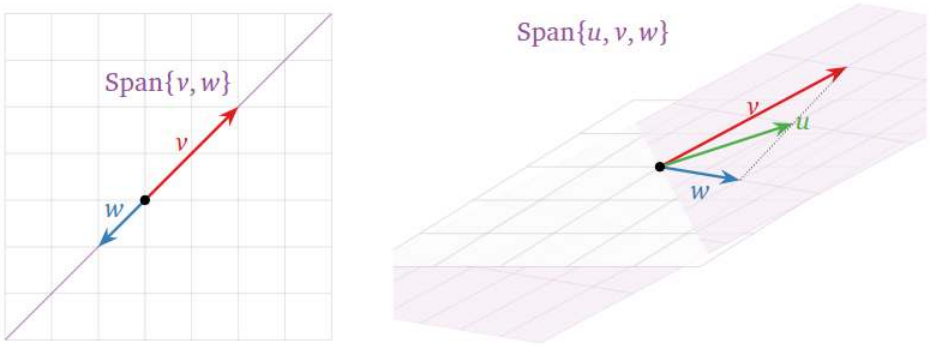


Figure **1.5.1**: Pictures of sets of vectors that are linearly dependent. Note that in each case, one vector is in the span of the others—so it doesn't make the span bigger.

This page titled 1.5: Linear Independence is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **2.5: Linear Independence** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.5.1: Defining Linear Independence

Definition: Linearly Independent and Linearly Dependent

A set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is ***linearly independent*** if the vector equation

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k = \vec{0}$$

has only the trivial solution $x_1 = x_2 = \cdots = x_k = 0$. The set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is ***linearly dependent*** otherwise.

In other words, $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly dependent if there exist numbers $x_1, x_2, \ldots, x_k$, not all equal to zero, such that

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_k\vec{v}_k = \vec{0}$$

This is called a ***linear dependence relation*** or ***equation of linear dependence***.

Note: 1.5.1.1

Note that linear dependence and linear independence are notions that apply to a *collection of vectors*. It does not make sense to say things like "this vector is linearly dependent on these other vectors," or "this matrix is linearly independent."

Example 1.5.1.1: Checking linear dependence

Is the set

$$\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\2\end{bmatrix}, \begin{bmatrix}3\\1\\4\end{bmatrix}\right\}$$

linearly independent?

Solution

Equivalently, we are asking if the homogeneous vector equation

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

has a nontrivial solution. We solve this by forming a matrix and row reducing (we do not need to augment when homogeneous):

$$\begin{bmatrix}1 & 1 & 3\\1 & -1 & 1\\1 & 2 & 4\end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix}1 & 0 & 2\\0 & 1 & 1\\0 & 0 & 0\end{bmatrix}$$

This says $x = -2z$ and $y = -z$. So there exist infinitely many solutions, most of which are nontrivial solutions.

For instance, taking $z = 1$ gives this equation of linear dependence:

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

Thus, the set is $\boxed{\text{linearly dependent}}$

Example 1.5.1.2: Checking linear independence

Is the set

$$\left\{\begin{bmatrix}1\\1\\-2\end{bmatrix},\begin{bmatrix}1\\-1\\2\end{bmatrix},\begin{bmatrix}3\\1\\4\end{bmatrix}\right\}$$

linearly independent?

Solution

Equivalently, we are asking if the homogeneous vector equation

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

has a nontrivial solution. We solve this by forming a matrix and row reducing:

$$\begin{bmatrix}1&1&3\\1&-1&1\\-2&2&4\end{bmatrix}\xrightarrow{\text{RREF}}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

This says $x=y=z=0$, i.e., the only solution is the trivial solution. We conclude that **the set is linearly independent**

Illustration 1.5.1.1: Vector parametric form

An important observation is that the vectors coming from the parametric vector form of the solution of a matrix equation $A\vec{x}=\vec{0}$ are linearly independent. Consider the solution set of $\begin{bmatrix}1&-1&2\\-2&2&-4\end{bmatrix}\vec{x}=\vec{0}$

$$\begin{bmatrix}1&-1&2\\-2&2&-4\end{bmatrix}\xrightarrow{\text{RREF}}\begin{bmatrix}1&-1&2\\0&0&0\end{bmatrix}$$

We put this solution into parametric vector form:

$$\vec{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=x_2\begin{bmatrix}1\\1\\0\end{bmatrix}+x_3\begin{bmatrix}-2\\0\\1\end{bmatrix}$$

So the vectors $\vec{v}_1=\begin{bmatrix}1\\1\\0\end{bmatrix}$ and $\vec{v}_2=\begin{bmatrix}-2\\0\\1\end{bmatrix}$ are clearly **linearly independent**

We show this by setting up a linear combination of them equal to the zero vector:

$$\begin{bmatrix}0\\0\\0\end{bmatrix}=x_2\begin{bmatrix}1\\1\\0\end{bmatrix}+x_3\begin{bmatrix}-2\\0\\1\end{bmatrix}=\begin{bmatrix}x_2-2x_3\\x_2\\x_3\end{bmatrix}$$

Comparing the second and third coordinates, we see that $x_2=x_3=0$. This reasoning will work in any example, since the entries corresponding to the free variables are all equal to 1 or 0, and are only equal to 1 in one of the vectors.

The above examples lead to the following recipe.

Recipe: Checking Linear Independence

A set of vectors $\{\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_k\}$ is linearly independent if and only if the vector equation

$$x_1\vec{v}_1+x_2\vec{v}_2+\cdots+x_k\vec{v}_k=\vec{0}$$

has only the trivial solution, if and only if the matrix equation $A\vec{x} = \vec{0}$ has only the trivial solution, where A is the matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$:

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k \\ | & | & & | \end{bmatrix}$$

This is true if and only if A has a pivot position in every column.

Solving the matrix equation $A\vec{x} = \vec{0}$ will either verify that the columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ are linearly independent (no free variables), or will produce a linear dependence relation by substituting any nonzero values for the free variables.

(Recall that $A\vec{x} = \vec{0}$ has a nontrivial solution if and only if A has a column without a pivot)

Suppose that A has more columns than rows. Then A cannot have a pivot in every column (it has at most one pivot per row), so its columns are automatically linearly dependent.

Note 1.5.1.2

A wide matrix (a matrix with more columns than rows) has linearly dependent columns.

For example, four vectors in $\mathbb{R}^3$ are automatically linearly dependent. Note that a tall matrix may or may not have linearly independent columns.

Fact 1.5.1.1: Facts About Linear Independence

1. Two vectors are linearly dependent if and only if they are collinear, i.e., one is a scalar multiple of the other.
2. Any set containing the zero vector is linearly dependent.
3. If a subset of $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly dependent, then $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly dependent as well.

Proof

1. If $\vec{v}_1 = c\vec{v}_2$ then $\vec{v}_1 - c\vec{v}_2 = \vec{0}$, so $\{\vec{v}_1, \vec{v}_2\}$ is linearly dependent. In the other direction, if $x_1\vec{v}_1 + x_2\vec{v}_2 = 0$, for $x_1 \neq 0$, we get $\vec{v}_1 = -\frac{x_2}{x_1}\vec{v}_2$
2. It is easy to produce a linear dependence relation if one vector is the zero vector: for instance, if $\vec{v}_1 = 0$ then
$$1 \cdot \vec{v}_1 + 0 \cdot \vec{v}_2 + \cdots + 0 \cdot \vec{v}_k = \vec{0}$$
3. After reordering, we may suppose that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r\}$ is linearly dependent, with $r < k$. This means that there is an equation of linear dependence
$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_r\vec{v}_r = \vec{0}$$
with at least one of $x_1, x_2, \ldots, x_r$ nonzero. This is also an equation of linear dependence among $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, since we can take the coefficients of $\vec{v}_{r+1}, \ldots, \vec{v}_k$ to all be zero.

With regard to the first fact, note that the zero vector is a multiple of any vector, so it is collinear with any other vector. Hence, facts 1 and 2 are consistent with each other.

1.5.1: Defining Linear Independence is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.5: Linear Independence** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.5.2: Understanding Linear Independence

In this subsection we give two criteria for a set of vectors to be linearly independent. Keep in mind, however, that the actual definition for linear independence, Definition **1.5.2.1**, is above.

Theorem 1.5.2.1

A set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly dependent if and only if one of the vectors is in the span of the other ones.

Any such vector may be removed without affecting the span.

Proof

Suppose, for instance, that $\vec{v}_3$ is in $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$, so we have an equation like

$$\vec{v}_3 = 2\vec{v}_1 - \frac{1}{2}\vec{v}_2 + 6\vec{v}_4$$

We can subtract $\vec{v}_3$ from both sides of the equation to get

$$\vec{0} = 2\vec{v}_1 - \frac{1}{2}\vec{v}_2 - \vec{v}_3 + 6\vec{v}_4$$

This is a linear dependence relation.

In this case, any linear combination of $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ is already a linear combination of $\vec{v}_1, \vec{v}_2, \vec{v}_4$:

$$\begin{aligned} x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 + x_4\vec{v}_4 &= x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\left(2\vec{v}_1 - \frac{1}{2}\vec{v}_2 + 6\vec{v}_4\right) + x_4\vec{v}_4 \\ &= (x_1 + 2x_3)\vec{v}_1 + \left(x_2 - \frac{1}{2}x_3\right)\vec{v}_2 + (x_4 + 6)\vec{v}_4 \end{aligned}$$

Therefore, $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is contained in $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$. Any linear combination of $\vec{v}_1, \vec{v}_2, \vec{v}_4$ is also a linear combination of $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4$ (with the $\vec{v}_3$-coefficient equal to zero), so $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_4\}$ is also contained in $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$, and thus they are equal.

In the other direction, if we have a linear dependence relation like

$$\vec{0} = 2\vec{v}_1 - \frac{1}{2}\vec{v}_2 + \vec{v}_3 - 6\vec{v}_4$$

then we can move any nonzero term to the left side of the equation and divide by its coefficient:

$$\vec{v}_1 = \frac{1}{2}\left(\frac{1}{2}\vec{v}_2 - \vec{v}_3 + 6\vec{v}_4\right)$$

This shows that $\vec{v}_1$ is in $\mathbf{Span}\{\vec{v}_2, \vec{v}_3, \vec{v}_4\}$.

We leave it to the reader to generalize this proof for any set of vectors.

Warning

In a linearly dependent set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$, it is not generally true that *any* vector $\vec{v}_j$ is in the span of the others, only that *at least one* of them is.

For example, the set $\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}2\\2\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$ is linearly dependent, but $\begin{bmatrix}0\\0\\1\end{bmatrix}$ is not in the span of the other two vectors.

The previous Theorem **1.5.2.1** makes precise in what sense a set of linearly dependent vectors is redundant.

Theorem 1.5.2.2: Increasing Span Criterion

A set of vectors $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly independent if and only if, for every j, the vector $\vec{v}_j$ is not in $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{j-1}\}$.

Proof

It is equivalent to show that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly dependent if and only if $\vec{v}_j$ is in $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{j-1}\}$ for some j. The "if" implication is an immediate consequence of the previous Theorem 1.5.2.1. Suppose then that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly dependent. This means that some $\vec{v}_j$ is in the span of the others. Choose the largest such j. We claim that this $\vec{v}_j$ is in $\mathbf{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{j-1}\}$. If not, then

$$\vec{v}_j = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_{j-1}\vec{v}_{j-1} + x_{j+1}\vec{v}_{j+1} + \cdots + x_k\vec{v}_k$$

with not all of $x_{j+1}, \ldots, x_k$ equal to zero. Suppose for simplicity that $x_k \neq 0$. Then we can rearrange:

$$\vec{v}_k = -\frac{1}{x_k}\left(x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_{j-1}\vec{v}_{j-1} - \vec{v}_j + x_{j+1}\vec{v}_{j+1} + \cdots + x_{p-1}\vec{v}_{p-1}\right)$$

This says that $\vec{v}_k$ is in the span of $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{p-1}\}$, which contradicts our assumption that $\vec{v}_j$ is the last vector in the span of the others.

We can rephrase this as follows:

Note 1.5.2.4

If you make a set of vectors by adding one vector at a time, and if the span got bigger every time you added a vector, then your set is linearly independent.

A set containing one vector $\{\vec{v}\}$ is linearly independent when $\vec{v} \neq 0$, since $x\vec{v} = 0$ implies $x = 0$.

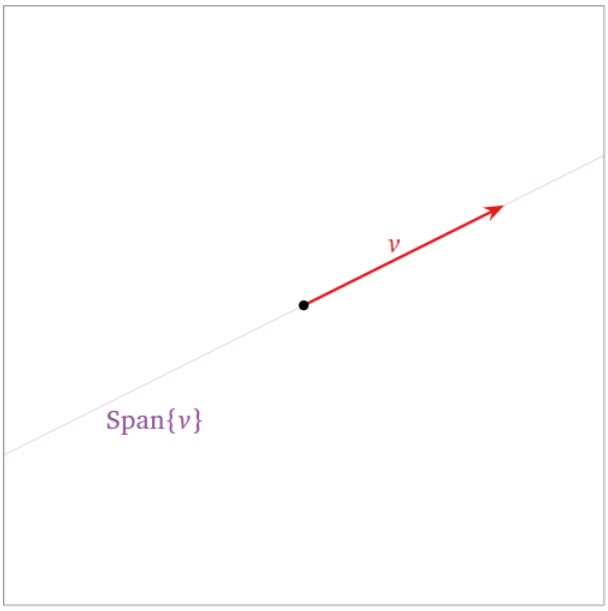


Figure 1.5.2.2

A set of two noncollinear vectors $\{\vec{v}, \vec{w}\}$ is linearly independent:

- Neither is in the span of the other, so we can apply the first criterion, Theorem 1.5.2.1.
- The span got bigger when we added $\vec{w}$, so we can apply the increasing span criterion, Theorem 1.5.2.2.

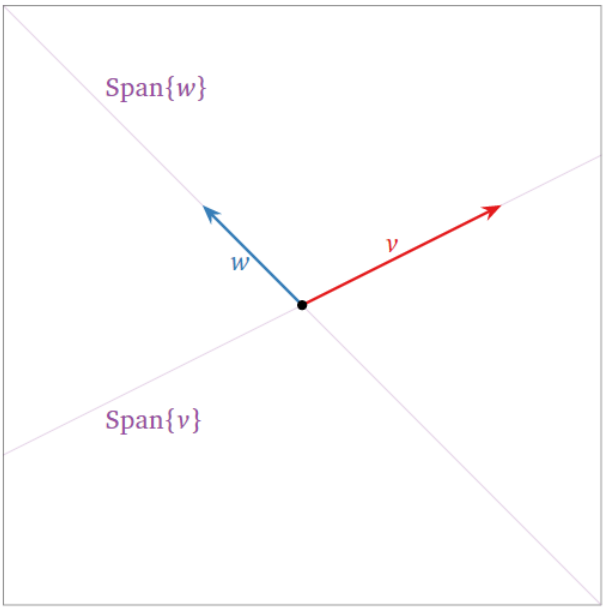


Figure 1.5.2.3

The set of three vectors $\{\vec{v},\vec{w},\vec{u}\}$ below is linearly dependent:

- $\vec{u}$ is in $\text{Span}\{\vec{v},\vec{w}\}$, so we can apply the first criterion, Theorem 1.5.2.1.
- The span did not increase when we added $\vec{u}$, so we can apply the increasing span criterion, Theorem 1.5.2.2.

In the picture below, note that $\vec{v}$ is in $\text{Span}\{\vec{u},\vec{w}\}$, and $\vec{w}$ is in $\text{Span}\{\vec{u},\vec{v}\}$, so we can remove any of the three vectors without shrinking the span.

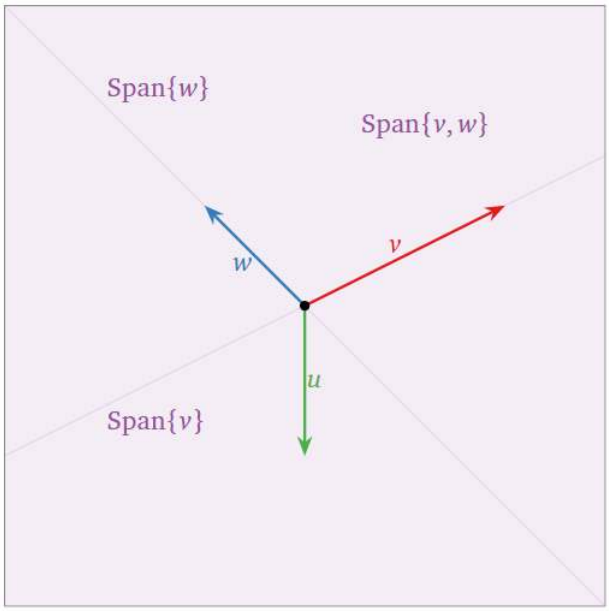


Figure 1.5.2.4

The three vectors $\{\vec{v},\vec{w},\vec{u}\}$ below are linearly independent: the span got bigger when we added $\vec{w}$, then again when we added $\vec{u}$, so we can apply the increasing span criterion, Theorem 1.5.2.2.

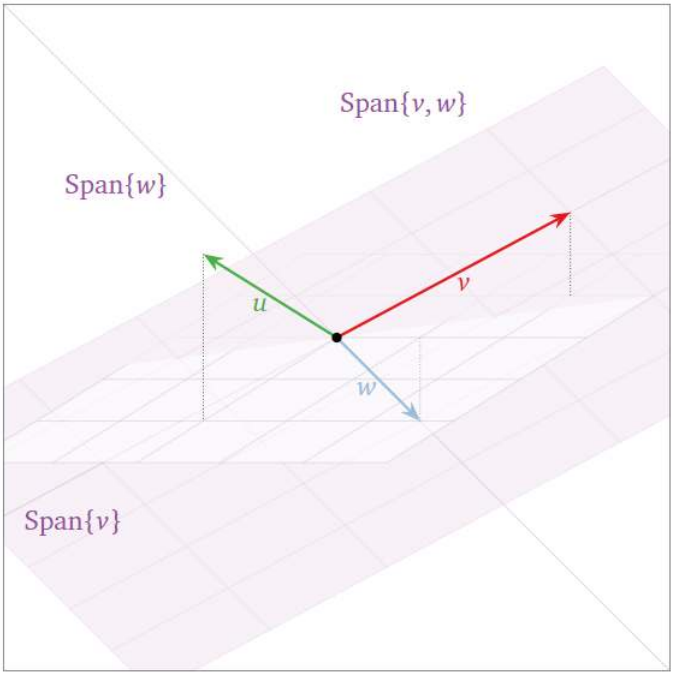


Figure 1.5.2.5

The three coplanar vectors $\{\vec{v}, \vec{w}, \vec{u}\}$ below are linearly dependent:

- $\vec{u}$ is in $\text{Span}\{\vec{v}, \vec{w}\}$, so we can apply the first criterion, Theorem 1.5.2.1.
- The span did not increase when we added $\vec{u}$, so we can apply the increasing span criterion, Theorem 1.5.2.2.

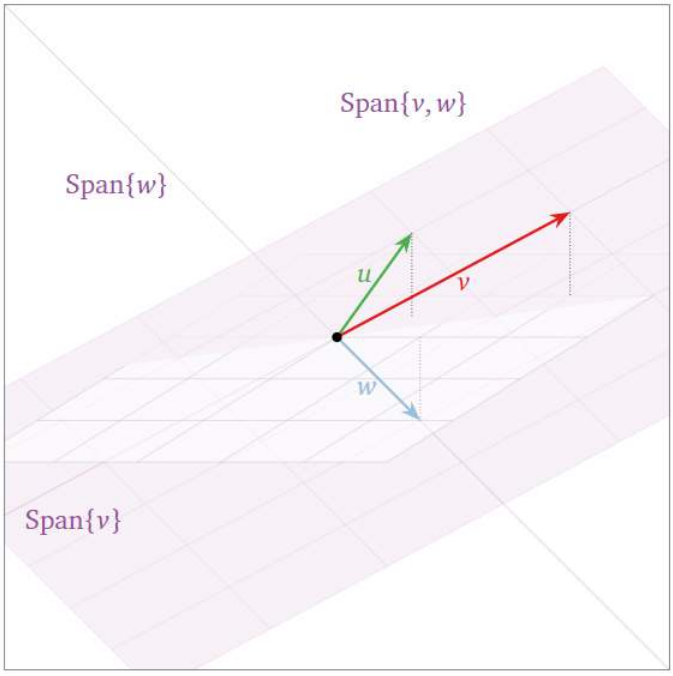


Figure 1.5.2.6

Note that three vectors are linearly dependent if and only if they are *coplanar*. Indeed, $\{\vec{v}, \vec{w}, \vec{u}\}$ is linearly dependent if and only if one vector is in the span of the other two, which is a plane (or a line) (or $\{\vec{0}\}$).

In light of this important note, Recipe: Checking Linear Independence, and this criterion, Theorem 1.5.2.1, it is natural to ask which columns of a matrix are redundant, i.e., which we can remove without affecting the column span.

Theorem 1.5.2.3

Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ be vectors in $\mathbb{R}^n$, and consider the matrix

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k \\ | & | & & | \end{bmatrix}$$

Then we can delete the columns of A *without* pivots (the columns corresponding to the free variables), without changing $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$.

The pivot columns are linearly independent, so we cannot delete any other columns without changing the span.

Proof

If we set up the homogeneous equation $A\vec{x} = \vec{0}$, we solve by putting A into RREF.

If there are no free variables, we must have $\vec{x} = \vec{0}$ the trivial solution only, and our columns are linearly independent.

If there is at least one free variable x_j, then there is a nontrivial solution $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_k \end{bmatrix}$ with $x_j = 1$. This gives us the linear dependence relation:

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + \vec{v}_j + \cdots + x_k\vec{v}_k = \vec{0}$$

This means $\vec{v}_j \in \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{j-1}, \vec{v}_{j+1}, \ldots, \vec{v}_k\}$

And thus, $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{j-1}, \vec{v}_{j+1}, \ldots, \vec{v}_k\} = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{j-1}, \vec{v}_j, \vec{v}_{j+1}, \ldots, \vec{v}_k\}$

Therefore, if we remove $\vec{v}_j$ the span will remain the same. We can do this for each free variable column until only pivot columns remain. These remaining columns will have the same span, but when put into RREF there will only be pivot columns so they must be linearly independent.

Note that it is necessary to row reduce A to find which are its pivot columns. However, the span of the columns of the row reduced matrix is generally *not* equal to the span of the columns of A: one must use the pivot columns of the *original* matrix.

Illustration 1.5.2.2

The matrix

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

has reduced row echelon form

$$\begin{bmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Therefore, the first two columns of A are the pivot columns, so we can delete the others without changing the span:

$$\text{Span}\left\{\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ -2 \end{bmatrix}\right\}.$$

Moreover, the first two columns are linearly independent.

Note 1.5.2.5: Pivot Columns and Dimension

Let d be the number of pivot columns in the matrix

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_k \\ | & | & & | \end{bmatrix}$$

- If $d = 1$ then $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is a line.
- If $d = 2$ then $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is a plane.
- If $d = 3$ then $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is a 3-space.
- Et cetera.

The number d is called the dimension. We will define this concept rigorously in the future!

There are two direct consequences of the Increasing Span Criterion:

If $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly independent, then any subset of these vectors forms a linearly independent set.
If $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is linearly dependent, then any set containing all of these vectors forms a linearly dependent set.

Example 1.5.2.3

Assume the set $\{\vec{u}, \vec{v}\}$ is linearly independent, and define $\vec{a} = \vec{u} + \vec{v}, \vec{b} = \vec{u} - \vec{v}$

Is $\{\vec{a}, \vec{b}\}$ linearly independent?

Solution

Since we are not given the values inside of $\vec{u}, \vec{v}$, we cannot manually calculate the values of $\vec{a}, \vec{b}$ to check for linear independence directly (by setting them as columns of a matrix and row reducing)

Instead, we must check if there is some linear combination

$$x\vec{a} + y\vec{b} = \vec{0}$$

This can be re-written as a linear combination of the vectors $\vec{u}, \vec{v}$:

$$x(\vec{u} + \vec{v}) + y(\vec{u} - \vec{v}) = \vec{0}$$

$$(x + y)\vec{u} + (x - y)\vec{v} = \vec{0}$$

Luckily, we know $\{\vec{u}, \vec{v}\}$ is linearly independent, so this can only be satisfied if $\begin{cases} x + y = 0 \\ x - y = 0 \end{cases}$

We can solve this system of equations, and quickly see it is only satisfied when $x = 0$ and $y = 0$.

This means the only linear combination of $\vec{a}$ and $\vec{b}$ that makes $\vec{0}$ is the trivial one.

This proves that $\{\vec{a}, \vec{b}\}$ is a **linearly independent set**

Example 1.5.2.4

Assume the set $\{\vec{u}, \vec{v}, \vec{w}\}$ is linearly independent, and define $\vec{a} = 2\vec{u} - 2\vec{v}, \vec{b} = \vec{v} - 2\vec{w}, \vec{c} = 2\vec{w} - \vec{u}$

Is $\{\vec{a}, \vec{b}, \vec{c}\}$ linearly independent?

Solution

We must check if there is some linear combination

$$x\vec{a} + y\vec{b} + z\vec{c} = \vec{0}$$

This can be re-written as a linear combination of the vectors $\vec{u}, \vec{v}, \vec{w}$:

$$x\left(2\vec{u}-2\vec{v}\right)+y\left(\vec{v}-2\vec{w}\right)+z\left(2\vec{w}-\vec{u}\right)=\vec{0}$$

$$\left(2x-z\right)\vec{u}+\left(-2x+y\right)\vec{v}+\left(-2y+2z\right)\vec{w}=\vec{0}$$

Again, we know $\{\vec{u},\vec{v},\vec{w}\}$ is linearly independent, so this can only be satisfied if $\begin{cases}2x-z=0\\-2x+y=0\\-2y+2z=0\end{cases}$

We can solve this system of equations, and quickly see it has parametric form $y=z$ and $x=\frac{1}{2}z$.

This means there is at least one nontrivial linear combination of $\vec{a},\vec{b}$ and $\vec{c}$ that makes $\vec{0}$.

For example, letting $z=2$, we get $x=1,y=2$:

$$1\vec{a}+2\vec{b}+2\vec{c}=1\left(2\vec{u}-2\vec{v}\right)+2\left(\vec{v}-2\vec{w}\right)+2\left(2\vec{w}-\vec{u}\right)=0\vec{u}+0\vec{v}+0\vec{w}=\vec{0}$$

This proves that $\{\vec{a},\vec{b},\vec{c}\}$ is a **linearly dependent set**

1.5.2: Understanding Linear Independence is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.5: Linear Independence** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.5.3: Practice Set

? Exercise 1.5.3.1*a*

Determine if the set $\left\{\begin{bmatrix}1\\2\\2\end{bmatrix}, \begin{bmatrix}2\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}1&2&1\\2&1&1\\2&1&1\end{bmatrix} \xrightarrow{RREF} \begin{bmatrix}1&0&\frac{1}{3}\\0&1&\frac{1}{3}\\0&0&0\end{bmatrix}$

No pivot in third column, this set is **linearly dependent.**

Observation: The sum of the first two columns is triple the third column, so the third column must be in the span of the first two. This is a linearly dependent set.

? Exercise 1.5.3.1*b*

Determine if the set $\left\{\begin{bmatrix}4\\0\\-2\end{bmatrix}, \begin{bmatrix}-2\\0\\1\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}4&-2\\0&0\\-2&1\end{bmatrix} \xrightarrow{RREF} \begin{bmatrix}1&-\frac{1}{2}\\0&0\\0&0\end{bmatrix}$

No pivot in second column, this set is **linearly dependent.**

Observation: The first column is a scalar multiple (-2) of the second column. This is a linearly dependent set.

? Exercise 1.5.3.1*c*

Determine if the set $\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}1&1&0\\0&1&1\\1&1&1\end{bmatrix} \xrightarrow{RREF} \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$

A pivot in every column, this set is **linearly independent.**

Observation: None of these vectors can be created by the other two. This is a linearly independent set.

? Exercise 1.5.3.1*d*

Determine if the set $\left\{\begin{bmatrix}4\\1\\7\end{bmatrix}, \begin{bmatrix}1\\0\\2\end{bmatrix}, \begin{bmatrix}0\\-1\\1\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}4&1&0\\1&0&-1\end{bmatrix} \xrightarrow{RREF} \begin{bmatrix}1&0&-1\\0&1&4\end{bmatrix}$

No pivot in third column, this set is **linearly dependent.**

Observation: Trickier to see, but the sum of the first and third column is exactly four times the center column. This is a linearly dependent set.

? Exercise 1.5.3.1e

Determine if the set $\left\{\begin{bmatrix}5\\-1\\4\\1\end{bmatrix},\begin{bmatrix}5\\1\\4\\1\end{bmatrix},\begin{bmatrix}5\\0\\4\\1\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}5&5&5\\-1&1&0\\4&4&4\\1&1&1\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0&\frac{1}{2}\\0&1&\frac{1}{2}\\0&0&0\\0&0&0\end{bmatrix}$

No pivot in third column, this set is **linearly dependent.**

Observation: The sum of the first and second column is exactly double times the third column. This is a linearly dependent set.

? Exercise 1.5.3.1f

Determine if the set $\left\{\begin{bmatrix}14\\23\end{bmatrix},\begin{bmatrix}15\\22\end{bmatrix},\begin{bmatrix}16\\25\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}14&15&16\\23&22&25\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0&\frac{23}{37}\\0&1&\frac{18}{37}\end{bmatrix}$

No pivot in third column, this set is **linearly dependent.**

Observation: This matrix is wider than it is tall, making it impossible to have a pivot in every column. This is a linearly dependent set.

? Exercise 1.5.3.1g

Determine if the set $\left\{\begin{bmatrix}\sqrt{2}\\\sqrt{3}\\\sqrt{5}\\\sqrt{6}\end{bmatrix},\begin{bmatrix}\sqrt[3]{2}\\\sqrt[3]{3}\\\sqrt[3]{5}\\\sqrt[3]{6}\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}\sqrt{2}&\sqrt[3]{2}\\\sqrt{3}&\sqrt[3]{3}\\\sqrt{5}&\sqrt[3]{5}\\\sqrt{6}&\sqrt[3]{6}\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0\\0&1\\0&0\\0&0\end{bmatrix}$

A pivot in every column, this set is **linearly independent.**

Observation: These two vectors are clearly not scalar multiplies of one another. This is a linearly independent set.

? Exercise 1.5.3.1*h*

Determine if the set $\left\{\begin{bmatrix}1\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\\1\end{bmatrix},\begin{bmatrix}1\\0\\0\\1\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}1&0&0&1\\1&1&0&0\\0&1&1&0\\0&0&1&1\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0&0&1\\0&1&0&-1\\0&0&1&1\\0&0&0&0\end{bmatrix}$

No pivot in every column, this set is **linearly dependent.**

Observation: The sum of the first and third vector is the same as the sum of the second and fourth vector.

? Exercise 1.5.3.1*i*

Determine if the set $\left\{\begin{bmatrix}1\\3\\0\\5\end{bmatrix},\begin{bmatrix}2\\1\\0\\1\end{bmatrix},\begin{bmatrix}11\\11\\0\\1\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}1&2&11\\3&1&11\\0&0&0\\5&1&1\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\\0&0&0\end{bmatrix}$

A pivot in every column, this set is **linearly independent.**

Observation: None of these vectors can be created by the other two. This is a linearly independent set.

? Exercise 1.5.3.1*j*

Determine if the set $\left\{\begin{bmatrix}6\\-4\\-3\\11\end{bmatrix},\begin{bmatrix}0\\0\\0\\0\end{bmatrix},\begin{bmatrix}-12\\13\\0\\5\end{bmatrix},\begin{bmatrix}100\\99\\-1\\-7\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix}6&0&-12&100\\-4&0&13&99\\-3&0&0&-1\\11&0&5&7\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&0&0&1\\0&0&0&0\end{bmatrix}$

No pivot in second column, this set is **linearly dependent.**

Observation: Any set containing the zero vector is dependent. This is a linearly dependent set.

? Exercise 1.5.3.1*k*

Determine if the set $\left\{\begin{bmatrix}\cos\left(\frac{\pi}{5}\right)\\\cot\left(\frac{\pi}{5}\right)\\1\end{bmatrix},\begin{bmatrix}1\\\csc\left(\frac{\pi}{5}\right)\\\sec\left(\frac{\pi}{5}\right)\end{bmatrix}\right\}$ is linearly independent or dependent.

Answer

Manually: $\begin{bmatrix} \cos\left(\frac{\pi}{5}\right) & 1 \\ \cot\left(\frac{\pi}{5}\right) & \csc\left(\frac{\pi}{5}\right) \\ 1 & \sec\left(\frac{\pi}{5}\right) \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & \sec\left(\frac{\pi}{5}\right) \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

No pivot in second column, this set is **linearly dependent.**

Observation: Trigonometrically inclined students should see that $\cos\left(\frac{\pi}{5}\right)$ times the second column is the first column. This is a linearly dependent set.

Exercise 1.5.3.2*a*

Assume that the set $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is linearly independent and each $\vec{v}_i \in \mathbb{R}^4$.

Determine if the set $\{2\vec{v}_1, 3\vec{v}_2, 4\vec{v}_3, 5\vec{v}_4\}$ is linearly independent or dependent.

Answer

We investigate the equation: $x_1(2\vec{v}_1) + x_2(3\vec{v}_2) + x_3(4\vec{v}_3) + x_4(5\vec{v}_4) = \vec{0}$

This can be written as $(2x_1)\vec{v}_1 + (3x_2)\vec{v}_2 + (4x_3)\vec{v}_3 + (5x_4)\vec{v}_4 = \vec{0}$

Since $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is independent, we know each coefficient must be zero, so

$$\begin{cases} 2x_1 = 0 \\ 3x_2 = 0 \\ 4x_3 = 0 \\ 5x_4 = 0 \end{cases} \xrightarrow{\text{matrix then RREF}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

There are no free variables, so the only solution is the trivial one $x_1 = x_2 = x_3 = x_4 = 0$, and this set is **linearly independent.**

Exercise 1.5.3.2*b*

Assume that the set $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is linearly independent and each $\vec{v}_i \in \mathbb{R}^4$.

Determine if the set $\{\vec{v}_1, \vec{v}_2, \vec{v}_2 + \vec{v}_4, \vec{v}_3 + \vec{v}_4\}$ is linearly independent or dependent.

Answer

We investigate the equation: $x_1(\vec{v}_1) + x_2(\vec{v}_2) + x_3(\vec{v}_2 + \vec{v}_4) + x_4(\vec{v}_3 + \vec{v}_4) = \vec{0}$

This can be written as $(x_1)\vec{v}_1 + (x_2 + x_3)\vec{v}_2 + (x_4)\vec{v}_3 + (x_3 + x_4)\vec{v}_4 = \vec{0}$

Since $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is independent, we know each coefficient must be zero, so

$$\begin{cases} x_1 = 0 \\ x_2 + x_3 = 0 \\ x_4 = 0 \\ x_3 + x_4 = 0 \end{cases} \xrightarrow{\text{matrix then RREF}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

There are no free variables, so the only solution is the trivial one $x_1 = x_2 = x_3 = x_4 = 0$, and this set is **linearly independent.**

Exercise 1.5.3.2*c*

Assume that the set $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\}$ is linearly independent and each $\vec{v}_i \in \mathbb{R}^4$.

Determine if the set $\{\vec{v}_1 - \vec{v}_2 - \vec{v}_3, \vec{v}_2 + \vec{v}_3, \vec{v}_1 + \vec{v}_4, \vec{v}_4\}$ is linearly independent or dependent.

Answer

We investigate the equation: $x_1(\vec{v}_1 - \vec{v}_2 - \vec{v}_3) + x_2(\vec{v}_2 + \vec{v}_3) + x_3(\vec{v}_1 + \vec{v}_4) + x_4(\vec{v}_4) = \vec{0}$

This can be written as $(x_1+x_3)\vec{v}_1+(-x_1+x_2)\vec{v}_2+(-x_1+x_2)\vec{v}_3+(x_3+x_4)\vec{v}_4=\vec{0}$

Since $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is independent, we know each coefficient must be zero, so

$$\begin{cases} x_1+x_3=0 \\ -x_1+x_2=0 \\ -x_1+x_2=0 \\ x_3+x_4=0 \end{cases} \xrightarrow{\text{matrix then RREF}} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

There is a free variable x_4, so there are nontrivial solutions (like $x_1=1, x_2=1, x_3=-1, x_4=1$), and this set is **linearly dependent.**

Note that $1(\vec{v}_1-\vec{v}_2-\vec{v}_3)+1(\vec{v}_2+\vec{v}_3)-1(\vec{v}_1+\vec{v}_4)+1(\vec{v}_4)=\vec{0}$ does hold.

? Exercise 1.5.3.2*d*

Assume that the set $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is linearly independent and each $\vec{v}_i\in\mathbb{R}^4$.

Determine if the set $\{\vec{v}_1-\vec{v}_4,\vec{v}_2-\vec{v}_1,\vec{v}_3-\vec{v}_2,\vec{v}_4-\vec{v}_3\}$ is linearly independent or dependent.

Answer

We investigate the equation: $x_1(\vec{v}_1-\vec{v}_4)+x_2(\vec{v}_2-\vec{v}_1)+x_3(\vec{v}_3-\vec{v}_2)+x_4(\vec{v}_4-\vec{v}_3)=\vec{0}$

This can be written as $(x_1-x_2)\vec{v}_1+(x_2-x_3)\vec{v}_2+(x_3-x_4)\vec{v}_3+(x_4-x_1)\vec{v}_4=\vec{0}$

Since $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is independent, we know each coefficient must be zero, so

$$\begin{cases} x_1-x_2=0 \\ x_2-x_3=0 \\ x_3-x_4=0 \\ -x_1+x_4=0 \end{cases} \xrightarrow{\text{matrix then RREF}} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

There is a free variable x_4, so there are nontrivial solutions (like $x_1=1, x_2=1, x_3=1, x_4=1$), and this set is **linearly dependent.**

Note that $1(\vec{v}_1-\vec{v}_4)+1(\vec{v}_2-\vec{v}_1)+1(\vec{v}_3-\vec{v}_2)+1(\vec{v}_4-\vec{v}_3)=\vec{0}$ does hold.

? Exercise 1.5.3.2*e*

Assume that the set $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is linearly independent and each $\vec{v}_i\in\mathbb{R}^4$.

Determine if the set $\{\vec{v}_2+\vec{v}_4,\vec{v}_1-\vec{v}_3,2\vec{v}_2+3\vec{v}_3\}$ is linearly independent or dependent.

Answer

We investigate the equation: $x_1(\vec{v}_2+\vec{v}_4)+x_2(\vec{v}_1-\vec{v}_3)+x_3(2\vec{v}_2+3\vec{v}_3)=\vec{0}$

This can be written as $(x_2)\vec{v}_1+(x_1+2x_3)\vec{v}_2+(-x_2+3x_3)\vec{v}_3+(x_1)\vec{v}_4=\vec{0}$

Since $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is independent, we know each coefficient must be zero, so

$$\begin{cases} x_2=0 \\ x_1+2x_3=0 \\ -x_2+3x_3=0 \\ x_1=0 \end{cases} \xrightarrow{\text{matrix then RREF}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

There are no free variables, so the only solution is the trivial one $x_1=x_2=x_3=0$, and this set is **linearly independent.**

? Exercise 1.5.3.2*f*

Assume that the set $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is linearly independent and each $\vec{v}_i\in\mathbb{R}^4$.

Determine if the set $\{\vec{v}_1+\vec{v}_2+\vec{v}_3+\vec{v}_4\}$ is linearly independent or dependent.

Answer

We investigate the equation: $x_1(\vec{v}_1+\vec{v}_2+\vec{v}_3+\vec{v}_4)=\vec{0}$

This can be written as $(x_1)\vec{v}_1+(x_1)\vec{v}_2+(x_1)\vec{v}_3+(x_1)\vec{v}_4=\vec{0}$

Since $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is independent, we know each coefficient must be zero, so

$$\begin{cases} x_1=0 \\ x_1=0 \\ x_1=0 \\ x_1=0 \end{cases}$$

Clearly there are no free variables, so the only solution is the trivial one $x_1=0$, and this set is **linearly independent.**

Exercise 1.5.3.2*g*

Assume that the set $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is linearly independent and each $\vec{v}_i\in\mathbb{R}^4$.

Determine if the set $\{\vec{v}_1,\vec{v}_1+\vec{v}_2,\vec{v}_1+\vec{v}_2+\vec{v}_3,\vec{v}_1-\vec{v}_4,\vec{v}_3+\vec{v}_4\}$ is linearly independent or dependent.

Answer

We investigate the equation: $x_1(\vec{v}_1)+x_2(\vec{v}_1+\vec{v}_2)+x_3(\vec{v}_1+\vec{v}_2+\vec{v}_3)+x_4(\vec{v}_1-\vec{v}_4)+x_5(\vec{v}_3+\vec{v}_4)=\vec{0}$

This can be written as $(x_1+x_2+x_3+x_4)\vec{v}_1+(x_2+x_3)\vec{v}_2+(x_3+x_5)\vec{v}_3+(-x_4+x_5)\vec{v}_4=\vec{0}$

Since $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is independent, we know each coefficient must be zero, so

$$\begin{cases} x_1+x_2+x_3+x_4=0 \\ x_2+x_3=0 \\ x_3+x_5=0 \\ -x_4+x_5=0 \end{cases} \xrightarrow{\text{matrix then RREF}} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

There is a free variable x_5, so there are nontrivial solutions (like $x_1=-1,x_2=1,x_3=-1,x_4=1,x_5=1$), and this set is **linearly dependent.**

Note that $-1(\vec{v}_1)+1(\vec{v}_1+\vec{v}_2)-1(\vec{v}_1+\vec{v}_2+\vec{v}_3)+1(\vec{v}_1-\vec{v}_4)+1(\vec{v}_3+\vec{v}_4)=\vec{0}$ does hold.

Exercise 1.5.3.2*h*

Assume that the set $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is linearly independent and each $\vec{v}_i\in\mathbb{R}^4$.

Determine if the set $\{\vec{v}_1,\vec{v}_2,37\vec{v}_2-15\vec{v}_1\}$ is linearly independent or dependent.

Answer

We investigate the equation: $x_1(\vec{v}_1)+x_2(\vec{v}_2)+x_3(37\vec{v}_2-15\vec{v}_1)=\vec{0}$

This can be written as $(x_1-15x_3)\vec{v}_1+(x_2+37x_3)\vec{v}_2+(0)\vec{v}_3+(0)\vec{v}_4=\vec{0}$

Since $\{\vec{v}_1,\vec{v}_2,\vec{v}_3,\vec{v}_4\}$ is independent, we know each coefficient must be zero, so

$$\begin{cases} x_1-15x_3=0 \\ x_2+37x_3=0 \\ 0=0 \\ 0=0 \end{cases} \xrightarrow{\text{matrix then RREF}} \begin{bmatrix} 1 & 0 & -15 \\ 0 & 1 & 37 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

There is a free variable x_3, so there are nontrivial solutions (like $x_1=15,x_2=-37,x_3=1$), and this set is **linearly dependent.**

Note that $15(\vec{v}_1)-37(\vec{v}_2)+1(37\vec{v}_2-15\vec{v}_1)=\vec{0}$ does hold.

? Exercise 1.5.3.3

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If $\{\vec{u},\vec{v}\}$ is linearly independent, then $\vec{u}+\vec{v}$ is non-zero.

(b) If $\{\vec{u},\vec{v}\}$ is linearly dependent, then $\vec{u}+\vec{v}$ is non-zero.

(c) If $\{\vec{u},\vec{v}\}$ are linearly independent vectors in $\mathbb{R}^2$, then they must span all of $\mathbb{R}^2$.

(d) If $\{\vec{u},\vec{v},\vec{w}\}$ are linearly independent vectors in $\mathbb{R}^3$, then they must span all of $\mathbb{R}^3$.

(e) If $\{\vec{u},\vec{v}\}$ is a linearly independent set and $\{\vec{u},\vec{w}\}$ is a linearly independent set, then $\{\vec{u},\vec{v},\vec{w}\}$ is linearly independent.

(f) If $\{\vec{u},\vec{v},\vec{w}\}$ are nonzero vectors satisfying $\vec{u}\notin \text{Span}\{\vec{v},\vec{w}\}$, then $\{\vec{u},\vec{v},\vec{w}\}$ is linearly independent.

(g) If $\{\vec{u},\vec{v}\}$ are vectors satisfying $\vec{u}\notin \text{Span}\{\vec{v}\}$, then $\{\vec{u},\vec{v}\}$ is linearly independent.

(h) If $\{\vec{u},\vec{v}\}$ is linearly dependent for every possible $\vec{v}\in\mathbb{R}^n$, then $\vec{u}=\vec{0}$.

Answer

(a) **Always True**

If it was zero, they would be linearly dependent by definition.

(b) **Sometimes True Sometimes False**

If $\vec{u}=\vec{v}\neq\vec{0}$, then the sum is nonzero but the set is dependent. Alternatively, if $\vec{u}=\vec{v}=\vec{0}$ the sum is zero and it is dependent.

(c) Always True

If they are columns of a 2×2 matrix A, it will row reduce to have a pivot in each column, and therefore a pivot in each row and $A\vec{x}=\vec{b}$ cannot be made inconsistent.

(d) **Always True**

If they are columns of a 3×3 matrix A, it will row reduce to have a pivot in each column, and therefore a pivot in each row and $A\vec{x}=\vec{b}$ cannot be made inconsistent.

(e) **Sometimes True Sometimes False**

This is true for the example $\vec{u}=\begin{bmatrix}1\\0\\0\end{bmatrix},\vec{v}=\begin{bmatrix}0\\1\\0\end{bmatrix},\vec{w}=\begin{bmatrix}0\\0\\1\end{bmatrix}$ and it is false for the example $\vec{u}=\begin{bmatrix}1\\0\\0\end{bmatrix},\vec{v}=\begin{bmatrix}0\\1\\0\end{bmatrix},\vec{w}=\begin{bmatrix}1\\1\\0\end{bmatrix}$.

(f) **Sometimes True Sometimes False**

When $\{\vec{v},\vec{w}\}$ is linearly independent, this statement is true. When $\{\vec{v},\vec{w}\}$ is linearly dependent, this statement is false.

(g) **Sometimes True Sometimes False**

This is true when $\vec{v}\neq\vec{0}$, but false for $\vec{v}=\vec{0}$.

(h) **Always True**

If $\vec{u}\neq\vec{0}$, then it spans a line in $\mathbb{R}^2$, so for any $\vec{v}$ not on that line, the set $\{\vec{u},\vec{v}\}$ will be linearly independent.

1.5.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.6: Subspaces

Objectives

1. Learn the definition of a subspace.
2. Learn to determine whether or not a subset is a subspace.
3. Learn the most important examples of subspaces.
4. Learn to write a given subspace as a column space or null space.
5. *Recipe:* compute a spanning set for a null space.
6. *Picture:* whether a subset of $\mathbb{R}^2$ or $\mathbb{R}^3$ is a subspace or not.
7. *Vocabulary words:* ***subspace***, ***column space***, ***null space***.

In this section we discuss *subspaces* of $\mathbb{R}^n$. A subspace turns out to be exactly the same thing as a span, except we don't have a particular set of spanning vectors in mind. This change in perspective is quite useful, as it is easy to produce subspaces that are not obviously spans. For example, the solution set of the equation $x + 3y + z = 0$ is a span because the equation is homogeneous, but we would have to compute the parametric vector form in order to write it as a span.

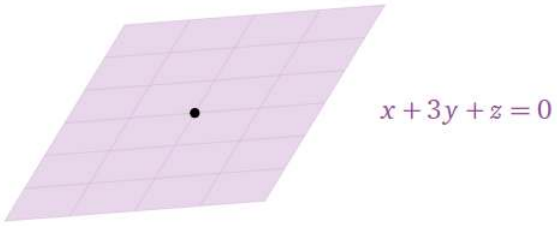


Figure **1.6.1**

(Spoiler: A subspace turns out to be the same thing as the solution set of a homogeneous system of equations.)

This page titled 1.6: Subspaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **2.6: Subspaces** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.6.1: Defining Subspaces

Definition: Subset

A ***subset*** of $\mathbb{R}^n$ is any collection of points in $\mathbb{R}^n$.

For instance, the unit circle

$$C = \{(x, y) \text{ in } \mathbb{R}^2 : x^2 + y^2 = 1\}$$

is a subset of $\mathbb{R}^2$.

Above, we expressed C in set builder notation: in English, it reads "C is the set of all ordered pairs (x, y) in $\mathbb{R}^2$ such that $x^2 + y^2 = 1$."

Definition: Subspace

A ***subspace*** of $\mathbb{R}^n$ is a subset V of $\mathbb{R}^n$ satisfying:

1. ***Non-emptiness:*** The zero vector is in V.
2. ***Closure under addition:*** If $\vec{u}$ and $\vec{v}$ are in V, then $\vec{u} + \vec{v}$ is also in V.
3. ***Closure under scalar multiplication:*** If $\vec{v}$ is in V and c is in $\mathbb{R}$, then $c\vec{v}$ is also in V.

As a consequence of these properties, we see:

- If $\vec{v}$ is a vector in V, then all scalar multiples of $\vec{v}$ are in V by the third property. In other words, the line through any nonzero vector in V is also contained in V.
- If $\vec{u}, \vec{v}$ are vectors in V and c, d are scalars, then $c\vec{u}, d\vec{v}$ are also in V by the third property, so $c\vec{u} + d\vec{v}$ is in V by the second property. Therefore, all of $\text{Span}\{\vec{u}, \vec{v}\}$ is contained in V.
- Similarly, if $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ are all in V, then $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is contained in V. In other words, *a subspace contains the span of any vectors in it.*

If you choose enough vectors, then eventually their span will fill up V, so we already see that a *subspace is a span*. We will shortly make this a precise statement.

Remark

Suppose that V is a non-empty subset of $\mathbb{R}^n$ that satisfies properties 2 and 3. Let v be any vector in V. Then $0\vec{v} = \vec{0}$ is in V by the third property, so V automatically satisfies property 1. It follows that the only subset of $\mathbb{R}^n$ that satisfies properties 2 and 3 but *not* property 1 is the empty subset $\{\}$. This is why we call the first property "non-emptiness".

Illustration 1.6.1.1

The set $\mathbb{R}^n$ is a subspace of itself: indeed, it contains zero, and is closed under addition and scalar multiplication.

Illustration 1.6.1.2

The set $\{\vec{0}\}$ containing only the zero vector is a subspace of $\mathbb{R}^n$: it contains zero, and if you add zero to itself or multiply it by a scalar, you always get zero.

Illustration 1.6.1.3: A line through the origin

A line L through the origin is a subspace.

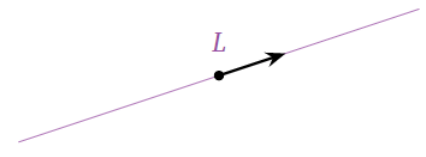


Figure 1.6.1.2

Indeed, L contains zero, and is easily seen to be closed under addition and scalar multiplication.

Illustration 1.6.1.4: A plane through the origin

A plane P through the origin is a subspace.

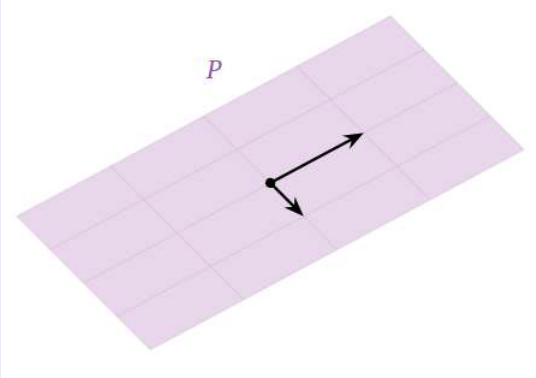


Figure 1.6.1.3

Indeed, P contains zero; the sum of two vectors in P is also in P; and any scalar multiple of a vector in P is also in P.

Illustration 1.6.1.5: Non-example (A line not containing the origin)

A line L (or any other subset) that does not contain the origin is not a subspace. It fails the first defining property: *every subspace contains the origin* by definition.

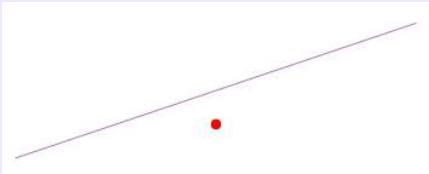


Figure 1.6.1.4

Illustration 1.6.1.6: Non-example (A circle)

The unit circle C is not a subspace. It fails all three defining properties: it does not contain the origin, it is not closed under addition, and it is not closed under scalar multiplication. In the picture, one red vector is the sum of the two black vectors (which are contained in C), and the other is a scalar multiple of a black vector.

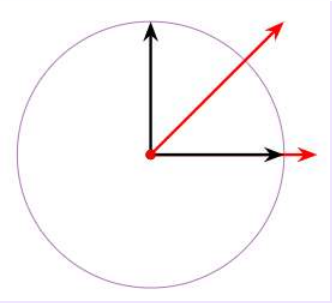


Figure 1.6.1.5

Illustration 1.6.1.7: Non-example (The first quadrant)

The first quadrant in $\mathbb{R}^2$ is not a subspace. It contains the origin and is closed under addition, but it is not closed under scalar multiplication (by negative numbers).

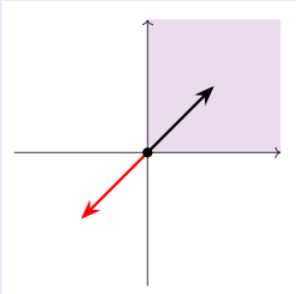


Figure 1.6.1.6

Illustration 1.6.1.8: Non-example (A line union a plane)

The union of a line and a plane in $\mathbb{R}^3$ is not a subspace. It contains the origin and is closed under scalar multiplication, but it is not closed under addition: the sum of a vector on the line and a vector on the plane is not contained in the line or in the plane.

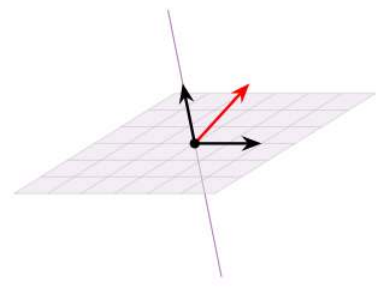


Figure 1.6.1.7

Note 1.6.1.1: Subsets versus Subspaces

A sub*set* of $\mathbb{R}^n$ is any collection of vectors whatsoever. For instance, the unit circle

$$C = \{(x, y) \text{ in } \mathbb{R}^2 : x^2 + y^2 = 1\}$$

is a subset of $\mathbb{R}^2$, but it is not a subspace. In fact, all of the non-examples above are still subsets of $\mathbb{R}^n$. A subspace is a subset that happens to satisfy the three additional defining properties.

In order to verify that a subset of $\mathbb{R}^n$ is in fact a subspace, one has to check the three defining properties. That is, unless the subset has already been verified to be a subspace: see this important note, Note **1.6.1.3**, below.

Example 1.6.1.1: Verifying that a subset is a subspace

Let

$$V = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \text{ in } \mathbb{R}^2 \, : \, 2a = 3b \right\}$$

Verify that V is a subspace.

Solution

First, we point out that the condition "$2a = 3b$" defines whether or not a vector is in V: that is, to say $\begin{bmatrix} a \\ b \end{bmatrix}$ is in V *means* that $2a = 3b$. In other words, a vector is in V if twice its first coordinate equals three times its second coordinate. So for instance, $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{3} \end{bmatrix}$ are in V, but $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ is not because $2 \cdot 2 \neq 3 \cdot 3$.

Let us check the first property. The subset V does contain the zero vector $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, because $2 \cdot 0 = 3 \cdot 0$

Next we check the second property. To show that V is closed under addition, we have to check that for *any* vectors $u = \begin{bmatrix} a \\ b \end{bmatrix}$ and $v = \begin{bmatrix} c \\ d \end{bmatrix}$ in V, the sum $\vec{u} + \vec{v}$ is in V. Since we cannot assume anything else about $\vec{u}$ and $\vec{v}$, we must treat them as *unknowns*.

We have

$$\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} a+c \\ b+d \end{bmatrix}$$

To say that $\begin{bmatrix} a+c \\ b+d \end{bmatrix}$ is contained in V means that $2(a+c) = 3(b+d)$, or $2a + 2c = 3b + 3d$. The one thing we are allowed to assume about $\vec{u}$ and $\vec{v}$ is that $2a = 3b$ and $2c = 3d$, so we see that $\vec{u} + \vec{v}$ is indeed contained in V.

Next we check the third property. To show that V is closed under scalar multiplication, we have to check that for any vector $v = \begin{bmatrix} a \\ b \end{bmatrix}$ in V and any scalar c in $\mathbb{R}$, the product $c\vec{v}$ is in V. Again, we must treat $\vec{v}$ and c as unknowns. We have

$$c \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ca \\ cb \end{bmatrix}$$

To say that $\begin{bmatrix} ca \\ cb \end{bmatrix}$ is contained in V means that $2(ca) = 3(cb)$, i.e., that $c \cdot 2a = c \cdot 3b$. The one thing we are allowed to assume about $\vec{v}$ is that $2a = 3b$, so $c\vec{v}$ is indeed contained in V.

Since V satisfies all three defining properties, it is a subspace

In fact, it is the line through the origin with slope $\frac{2}{3}$, pictured below.

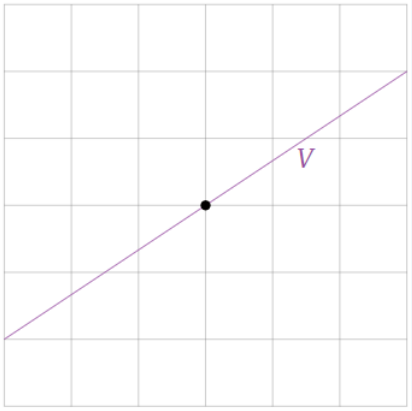


Figure 1.6.1.8

Example 1.6.1.2: Showing that a subset is not a subspace

Let

$$V = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \text{ in } \mathbb{R}^2 \,:\, ab = 0 \right\}$$

Is V a subspace?

Solution

First, we point out that the condition "$ab = 0$" defines whether or not a vector is in V: that is, to say $\begin{bmatrix} a \\ b \end{bmatrix}$ is in V *means* that $ab = 0$. In other words, a vector is in V if the product of its coordinates is zero, i.e., if one (or both) of its coordinates are zero. So for instance, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ are in V, but $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is not because $1 \cdot 1 \neq 0$.

Let us check the first property. The subset V does contain the zero vector $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, because $0 \cdot 0 = 0$.

Next we check the third property. To show that V is closed under scalar multiplication, we have to check that for *any* vector $\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ in V and *any* scalar c in $\mathbb{R}$, the product $c\vec{v}$ is in V. Since we cannot assume anything else about $\vec{v}$ and c, we must treat them as *unknowns*.

We have

$$c\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ca \\ cb \end{bmatrix}$$

To say that $\begin{bmatrix} ca \\ cb \end{bmatrix}$ is contained in V means that $(ca)(cb) = 0$. Rewriting, this means $c^2(ab) = 0$. The one thing we are allowed to assume about $\vec{v}$ is that $ab = 0$, so we see that $c\vec{v}$ is indeed contained in V.

Next we check the second property. It turns out that V is not closed under addition; to verify this, we must show that there exists *some* vectors $\vec{u}, \vec{v}$ in V such that $\vec{u} + \vec{v}$ is not contained in V. The easiest way to do so is to produce examples of such vectors. We can take $u = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$; these are contained in V because the products of their coordinates are zero, but

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

is not contained in V because $1 \cdot 1 \neq 0$.

Since V does not satisfy the second property (it is not closed under addition), we conclude that $\boxed{V \text{ is not a subspace}}$

Indeed, it is the union of the two coordinate axes, which is not a span.

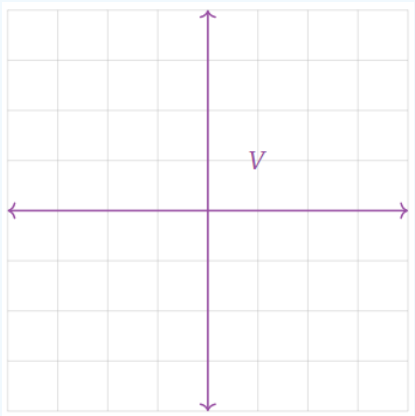


Figure 1.6.1.9

1.6.1: Defining Subspaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.6: Subspaces** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.6: Common Subspaces

Theorem 1.6.1: Spans are Subspaces and Subspaces are Spans

If $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$ are any vectors in $\mathbb{R}^n$, then $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ is a subspace of $\mathbb{R}^n$. Moreover, any subspace of $\mathbb{R}^n$ can be written as a span of a set of p linearly independent vectors in $\mathbb{R}^n$ for $p \leq n$.

Proof

To show that $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ is a subspace, we have to verify the three defining properties.

1. The zero vector $\vec{0} = 0\vec{v}_1 + 0\vec{v}_2 + \cdots + 0\vec{v}_p$ is in the span.
2. If $\vec{u} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_p\vec{v}_p$ and $\vec{v} = b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_p\vec{v}_p$ are in $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$, then
$$\vec{u} + \vec{v} = (a_1 + b_1)\vec{v}_1 + (a_2 + b_2)\vec{v}_2 + \cdots + (a_p + b_p)\vec{v}_p$$
is also in $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$.
3. If $\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_p\vec{v}_p$ is in $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ and c is a scalar, then
$$c\vec{v} = ca_1\vec{v}_1 + ca_2\vec{v}_2 + \cdots + ca_p\vec{v}_p$$
is also in $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$.

Since $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ satisfies the three defining properties of a subspace, it is a subspace.

Now let V be a subspace of $\mathbb{R}^n$. If V is the zero subspace, then it is the span of the empty set, so we may assume V is nonzero. Choose a nonzero vector $\vec{v}_1$ in V. If $V = \text{Span}\{\vec{v}_1\}$, then we are done. Otherwise, there exists a vector $\vec{v}_2$ that is in V but not in $\text{Span}\{\vec{v}_1\}$. Then $\text{Span}\{\vec{v}_1, \vec{v}_2\}$ is contained in V, and by the increasing span criterion in the previous section, the set $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent. If $V = \text{Span}\{\vec{v}_1, \vec{v}_2\}$ then we are done. Otherwise, we continue in this fashion until we have written $V = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ for some linearly independent set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$. This process terminates after at most n steps, since we know there cannot be more than n linearly independent vectors in $\mathbb{R}^n$.

If $V = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$, we say that V is the subspace ***spanned by*** or ***generated by*** the vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$. We call $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p\}$ a ***spanning set*** for V.

Any matrix naturally gives rise to *two* subspaces.

Definition: Column Space and Null Space

Let A be an $m \times n$ matrix.

- The ***column space*** of A is the subspace of $\mathbb{R}^m$ spanned by the columns of A. It is written $\text{Col}(A)$.
- The ***null space*** of A is the subspace of $\mathbb{R}^n$ consisting of all solutions of the homogeneous equation $A\vec{x} = \vec{0}$:
$$\text{Null}(A) = \{\vec{x} \text{ in } \mathbb{R}^n : A\vec{x} = \vec{0}\}$$

The column space is defined to be a span, so it is a subspace by the above Theorem, Theorem 1.6.1. We need to verify that the null space is really a subspace. We have already seen that the set of solutions of $A\vec{x} = \vec{0}$ is always a span, so the fact that the null space is a subspace should not come as a surprise.

Proof

We have to verify the three defining properties. To say that a vector $\vec{v}$ is in $\text{Null}(A)$ *means* that $A\vec{v} = \vec{0}$.

1. The zero vector is in $\text{Null}(A)$ because $A\vec{0} = \vec{0}$.
2. Suppose that $\vec{u}, \vec{v}$ are in $\text{Null}(A)$. This means that $A\vec{u} = \vec{0}$ and $A\vec{v} = \vec{0}$. Hence,

$$A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v} = \vec{0} + \vec{0} = \vec{0}$$

Therefore, $\vec{u}+\vec{v}$ is in $\text{Null}(A)$.

3. Suppose that $\vec{v}$ is in $\text{Null}(A)$ and c is a scalar. Then

$$A(c\vec{v}) = cA\vec{v} = c \cdot \vec{0} = \vec{0}$$

Therefore, $c\vec{v}$ is also in $\text{Null}(A)$.

Since $\text{Null}(A)$ satisfies the three defining properties of a subspace, it is a subspace.

Example 1.6.3

Describe the column space and the null space of

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Solution

The column space is the span of the columns of A:

$$\boxed{\text{Col}(A) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}}$$

This is a line in $\mathbb{R}^3$.

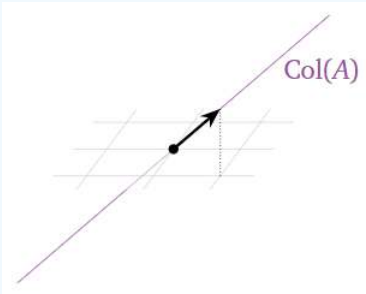


Figure 1.6.10

The null space is the solution set of the homogeneous system $A\vec{x} = \vec{0}$. To compute this, we need to row reduce A. Its reduced row echelon form is

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

This gives the equation $x + y = 0$, or

$$\begin{cases} x = -y \\ y = y \end{cases} \xrightarrow{\text{parametric vector form}} \begin{bmatrix} x \\ y \end{bmatrix} = y\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Hence, the null space is $\boxed{\text{Null}(A) = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}}$, which is a line in $\mathbb{R}^2$.

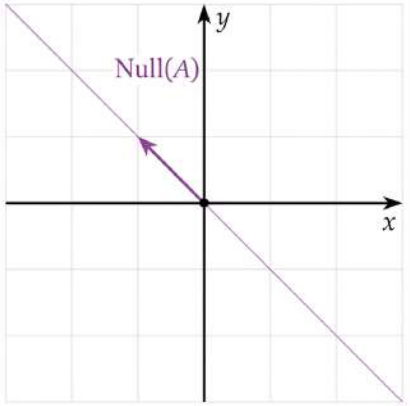


Figure 1.6.11

Notice that the column space is a subspace of $\mathbb{R}^3$, whereas the null space is a subspace of $\mathbb{R}^2$. This is because A has three rows and two columns.

The column space and the null space of a matrix are both subspaces, so they are both spans. The column space of a matrix A is defined to be the span of the columns of A. The null space is defined to be the solution set of $A\vec{x} = \vec{0}$, so this is a good example of a kind of subspace that we can define *without any spanning set in mind.* In other words, it is easier to show that the null space is a subspace than to show it is a span—see the proof above. In order to do computations, however, it is usually necessary to find a spanning set.

Note 1.6.2: Null Spaces are Solution Sets

The null space of a matrix is the solution set of a homogeneous system of equations. For example, the null space of the matrix

$$A = \begin{bmatrix} 1 & 7 & 2 \\ -2 & 1 & 3 \\ 4 & -2 & -3 \end{bmatrix}$$

is the solution set of $A\vec{x} = \vec{0}$, i.e., the solution set of the system of equations

$$\begin{cases} x + 7y + 2z = 0 \\ -2x + y + 3z = 0 \\ 4x - 2y - 3z = 0 \end{cases}$$

Conversely, the solution set of any homogeneous system of equations is precisely the null space of the corresponding coefficient matrix.

To find a spanning set for the null space, one has to solve a system of homogeneous equations.

Recipe: Compute a Spanning Set for Null Space

To find a spanning set for $\text{Null}(A)$, compute the parametric vector form of the solutions to the homogeneous equation $A\vec{x} = \vec{0}$. The vectors attached to the free variables form a spanning set for $\text{Null}(A)$.

Example 1.6.4: Two free variables

Find a spanning set for the null space of the matrix

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

Solution

We compute the parametric vector form of the solutions of $A\vec{x} = \vec{0}$. The reduced row echelon form of A is

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

The free variables are x_3 and x_4; the parametric form of the solution set is

$$\begin{cases} x_1 = x_3 - 2x_4 \\ x_2 = 2x_3 + 3x_4 \\ x_3 = x_3 \\ x_4 = x_4 \end{cases} \xrightarrow[\text{vector form}]{\text{parametric}} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

Therefore,

$$\boxed{\text{Null}(A) = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ 0 \\ 1 \end{bmatrix} \right\}}$$

Example 1.6.5: No free variables

Find a spanning set for the null space of the matrix

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

Solution

We compute the parametric vector form of the solutions of $A\vec{x} = \vec{0}$. The reduced row echelon form of A is

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

There are no free variables; hence, the only solution of $A\vec{x} = \vec{0}$ is the trivial solution. In other words,

$$\boxed{\text{Null}(A) = \{\vec{0}\} = \text{Span}\{\vec{0}\}}$$

It is natural to define $\text{Span}\{\} = \{\vec{0}\}$, so that we can take our spanning set to be empty. This is consistent with the definition of dimension that we will refine in the future!

A subspace can be given to you in many different forms. In practice, computations involving subspaces are much easier if your subspace is the column space or null space of a matrix. The simplest example of such a computation is finding a spanning set: a column space is by definition the span of the columns of a matrix, and we showed above how to compute a spanning set for a null space using parametric vector form. For this reason, it is useful to rewrite a subspace as a column space or a null space before trying to answer questions about it.

Note 1.6.3

When asking questions about a subspace, it is usually best to rewrite the subspace as a column space or a null space.

This also applies to the question "is my subset a subspace?" If your subset is a column space or null space of a matrix, then the answer is yes.

Illustration 1.6.9

Let

$$V = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \text{ in } \mathbb{R}^2 \,:\, 2a = 3b \right\}$$

be the subset of a previous Example 1.6.1. The subset V is exactly the solution set of the homogeneous equation $2x - 3y = 0$. Therefore,

$$V = \text{Null}\left(\begin{bmatrix} 2 & -3 \end{bmatrix}\right)$$

Note that it could also be the null space of other matrices, for example: $V = \text{Null}\left(\begin{bmatrix} 2 & -3 \\ -2 & 3 \\ 4\pi & -6\pi \end{bmatrix}\right)$

In particular, it is a subspace. The reduced row echelon form of $\begin{bmatrix} 2 & -3 \end{bmatrix}$ is $\begin{bmatrix} 1 & -\frac{3}{2} \end{bmatrix}$, so the parametric form of V is $x = \frac{3}{2}y$, so the parametric vector form is $\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} \frac{3}{2} \\ 1 \end{bmatrix}$, and hence, $\left\{ \begin{bmatrix} \frac{3}{2} \\ 1 \end{bmatrix} \right\}$ spans V.

Example 1.6.6

Let V be the plane in $\mathbb{R}^3$ defined by

$$V = \left\{ \begin{bmatrix} 2x + y \\ x - y \\ 3x - 2y \end{bmatrix} \,:\, x, y \text{ are in } \mathbb{R} \right\}$$

Write V as the column space of a matrix.

Solution

Since

$$\begin{bmatrix} 2x + y \\ x - y \\ 3x - 2y \end{bmatrix} = x \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + y \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$$

we notice that V is exactly the span of the vectors

$$\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$$

Hence,

$$\boxed{V = \text{Col}\left(\begin{bmatrix} 2 & 1 \\ 1 & -1 \\ 3 & -2 \end{bmatrix}\right)}$$

1.6: Common Subspaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.6: Subspaces** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.6.3: Practice Set

? Exercise 1.6.3.1a

Determine if the set $A = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : ab = 0 \right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix} 0 \\ 0 \end{bmatrix} \in A$:

Since $a = 0, b = 0 \implies ab = 0 \cdot 0 = 0$ we know $\vec{0} \in A$.

This set contains the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \in A \implies \begin{bmatrix} a+\alpha \\ b+\beta \end{bmatrix} \in A$:

We assume $ab = 0$ and $\alpha\beta = 0$, and consider $(a+\alpha)(b+\beta) = ab + \alpha b + \beta a + \alpha\beta = \alpha b + \beta a$

If we pick $a = 1, b = 0, \alpha = 0, \beta = 1$ this will be nonzero, so we can say that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in A$ but their sum $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin A$.

This set is not closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix} a \\ b \end{bmatrix} \in A, r \in \mathbb{R} \implies \begin{bmatrix} ra \\ rb \end{bmatrix} \in A$:

We assume $ab = 0$ and consider $(ra)(rb) = r^2ab = r^2 \cdot 0 = 0$. Therefore we know $\begin{bmatrix} ra \\ rb \end{bmatrix} \in A$.

This set is closed under scalar multiplication.

The set A is **not a subspace.**

? Exercise 1.6.3.1b

Determine if the set $B = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a^2 + b^2 = 0 \right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix} 0 \\ 0 \end{bmatrix} \in B$:

Since $a = 0, b = 0 \implies a^2 + b^2 = 0^2 + 0^2 = 0$ we know $\vec{0} \in B$.

This set contains the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \in B \implies \begin{bmatrix} a+\alpha \\ b+\beta \end{bmatrix} \in B$:

In this set, we know that if $a^2 + b^2 = 0$ then $a = b = 0$ must hold, so $\alpha = \beta = 0$ is also true. Therefore $a + \alpha = b + \beta = 0$ and the sum is in this set.

This set is closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix} a \\ b \end{bmatrix} \in B, r \in \mathbb{R} \implies \begin{bmatrix} ra \\ rb \end{bmatrix} \in B$:

We assume $a^2 + b^2 = 0$ and consider $(ra)^2 + (rb)^2 = r^2(a^2 + b^2) = r^2 \cdot 0 = 0$. Therefore we know $\begin{bmatrix} ra \\ rb \end{bmatrix} \in B$.

This set is closed under scalar multiplication.

The set B is **a subspace** (containing only the zero vector)**.**

Exercise 1.6.3.1c

Determine if the set $C = \left\{ \begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2 : a^2 + b^2 \leq 1 \right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix} 0 \\ 0 \end{bmatrix} \in C$:

Since $a = 0, b = 0 \implies a^2 + b^2 = 0^2 + 0^2 \leq 1$ we know $\vec{0} \in C$.

This set contains the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \in C \implies \begin{bmatrix} a+\alpha \\ b+\beta \end{bmatrix} \in C$:

We assume $a^2 + b^2 \leq 1$ and $\alpha^2 + \beta^2 \leq 1$, and consider $(a+\alpha)^2 + (b+\beta)^2 \leq 1$?

If we pick $a = \frac{1}{2}, b = \frac{1}{3}, \alpha = \frac{1}{2}, \beta = \frac{2}{3}$ this will break, so we can say that $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ \frac{2}{3} \end{bmatrix} \in C$ but their sum $\begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin C$.

This set is not closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix} a \\ b \end{bmatrix} \in C, r \in \mathbb{R} \implies \begin{bmatrix} ra \\ rb \end{bmatrix} \in C$:

We assume $a^2 + b^2 \leq 1$ and consider $(ra)^2 + (rb)^2 = r^2(a^2 + b^2) \leq 1$? If we pick $a = 1, b = 0, r = 100$ this will break, so we can say that $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \in C$ but its multiple $\begin{bmatrix} 100 \\ 0 \end{bmatrix} \notin C$.

This set is not closed under scalar multiplication.

The set C is **not a subspace.**

Exercise 1.6.3.1d

Determine if the set $D = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3 : a + b + c = 1, a - c = 1 \right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in D$:

Since $a = 0, b = 0, c = 0 \implies a + b + c = 0 \neq 1, a - c = 0 \neq 1$ we know $\vec{0} \notin D$.

This set does not contain the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix} a \\ b \\ c \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \in D \implies \begin{bmatrix} a+\alpha \\ b+\beta \\ c+\gamma \end{bmatrix} \in D$:

We know that both $a + b + c = 1, a - c = 1$ and $\alpha + \beta + \gamma = 1, \alpha - \gamma = 1$ must hold. Notice that:

$(a+\alpha) + (b+\beta) + (c+\gamma) = (a+b+c) + (\alpha+\beta+\gamma) = 1 + 1 = 2 \neq 1$

Similarly $(a+\alpha) - (c+\gamma) = 2 \neq 1$ so the sum of any two vectors in this set is never inside of this set.

This set is not closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in D, r \in \mathbb{R} \implies \begin{bmatrix} ra \\ rb \\ rc \end{bmatrix} \in D$:

We assume $a+b+c=1, a-c=1$ and consider $(ra)+r(b)+r(c)=r(a+b+c)=r\cdot 1=r\neq 1$. Therefore we know $\begin{bmatrix} ra \\ rb \\ rc \end{bmatrix} \notin D$ for any $r\neq 1$.

This set is not closed under scalar multiplication.

The set D is **not a subspace.**

Exercise 1.6.3.1e

Determine if the set $E=\left\{\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 : c=\pi\right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \in E$:

Since $c=0 \implies c\neq\pi$ we know $\vec{0}\notin E$.

This set does not contain the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} \in E \implies \begin{bmatrix} a+\alpha \\ b+\beta \\ c+\gamma \\ d+\delta \end{bmatrix} \in E$:

We know that both $c=\pi$ and $\gamma=\pi$ must hold. Notice that:

$c+\gamma=\pi+\pi=2\pi\neq\pi$ so the sum of any two vectors in this set is never inside of this set.

This set is not closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in E, r\in\mathbb{R} \implies \begin{bmatrix} ra \\ rb \\ rc \\ rd \end{bmatrix} \in E$:

We assume $c=\pi$ and consider $rc=r\pi\neq\pi$. Therefore we know $\begin{bmatrix} ra \\ rb \\ rc \\ rd \end{bmatrix} \notin E$ for any $r\neq 1$.

This set is not closed under scalar multiplication.

The set E is **not a subspace.**

Exercise 1.6.3.1f

Determine if the set $F=\left\{\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 : \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \in F$:

Since $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 5 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 6 \end{bmatrix} + 0\begin{bmatrix} 3 \\ 7 \end{bmatrix} + 0\begin{bmatrix} 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ we know $\vec{0} \in F$.

This set contains the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} \in F \implies \begin{bmatrix} a+\alpha \\ b+\beta \\ c+\gamma \\ d+\delta \end{bmatrix} \in F$:

We assume $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

It is easy to see $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} a+\alpha \\ b+\beta \\ c+\gamma \\ d+\delta \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} + \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

This set is closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in F, r \in \mathbb{R} \implies \begin{bmatrix} ra \\ rb \\ rc \\ rd \end{bmatrix} \in F$:

We assume $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

And see that $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} ra \\ rb \\ rc \\ rd \end{bmatrix} = r\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = r\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

This set is closed under scalar multiplication.

The set F is **a subspace.**

? Exercise 1.6.3.1*g*

Determine if the set $G = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 : \begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = 0 \right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \in G$:

Since $\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = 0 \cdot 1 + 0 \cdot 4 + 0 \cdot 5 + 0 \cdot 7 = 0$ we know $\vec{0} \in G$.

This set contains the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} \in G \implies \begin{bmatrix} a+\alpha \\ b+\beta \\ c+\gamma \\ d+\delta \end{bmatrix} \in G$:

We assume $\begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = 0$ and $\begin{bmatrix} \alpha & \beta & \gamma & \delta \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = 0$

It is easy to see $\begin{bmatrix} (a+\alpha) & (b+\beta) & (c+\gamma) & (d+\delta) \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} + \begin{bmatrix} \alpha & \beta & \gamma & \delta \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = 0$

This set is closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in G, r \in \mathbb{R} \implies \begin{bmatrix} ra \\ rb \\ rc \\ rd \end{bmatrix} \in G$:

We assume $\begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = 0$

And see that $\begin{bmatrix} ra & rb & rc & rd \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = r \begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = r \cdot 0 = 0$

This set is closed under scalar multiplication.

The set G is **a subspace.**

? Exercise 1.6.3.1h

Determine if the set $H = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3 : ab = c \right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in H$:

Since $a = 0, b = 0, c = 0 \implies ab = 0 = c$ we know $\vec{0} \in H$.

This set contains the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix} a \\ b \\ c \end{bmatrix}, \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \in H \implies \begin{bmatrix} a+\alpha \\ b+\beta \\ c+\gamma \end{bmatrix} \in H$:

We assume $ab = c$ and $\alpha\beta = \gamma$, and consider $(a+\alpha)(b+\beta) = ab + \alpha b + \beta a + \alpha\beta = c + \gamma$?

If we pick $a=1, b=0, c=0, \alpha=0, \beta=1, \gamma=0$ this will fail, so we can say that $\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix} \in H$ but their sum $\begin{bmatrix}1\\1\\0\end{bmatrix} \notin H$.

This set is not closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix}a\\b\\c\end{bmatrix} \in H, r \in \mathbb{R} \implies \begin{bmatrix}ra\\rb\\rc\end{bmatrix} \in H$:

We assume $ab = c$ and consider $(ra)(rb) = r^2ab = r^2c \neq rc$ for any $r \neq 1$.

This set is not closed under scalar multiplication.

The set H is **not a subspace.**

Exercise 1.6.3.1*i*

Determine if the set $I = \left\{ \begin{bmatrix}2a+b\\b+2c\\a-c\end{bmatrix} \in \mathbb{R}^3 : a, b, c \in \mathbb{R} \right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix}0\\0\\0\end{bmatrix} \in I$:

When $a=b=c=0$ we get $\begin{bmatrix}2\cdot 0+0\\0+2\cdot 0\\0-0\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$ we know $\vec{0} \in I$.

This set contains the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix}2a+b\\b+2c\\a-c\end{bmatrix}, \begin{bmatrix}2\alpha+\beta\\\beta+2\gamma\\\alpha-\gamma\end{bmatrix} \in I \implies \begin{bmatrix}2a+b+2\alpha+\beta\\b+2c+\beta+2\gamma\\a-c+\alpha-\gamma\end{bmatrix} \in I$:

Note that $\begin{bmatrix}2a+b+2\alpha+\beta\\b+2c+\beta+2\gamma\\a-c+\alpha-\gamma\end{bmatrix} = \begin{bmatrix}2(a+\alpha)+(b+\beta)\\(b+\beta)+2(c+\gamma)\\(a+\alpha)-(c+\gamma)\end{bmatrix}$

This set is closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix}2a+b\\b+2c\\a-c\end{bmatrix} \in I, r \in \mathbb{R} \implies \begin{bmatrix}r(2a+b)\\r(b+2c)\\r(a-c)\end{bmatrix} \in I$:

Note that $\begin{bmatrix}r(2a+b)\\r(b+2c)\\r(a-c)\end{bmatrix} = \begin{bmatrix}2(ra)+(rb)\\(rb)+2(rc)\\(ra)-(rc)\end{bmatrix}$

This set is closed under scalar multiplication.

The set I is **a subspace.**

Exercise 1.6.3.1*j*

Determine if the set $J = \left\{ \begin{bmatrix}a+1\\b+2\\c+3\end{bmatrix} \in \mathbb{R}^3 : a, b, c \in \mathbb{R} \right\}$ is a subspace.

Answer

1. *Contains the zero vector?* We check if $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in J$:

When $a=-1, b=-2, c=-3$ we get $\begin{bmatrix} -1+1 \\ -2+2 \\ -3+3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ we know $\vec{0} \in J$.

This set contains the zero vector.

2. *Closed under addition?* We check if $\begin{bmatrix} a+1 \\ b+2 \\ c+3 \end{bmatrix}, \begin{bmatrix} \alpha+1 \\ \beta+2 \\ \gamma+3 \end{bmatrix} \in J \implies \begin{bmatrix} a+1+\alpha+1 \\ b+2+\beta+2 \\ c+3+\gamma+3 \end{bmatrix} \in J$:

Note that $\begin{bmatrix} a+1+\alpha+1 \\ b+2+\beta+2 \\ c+3+\gamma+3 \end{bmatrix} = \begin{bmatrix} (a+1+\alpha)+1 \\ (b+2+\beta)+2 \\ (c+3+\gamma)+3 \end{bmatrix}$

This set is closed under addition.

3. *Closed under scalar multiplication?* We check if $\begin{bmatrix} a+1 \\ b+2 \\ c+3 \end{bmatrix} \in J, r \in \mathbb{R} \implies \begin{bmatrix} r(a+1) \\ r(b+2) \\ r(c+3) \end{bmatrix} \in J$:

Note that $\begin{bmatrix} r(a+1) \\ r(b+2) \\ r(c+3) \end{bmatrix} = \begin{bmatrix} (ra+r-1)+1 \\ (rb+2r-2)+2 \\ (rc+3r-3)+3 \end{bmatrix}$

This set is closed under scalar multiplication.

The set J is **a subspace.**

? Exercise 1.6.3.2*a*

Find the (i) Null Space and (ii) Column Space of $A = \begin{bmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & -1 & 0 \\ 4 & 1 & 1 & -2 \end{bmatrix}$

Answer

(i) $\begin{bmatrix} 1 & 0 & 1 & -1 \\ 2 & 1 & -1 & 0 \\ 4 & 1 & 1 & -2 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

So our homogeneous solutions satisfy $\begin{cases} x_1 + x_3 - x_4 = 0 \\ x_2 - 3x_3 + 2x_4 = 0 \\ x_3, x_4 \text{ free} \end{cases} \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ 3 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ -2 \\ 0 \\ 1 \end{bmatrix}$

Therefore $\mathbf{Null}(A) = \mathbf{Span}\left\{ \begin{bmatrix} -1 \\ 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 0 \\ 1 \end{bmatrix} \right\}$

(ii) We can easily see that $\mathbf{Col}(A) = \mathbf{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix} \right\}$

? Exercise 1.6.3.2*b*

Find the (i) Null Space and (ii) Column Space of $B = \begin{bmatrix} 6 & 3 \\ 4 & 2 \\ -8 & -4 \\ 0 & 0 \\ 10 & 5 \end{bmatrix}$

Answer

(i) $\begin{bmatrix} 6 & 3 \\ 4 & 2 \\ -8 & -4 \\ 0 & 0 \\ 10 & 5 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

So our homogeneous solutions satisfy $\begin{cases} x_1 + \frac{1}{2}x_2 = 0 \\ x_2 \text{ free} \end{cases} \implies \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -\frac{1}{2} \\ 1 \end{bmatrix}$

Therefore $\mathbf{Null}(B) = \mathbf{Span}\left\{\begin{bmatrix} -\frac{1}{2} \\ 1 \end{bmatrix}\right\}$

(ii) We can easily see that $\mathbf{Col}(B) = \mathbf{Span}\left\{\begin{bmatrix} 6 \\ 4 \\ -8 \\ 0 \\ 10 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ -4 \\ 0 \\ 5 \end{bmatrix}\right\}$

? Exercise 1.6.3.2*c*

Find the (i) Null Space and (ii) Column Space of $C = \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 2 \\ 1 & 2 & 1 & 2 & 0 & -1 \\ 3 & 6 & 4 & 7 & 0 & 1 \end{bmatrix}$

Answer

(i) $\begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 2 \\ 1 & 2 & 1 & 2 & 0 & -1 \\ 3 & 6 & 4 & 7 & 0 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

So our homogeneous solutions satisfy $\begin{cases} x_1 + 2x_2 + x_4 = 0 \\ x_3 + x_4 = 0 \\ x_6 = 0 \\ x_2, x_4, x_5 \text{ free} \end{cases} \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$

Therefore $\mathbf{Null}(C) = \mathbf{Span}\left\{\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}\right\}$

(ii) We can easily see that $\mathbf{Col}(C) = \mathbf{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 6 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}\right\}$

? Exercise 1.6.3.2*d*

Find the (i) Null Space and (ii) Column Space of $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 3 & 0 & 2 \\ -1 & 1 & 1 \end{bmatrix}$

Answer

(i) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 3 & 0 & 2 \\ -1 & 1 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

So our homogeneous solutions satisfy $\begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = 0 \end{cases} \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Therefore $\mathbf{Null}(D) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$

(ii) We can easily see that $\mathbf{Col}(D) = \mathbf{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix} \right\}$

? Exercise 1.6.3.2*e*

Find the (i) Null Space and (ii) Column Space of $E = \begin{bmatrix} 1 & 1 & 0 & -2 \\ 8 & 8 & 8 & 8 \\ 1 & 1 & 0 & -2 \\ 1 & 1 & 0 & -2 \end{bmatrix}$

Answer

(i) $\begin{bmatrix} 1 & 1 & 0 & -2 \\ 8 & 8 & 8 & 8 \\ 1 & 1 & 0 & -2 \\ 1 & 1 & 0 & -2 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

So our homogeneous solutions satisfy $\begin{cases} x_1 + x_2 - 2x_4 = 0 \\ x_3 + 3x_4 = 0 \\ x_2, x_4 \text{ free} \end{cases} \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 0 \\ -3 \\ 1 \end{bmatrix}$

Therefore $\mathbf{Null}(E) = \mathbf{Span} \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -3 \\ 1 \end{bmatrix} \right\}$

(ii) We can easily see that $\mathbf{Col}(E) = \mathbf{Span} \left\{ \begin{bmatrix} 1 \\ 8 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 8 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 8 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 8 \\ -2 \\ -2 \end{bmatrix} \right\}$

? Exercise 1.6.3.2*f*

Find the (i) Null Space and (ii) Column Space of $F = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix}$

Answer

(i) $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

So our homogeneous solutions satisfy $\begin{cases} x_1+2x_2+3x_3+4x_4+5x_5=0 \\ x_2,x_3,x_4,x_5 \text{ free} \end{cases} \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_2\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4\begin{bmatrix} -4 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5\begin{bmatrix} -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

Therefore $\text{Null}(F) = \text{Span}\left\{\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$

(ii) We can easily see that $\text{Col}(F) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 4 \\ 4 \\ 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \\ 5 \\ 5 \\ 5 \end{bmatrix}\right\}$

? Exercise 1.6.3.3*a*

Write $A = \left\{\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 : a+b-c=d, 2a=3c, 2d=-b\right\}$ as either the Column Space or Null Space of a matrix of your choosing.

Answer

$$\begin{cases} a+b-c=d \\ 2a=3c \\ 2d=-b \end{cases} \leftrightarrow \begin{cases} a+b-c-d=0 \\ 2a-3c=0 \\ b+2d=0 \end{cases}$$

Thus $A = \text{Null}\left(\begin{bmatrix} 1 & 1 & -1 & -1 \\ 2 & 0 & -3 & 0 \\ 0 & 1 & 0 & 2 \end{bmatrix}\right)$

? Exercise 1.6.3.3*b*

Write $B = \left\{\begin{bmatrix} a+b+c \\ a-b \\ b+c \end{bmatrix} \in \mathbb{R}^3 : a,b,c \in \mathbb{R}\right\}$ as either the Column Space or Null Space of a matrix of your choosing.

Answer

$$\begin{bmatrix} a+b+c \\ a-b \\ b+c \end{bmatrix} = a\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} + c\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Thus $B = \text{Col}\left(\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{bmatrix}\right)$

? Exercise 1.6.3.3c

Write $C = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3 : a = b = c \right\}$ as either the Column Space or Null Space of a matrix of your choosing.

Answer

Note that $a = b = c$ is equivalent to saying $a = c, b = c, c = c$, or equivalently $\begin{cases} a - c = 0 \\ b - c = 0 \end{cases}$

Thus, $C = \text{Null}\left(\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \right)$

? Exercise 1.6.3.3d

Write $D = \left\{ \vec{0} \in \mathbb{R}^5 \right\}$ as either the Column Space or Null Space of a matrix of your choosing.

Answer

The zero vector is equivalent to its own span, so any matrix with 5 rows of all zeros will work for a column space:

$$D = \text{Col}\left(\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \right)$$

? Exercise 1.6.3.3e

Write $E = \mathbb{R}^3$ as either the Column Space or Null Space of a matrix of your choosing.

Answer

To span all of $E = \mathbb{R}^3$, any three linearly independent vectors will do:

$$E = \text{Col}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right)$$

? Exercise 1.6.3.3f

Write $F = \left\{ \begin{bmatrix} a+b \\ b+c \\ c+d \\ a+d \end{bmatrix} \in \mathbb{R}^4 : a + b + c = 0 \right\}$ as either the Column Space or Null Space of a matrix of your choosing.

Answer

This one requires us to simplify F first: $a + b + c = 0 \implies c = -a - b$ so

$$F = \left\{ \begin{bmatrix} a+b \\ -a \\ -a-b+d \\ a+d \end{bmatrix} \in \mathbb{R}^4 : a, b, d \in \mathbb{R} \right\}$$

$$\begin{bmatrix} a+b \\ -a \\ -a-b+d \\ a+d \end{bmatrix} = a \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ -1 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ d \end{bmatrix}$$

Thus $F = \text{Col}\left(\begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 0 \\ -1 & -1 & 1 \\ 1 & 0 & 1 \end{bmatrix}\right)$

1.6.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

1.7: Basis and Dimension

Objectives

1. Understand the definition of a basis of a subspace and rank and nullity of a matrix.
2. Understand the basis and rank theorems.
3. *Recipes:* basis for a column space, basis for a null space, basis of a span.
4. *Picture:* basis of a subspace of $\mathbb{R}^2$ or $\mathbb{R}^3$.
5. *Theorems:* basis theorem, rank theorem.
6. *Essential vocabulary words:* ***basis***, ***dimension, rank***, ***nullity***.

We conclude our first chapter with a definition of a basis as an efficient spanning set of a subspace. We then use this to find relationships between the "size" of the null space for a given matrix (the size of a solution set to $A\vec{x} = \vec{b}$ when it exists), and the "number of different" vectors $\vec{b}$ which make such an equation consistent.

This page titled 1.7: Basis and Dimension is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **2.7: Basis and Dimension** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.7.1: Defining Basis

As we discussed earlier, a subspace is the same as a span, except we do not need to have a set of spanning vectors in mind. There are infinitely many choices of spanning sets for a nonzero subspace; to avoid redundancy, usually it is most convenient to choose a spanning set with the *minimal* number of vectors in it. This is the idea behind the notion of a basis.

Definition: Basis

Let V be a subspace of $\mathbb{R}^n$. A ***basis*** of V is a set of vectors $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ in V such that:

1. $V = \text{Span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$
2. The set $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ is linearly independent.

Recall that a set of vectors is *linearly independent* if and only if, when you remove any vector from the set, the span shrinks. In other words, if $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ is a basis of a subspace V, then no proper subset of $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ will span V: it is a *minimal* spanning set. Any subspace with a finite spanning set admits a basis.

Note 1.7.1.1

A nonzero subspace has *infinitely many* different bases, but they all contain the same number of vectors.

We leave it as an exercise to prove that any two bases have the same number of vectors; one might want to wait until after learning the invertible matrix theorem in the next chapter.

Definition: Dimension

Let V be a subspace of $\mathbb{R}^n$. The number of vectors in any basis of V is called the ***dimension*** of V, and is written $\dim(V)$.

Example 1.7.1.1: A basis of $\mathbb{R}^2$

Find a basis of $\mathbb{R}^2$.

Solution

We need to find two vectors in $\mathbb{R}^2$ that span $\mathbb{R}^2$ and are linearly independent. One such basis is $\boxed{\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\}}$

1. They span because any vector $\begin{bmatrix}a\\b\end{bmatrix}$ can be written as a linear combination of $\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}$:

$$\begin{bmatrix}a\\b\end{bmatrix} = a\begin{bmatrix}1\\0\end{bmatrix} + b\begin{bmatrix}0\\1\end{bmatrix}$$

2. They are linearly independent: if

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

then $x = y = 0$.

This shows that the plane $\mathbb{R}^2$ has dimension 2.

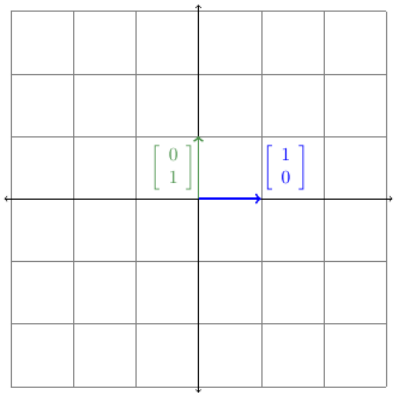


Figure 1.7.1.1

Illustration 1.7.1.1: All bases of $\mathbb{R}^2$

What are all possible bases of $\mathbb{R}^2$?

We know from the previous Example 1.7.1.1 that $\mathbb{R}^2$ has dimension 2, so any basis of $\mathbb{R}^2$ has two vectors in it. Let $\vec{v}_1, \vec{v}_2$ be vectors in $\mathbb{R}^2$, and let A be the matrix with columns $\vec{v}_1, \vec{v}_2$.

1. To say that $\{\vec{v}_1, \vec{v}_2\}$ spans $\mathbb{R}^2$ means that A has a pivot in every *row*.
2. To say that $\{\vec{v}_1, \vec{v}_2\}$ is linearly independent means that A has a pivot in every *column*.

Since A is a 2×2 matrix, it has a pivot in every row exactly when it has a pivot in every column.

Hence, **any two non-collinear vectors form a basis of $\mathbb{R}^2$**

For example, $\left\{ \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ is one such a basis.

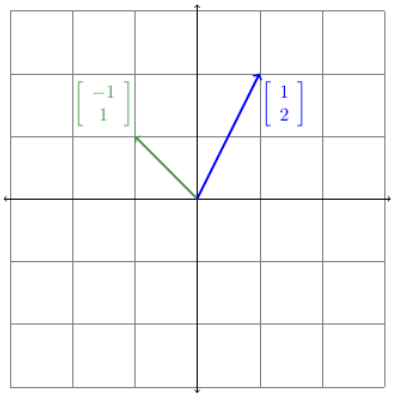


Figure 1.7.1.2

Illustration 1.7.1.2: The standard basis of $\mathbb{R}^n$

One shows exactly as in the above Illustration 1.7.1.1 that the standard coordinate vectors

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \cdots, \quad \vec{e}_{n-1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

form a basis for $\mathbb{R}^n$. This is sometimes known as the ***standard basis***.

In particular, $\mathbb{R}^n$ has dimension n.

Note 1.7.1.2

Illustration 1.7.1.1 implies that any basis for $\mathbb{R}^n$ has n vectors in it. Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be vectors in $\mathbb{R}^n$, and let A be the $n \times n$ matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$.

1. To say that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ spans $\mathbb{R}^n$ means that A has a pivot position in every *row*.
2. To say that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly independent means that A has a pivot position in every *column*.

Since A is a square matrix, it has a pivot in every row if and only if it has a pivot in every column. We will see in the future that the above two conditions are equivalent to the *invertibility* of the matrix A.

Illustration 1.7.1.3

Let

$$V = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \text{ in } \mathbb{R}^3 : x + 3y + z = 0 \right\} \quad \mathcal{B} = \left\{ \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

We will verify that V is a subspace, and show directly that $\mathcal{B}$ is a basis for V.

First, we observe that V is the solution set of the homogeneous equation $x + 3y + z = 0$, so it is a null space and therefore, a subspace.

To show that $\mathcal{B}$ is a basis, we really need to verify three things:

1. Both vectors are in V:

$$\begin{array}{rcrcrcl} (-3) & + & 3(1) & + & (0) & = & 0 \\ (-1) & + & 3(0) & + & (1) & = & 0 \end{array}$$

2. The vectors span V:

Since $V = \mathbf{Nul}\begin{pmatrix} \begin{bmatrix} 1 & 3 & 1 \end{bmatrix} \end{pmatrix}$, this is already in RREF so $x = -3y - z$ with y, z free variables.

Putting this into parametric vector form we see all elements of V look like

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

This clearly shows that $V = \mathbf{Span}\left\{ \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$ as desired.

3. The vectors are linearly independent:

$$c_1 \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \vec{0} \implies \begin{bmatrix} -3c_1 - c_2 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \implies c_1 = c_2 = 0$$

Alternatively, one can observe that the two vectors are not collinear.

This proves that these two vectors form a basis for V. Since V has a basis with two vectors, it has dimension two: it is a *plane*.

1.7.1: Defining Basis is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.7: Basis and Dimension** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.7: Finding a Basis of a Subspace

Now we show how to find bases for the column space of a matrix and the null space of a matrix. In order to find a basis for a given subspace, it is usually easiest to rewrite the subspace as a column space or a null space first.

First, we show how to compute a basis for the column space of a matrix.

Theorem 1.7.1

The pivot columns of a matrix A form a basis for $\mathrm{Col}(A)$.

Note 1.7.3

The above theorem is referring to the pivot columns in the *original* matrix, not its reduced row echelon form. Indeed, a matrix and its reduced row echelon form generally have different column spaces. For example, consider the matrix A below:

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ -2 & -3 & 4 & 5 \\ 2 & 4 & 0 & -2 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

pivot columns = basis ⟷ pivot columns in RREF

Figure 1.7.4

The pivot columns are the first two columns, so a basis for $\mathrm{Col}(A)$ is

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} \right\}$$

The first two columns of the reduced row echelon form certainly span a different subspace, as

$$\mathrm{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} : a, b \text{ in } \mathbb{R} \right\} = (xy\text{-plane with } z = 0)$$

but $\mathrm{Col}(A)$ contains vectors whose last coordinate is nonzero.

Corollary 1.7.1

The dimension of $\mathrm{Col}(A)$ is the number of pivots of A.

Computing a basis for a span is the same as computing a basis for a column space. Indeed, the span of finitely many vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ *is* the column space of a matrix, namely, the matrix A whose columns are $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$:

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix}$$

Example 1.7.2: A basis of a span

Find a basis of the subspace

$$V = \mathrm{Span}\left\{ \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ -2 \end{bmatrix} \right\}$$

Solution

The subspace V is the column space of the matrix

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

The reduced row echelon form of this matrix is

$$\begin{bmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The first two columns are pivot columns, so a basis for V is

$$\boxed{\left\{ \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} \right\}}$$

Example 1.7.3: Another basis of the same span

Find a basis of the subspace

$$V = \text{Span}\left\{ \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ -2 \end{bmatrix} \right\}$$

which does not consist of the first two vectors, as in the previous Example 1.7.6.

Solution

The point of this example is that the above Theorem 1.7.1 gives *one* basis for V; as always, there are infinitely more.

Reordering the vectors, we can express V as the column space of

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

The reduced row echelon form of this matrix is

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

The first two columns are pivot columns, so a basis for V is

$$\boxed{\left\{ \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 5 \\ -2 \end{bmatrix} \right\}}$$

These are the *last* two vectors in the given spanning set.

Of course, a cheaper way to answer this question is to scale our previous answer $\left\{ \begin{bmatrix} 10 \\ -20 \\ 20 \end{bmatrix}, \begin{bmatrix} 200 \\ -300 \\ 400 \end{bmatrix} \right\}$ to span the same subspace.

In order to compute a basis for the null space of a matrix, one has to find the parametric vector form of the solutions of the homogeneous equation $A\vec{x} = \vec{0}$.

Theorem 1.7.2

The vectors attached to the free variables in the parametric vector form of the solution set of $A\vec{x} = \vec{0}$ form a basis of $\text{Null}(A)$.

The proof of the theorem has two parts. The first part is that every solution lies in the span of the given vectors. This is automatic: the vectors are exactly chosen so that every solution is a linear combination of those vectors. The second part is that the vectors are linearly independent, which we have already discussed to be true.

As mentioned at the beginning of this subsection, when given a subspace written in a different form, in order to compute a basis it is usually best to rewrite it as a column space or null space of a matrix.

Example 1.7.4: A basis of a subspace

Let V be the subspace defined by

$$V = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + 2y = z \right\}$$

Find $\dim(V)$

Solution

First, we notice that V is exactly the solution set of the homogeneous linear equation $x + 2y - z = 0$. Hence, $V = \text{Null}\left(\begin{bmatrix} 1 & 2 & -1 \end{bmatrix}\right)$. This matrix is in reduced row echelon form; the parametric form of the general solution is $x = -2y + z$, so the parametric vector form is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = y \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

It follows that a basis for V is $\left\{ \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$

Since V has a basis with two vectors, we have $\boxed{\dim(V) = 2}$: it is a plane.

Recall that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ forms a basis for $\mathbb{R}^n$ if and only if the matrix A with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ has a pivot in every row and column (see this Example 1.7.4). Since A is an $n \times n$ matrix, these two conditions are equivalent: the vectors span if and only if they are linearly independent. The basis theorem is an abstract version of the preceding statement, that applies to any subspace.

Theorem 1.7.3: Basis Theorem

Let V be a subspace of dimension m. Then:

- Any m linearly independent vectors in V form a basis for V.
- Any m vectors that span V form a basis for V.

Proof

Suppose that $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ is a set of linearly independent vectors in V. In order to show that $\mathcal{B}$ is a basis for V, we must prove that $V = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$. If not, then there exists some vector $\vec{v}_{m+1}$ in V that is not contained in $\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$. By the increasing span criterion, the set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m, \vec{v}_{m+1}\}$ is also linearly independent. Continuing in this way, we keep choosing vectors until we eventually do have a linearly independent spanning set: say $V = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m, \ldots, \vec{v}_{m+k}\}$. Then $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{m+k}\}$ is a basis for V, which implies that $\dim(V) = m + k > m$. But we were assuming that V has dimension m, so $\mathcal{B}$ must have already been a basis.

Now suppose that $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ spans V. If $\mathcal{B}$ is not linearly independent, then we can remove some number of vectors from $\mathcal{B}$ without shrinking its span. After reordering, we can assume that we removed the last k vectors without

shrinking the span, and that we cannot remove any more. Now $V = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{m-k}\}$, and $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{m-k}\}$ is a basis for V because it is linearly independent. This implies that $\dim V = m - k < m$. But we were assuming that $\dim V = m$, so $\mathcal{B}$ must have already been a basis.

In other words, if you *already* know that $\dim(V) = m$, and if you have a set of m vectors $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ in V, then you only have to check *one* of:

1. $\mathcal{B}$ is linearly independent, *or*
2. $\mathcal{B}$ spans V

In order for $\mathcal{B}$ to be a basis of V. If you did not already know that $\dim(V) = m$, then you would have to check *both* properties.

To put it yet another way, suppose we have a set of vectors $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ in a subspace V. Then if any two of the following statements is true, the third must also be true:

1. $\mathcal{B}$ is linearly independent
2. $\mathcal{B}$ spans V
3. $\dim(V) = m$

For example, if V is a plane, then any two non-collinear vectors in V form a basis.

Example 1.7.5: Finding a basis

Find a basis for the subspace

$$V = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} : x_1 + x_2 = x_4; x_2 + 2x_3 = x_4 \right\}$$

Solution

First, note that V is the null space of the matrix $\begin{bmatrix} 1 & 1 & 0 & -1 \\ 0 & 1 & 2 & -1 \end{bmatrix}$

We put this matrix in reduced row echelon form $\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 2 & -1 \end{bmatrix}$ and see it has two free variables, so V is indeed a plane.

We write our solution in parametric vector form:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

This shows $V = \text{Span}\left\{ \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$. Since these are clearly not collinear, they are linearly independent; so $\boxed{\left\{ \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}}$ is a basis for V.

1.7: Finding a Basis of a Subspace is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.7: Basis and Dimension** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.7.3: Rank and Nullity

Since each matrix A has two important subspaces related to it (the column space and null space), we give the dimensions of those subspaces names.

Definition: Rank and Nullity

The ***rank*** of a matrix A, written $\text{rank}(A)$, is the dimension of the column space $\text{Col}(A)$.

The ***nullity*** of a matrix A, written $\text{nullity}(A)$, is the dimension of the null space $\text{Null}(A)$.

The rank of a matrix A gives us important information about the solutions to $A\vec{x} = \vec{b}$. Recall that $A\vec{x} = \vec{b}$ is consistent exactly when $\vec{b}$ is in the span of the columns of A, in other words when $\vec{b}$ is in the column space of A. Thus, $\text{rank}(A)$ is the dimension of the set of b with the property that $A\vec{x} = \vec{b}$ is consistent.

We know that the rank of A is equal to the number of pivot columns, and the nullity of A is equal to the number of free variables, which is the number of columns without pivots. To summarize:

$$\begin{aligned} \text{rank}(A) = \dim(\text{Col}(A)) &= \text{the number of columns with pivots} \\ \text{nullity}(A) = \dim(\text{Null}(A)) &= \text{the number of free variables} \\ &= \text{the number of columns without pivots} \end{aligned}$$

Clearly

$$(\text{columns with pivots}) + (\text{columns without pivots}) = (\text{columns})$$

so we have proved the following theorem.

Theorem 1.7.3.4: Rank-Nullity Theorem

If A is a matrix with n columns, then

$$\text{rank}(A) + \text{nullity}(A) = n$$

In other words, for any consistent system of linear equations,

$$(\text{dim of column span}) + (\text{dim of solution set}) = (\text{number of variables})$$

The rank theorem is really the culmination of this chapter, as it gives a strong relationship between the null space of a matrix (the solution set of $A\vec{x} = 0$) with the column space (the set of vectors $\vec{b}$ making $A\vec{x} = \vec{b}$ consistent), our two primary objects of interest. The more freedom we have in choosing $\vec{x}$ the less freedom we have in choosing $\vec{b}$ and vice versa.

Illustration 1.7.3.4: Rank and nullity

Here is a concrete example of the rank theorem and the interplay between the degrees of freedom we have in choosing $\vec{x}$ and $\vec{b}$ in a matrix equation $A\vec{x} = \vec{b}$.

Consider the matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

If we multiply a vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in $\mathbb{R}^3$ by A and B we obtain the vectors $A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$ and $B\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$. So we can think of multiplication by A as giving the latitude and longitude of a point in $\mathbb{R}^3$ and we can think of multiplication by B as giving the height of a point in $\mathbb{R}^3$. The rank of A is 2 and the nullity is 1. Similarly, the rank of B is 1 and the nullity is 2.

These facts have natural interpretations. For the matrix A: the set of all latitudes and longitudes in $\mathbb{R}^3$ is a plane, and the set of points with the same latitude and longitude in $\mathbb{R}^3$ is a line; and for the matrix B: the set of all heights in $\mathbb{R}^3$ is a line, and the set of points at a given height in $\mathbb{R}^3$ is a plane. As the rank theorem tells us, we "trade off" having more choices for $\vec{x}$ for having more choices for $\vec{b}$, and vice versa.

The rank theorem is a prime example of how we use the theory of linear algebra to say something qualitative about a system of equations without ever solving it. This is, in essence, the power of the subject.

Illustration 1.7.3.5: The rank is 2 and the nullity is 2

Consider the following matrix and its reduced row echelon form:

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ -2 & -3 & 4 & 5 \\ 2 & 4 & 0 & -2 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

basis of Col(A) — free variables

Figure 1.7.3.4

A basis for $\mathbf{Col}(A)$ is given by the pivot columns:

$$\left\{ \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} \right\}$$

so $\mathbf{rank}(A) = \mathbf{dim}\,\mathbf{Col}(A) = 2$.

Since there are two free variables x_3, x_4, the null space of A has two vectors:

$$\left\{ \begin{bmatrix} 8 \\ -4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ -3 \\ 0 \\ 1 \end{bmatrix} \right\}$$

so $\mathbf{nullity}(A) = 2$.

In this case, the rank theorem says that $2 + 2 = 4$, where 4 is the number of columns.

Example 1.7.3.6: Using rank and nullity

Consider the system of three equations using five variables $\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 = \alpha \\ b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 = \beta \\ c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 = \gamma \end{cases}$

If we know that this system is consistent for all possible choices of α, β, γ, how many linearly independent solutions to the homogeneous system of equations below?

$$\begin{cases} a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 = 0 \\ b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 = 0 \\ c_1x_1 + c_2x_2 + c_3x_3 + c_4x_4 + c_5x_5 = 0 \end{cases}$$

Solution

We define $A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 \\ b_1 & b_2 & b_3 & b_4 & b_5 \\ c_1 & c_2 & c_3 & c_4 & c_5 \end{bmatrix}$ to make things easier to describe.

Since the system $A\vec{x} = \vec{b}$ is consistent for all $\vec{b} = \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix}$ in $\mathbb{R}^3$, we know that $\text{Col}(A) = \mathbb{R}^3$

That means $\dim(\text{Col}(A)) = \text{rank}(A) = 3$, and since it has 5 columns, we know $\text{nullity}(A) = 5 - 3 = 2$

Therefore, any basis for $\text{Null}(A)$ must contain exactly two linearly independent vectors.

This is exactly the same as the collection of solutions to $A\vec{x} = \vec{0}$, so our answer is $\boxed{\text{two linearly independent solutions}}$

1.7.3: Rank and Nullity is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.9: The Rank Theorem** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

1.7.4: Practice Set

Exercise 1.7.4.1a

Find a basis for $A = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x = -\frac{2}{3}y \right\}$ and determine its dimension.

Answer

$$A = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2 : x = -\tfrac{2}{3}y \right\} = \left\{ \begin{bmatrix} -\frac{2}{3}y \\ y \end{bmatrix} \right\} = \left\{ y \begin{bmatrix} -\frac{2}{3} \\ 1 \end{bmatrix} \right\} = \text{Span} \left\{ \begin{bmatrix} -\frac{2}{3} \\ 1 \end{bmatrix} \right\}$$

Since this vector is non-zero, it is linearly independent and thus forms a basis for A.

Since there is one vector in a basis for A, we have $\dim(A) = 1$

Exercise 1.7.4.1b

Find a basis for $B = \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \in \mathbb{R}^3 : 4x + y = 0 \right\}$ and determine its dimension.

Answer

$$B = \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \in \mathbb{R}^3 : 4x + y = 0 \right\} = \left\{ \begin{bmatrix} x \\ -4x \\ 0 \end{bmatrix} \right\} = \left\{ x \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix} \right\} = \text{Span} \left\{ \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix} \right\}$$

Since this vector is non-zero, it is linearly independent and thus forms a basis for B.

Since there is one vector in a basis for B, we have $\dim(B) = 1$

Exercise 1.7.4.1c

Find a basis for $C = \left\{ \begin{bmatrix} a + 10b \\ b + 10c \\ c + 10a \end{bmatrix} \in \mathbb{R}^3 : a, b, c \in \mathbb{R} \right\}$ and determine its dimension.

Answer

$$C = \left\{ \begin{bmatrix} a + 10b \\ b + 10c \\ c + 10a \end{bmatrix} \in \mathbb{R}^3 : a, b, c \in \mathbb{R} \right\} = \left\{ a \begin{bmatrix} 1 \\ 0 \\ 10 \end{bmatrix} + b \begin{bmatrix} 10 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 10 \\ 1 \end{bmatrix} \right\} = \text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 10 \end{bmatrix}, \begin{bmatrix} 10 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 10 \\ 1 \end{bmatrix} \right\}$$

These three vectors are a spanning set, and $\begin{bmatrix} 1 & 10 & 0 \\ 0 & 1 & 10 \\ 10 & 0 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Therefore they are linearly independent, and are a basis of C

Since there are three vectors in a basis for C, we have $\dim(C) = 3$

Exercise 1.7.4.1d

Find a basis for $D = \left\{ \begin{bmatrix} a \\ b + c \\ a + c \\ d \end{bmatrix} \in \mathbb{R}^4 : a + b = d, a + c = d \right\}$ and determine its dimension.

Answer

$\begin{cases} a + b = d \end{cases}$ $\begin{cases} b = d - a \end{cases}$

First we simplify $D = \left\{ \begin{bmatrix} a \\ b+c \\ a+c \\ d \end{bmatrix} \in \mathbb{R}^4 : a+b=d, a+c=d \right\} \begin{cases} a+c=d \\ \end{cases} \implies \begin{cases} c=d-a \\ \end{cases} = \left\{ \begin{bmatrix} a \\ 2d-2a \\ d \\ d \end{bmatrix} \right\} = \left\{ a \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ 2 \\ 1 \\ 1 \end{bmatrix} \right\} = \text{Span} \left\{ \begin{bmatrix} 1 \\ -2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 1 \\ 1 \end{bmatrix} \right\}$

These two vectors are clearly not multiples of one another so they must be linearly independent and form a basis for D.

Since there are two vectors in a basis for D, we have $\dim(D) = 2$

Exercise 1.7.4.1e

Find a basis for $E = \left\{ \begin{bmatrix} 3a+2b-c \\ 2a-7b+c \\ a+4b+2c \\ 2a+13b \end{bmatrix} \in \mathbb{R}^4 : a, b, c \in \mathbb{R} \right\}$ and determine its dimension.

Answer

$$E = \left\{ \begin{bmatrix} 3a+2b-c \\ 2a-7b+c \\ a+4b+2c \\ 2a+13b \end{bmatrix} \in \mathbb{R}^4 : a, b, c \in \mathbb{R} \right\} = \left\{ a \begin{bmatrix} 3 \\ 2 \\ 1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 2 \\ -7 \\ 4 \\ 13 \end{bmatrix} + c \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix} \right\} = \text{Span} \left\{ \begin{bmatrix} 3 \\ 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -7 \\ 4 \\ 13 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix} \right\}$$

These three vectors are a spanning set, and $\begin{bmatrix} 3 & 2 & -1 \\ 2 & -7 & 1 \\ 1 & 4 & 2 \\ 2 & 13 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Therefore they are linearly independent, and are a basis of E

Since there are three vectors in a basis for E, we have $\dim(E) = 3$

Exercise 1.7.4.1f

Find a basis for $F = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 : \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$ and determine its dimension.

Answer

First we simplify $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \begin{cases} a+2b+3c+4d=0 \\ 5a+6b+7c+8d=0 \end{cases} \implies \begin{cases} a=c+2d \\ b=-2c-3d \end{cases}$

Now $F = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 : \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} c+2d \\ -2c-3d \\ c \\ d \end{bmatrix} \right\} = \left\{ c \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + d \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} \right\}$

$= \text{Span} \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix} \right\}$

These two vectors are clearly not multiples of one another so they must be linearly independent and form a basis for F.

Since there are two vectors in a basis for F, we have $\dim(F) = 2$

Exercise 1.7.4.1g

Find a basis for $G = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 : \begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = 0 \right\}$ and determine its dimension.

Answer

First we simplify $\begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = 0 \implies a + 4b + 5c + 7d = 0 \implies a = -4b - 5c - 7d$

Now $G = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 : \begin{bmatrix} a & b & c & d \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 5 \\ 7 \end{bmatrix} = 0 \right\} = \left\{ \begin{bmatrix} -4b - 5c - 7d \\ b \\ c \\ d \end{bmatrix} \right\} = \left\{ b \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c \begin{bmatrix} -5 \\ 0 \\ 1 \\ 0 \end{bmatrix} + d \begin{bmatrix} -7 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

$= \text{Span} \left\{ \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -7 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

These three vectors are a spanning set, and $\begin{bmatrix} -4 & -5 & -7 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Therefore they are linearly independent, and are a basis of G

Since there are three vectors in a basis for G, we have $\dim(G) = 3$

Exercise 1.7.4.1h

Find a basis for $H = \left\{ \begin{bmatrix} 2a + b \\ b + 2c \\ a - c \end{bmatrix} \in \mathbb{R}^3 : a, b, c \in \mathbb{R} \right\}$ and determine its dimension.

Answer

$H = \left\{ \begin{bmatrix} 2a + b \\ b + 2c \\ a - c \end{bmatrix} \in \mathbb{R}^3 : a, b, c \in \mathbb{R} \right\} = \left\{ a \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} \right\} = \text{Span} \left\{ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} \right\}$

These three vectors are a spanning set, and $\begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \\ 1 & 0 & -1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

Therefore they are linearly dependent, and removing the free variable column creates a basis for H, which is: $\left\{ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$

Since there are two vectors in a basis for H, we have $\dim(H) = 2$

Exercise 1.7.4.2a

For the matrix $A = \begin{bmatrix} 1 & 1 & 0 & -1 & 0 \\ 1 & 1 & 2 & 3 & 1 \\ 1 & 1 & 3 & 5 & 2 \end{bmatrix}$

(i) Find a basis for $\text{Col}(A)$

(ii) Find a basis for $\mathbf{Null}(A)$

(iii) Verify the Rank-Nullity Equation holds

Answer

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

(i) We keep only the columns in A which became pivots in RREF.

Thus $\mathbf{Col}(A)$ has basis $\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right\}$

(ii) $$\begin{cases} x_1 + x_2 - x_4 = 0 \\ x_3 + 2x_4 = 0 \\ x_5 = 0 \\ x_2, x_4 \text{ free} \end{cases} \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}$$

Thus $\mathbf{Null}(A)$ has basis $\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} \right\}$

(iii) We can see that $\mathbf{Rank}(A) = \dim(\mathrm{Col}(A)) = 3$ and $\mathbf{Nullity}(A) = \dim(\mathbf{Null}(A)) = 2$:

$\mathbf{Rank}(A) + \mathbf{Nullity}(A) = 3 + 2 = 5 = \mathbf{Number\ of\ columns\ in}\ A$

? Exercise 1.7.4.2*b*

For the matrix $B = \begin{bmatrix} 4 & 8 & 0 & 29 & 0 \\ 2 & 3 & 0 & 13 & 2 \\ 1 & 4 & 0 & 11 & -4 \end{bmatrix}$

(i) Find a basis for $\mathbf{Col}(B)$

(ii) Find a basis for $\mathbf{Null}(B)$

(iii) Verify the Rank-Nullity Equation holds

Answer

$$\begin{bmatrix} 4 & 8 & 0 & 29 & 0 \\ 2 & 3 & 0 & 13 & 2 \\ 1 & 4 & 0 & 11 & -4 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 & 0 & 4 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

(i) We keep only the columns in B which became pivots in RREF.

Thus $\mathbf{Col}(B)$ has basis $\left\{ \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 29 \\ 13 \\ 11 \end{bmatrix} \right\}$

(ii) $$\begin{cases} x_1 + 4x_5 = 0 \\ x_2 - 2x_5 = 0 \\ x_4 = 0 \\ x_3, x_5 \text{ free} \end{cases} \implies \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -4 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Thus $\mathbf{Null}(B)$ has basis $\left\{\begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-4\\2\\0\\0\\1\end{bmatrix}\right\}$

(iii) We can see that $\mathbf{Rank}(B) = \mathbf{dim}\,(\mathbf{Col}(B)) = 3$ and $\mathbf{Nullity}(B) = \mathbf{dim}\,(\mathbf{Null}(B)) = 2$:

$\mathbf{Rank}(B) + \mathbf{Nullity}(B) = 3 + 2 = 5 = \mathbf{Number\ of\ columns\ in}\ B$

? Exercise 1.7.4.2*c*

For the matrix $C = \begin{bmatrix}1 & 1 & 2\\2 & -1 & 4\\1 & 0 & 2\end{bmatrix}$

(i) Find a basis for $\mathbf{Col}(C)$

(ii) Find a basis for $\mathbf{Null}(C)$

(iii) Verify the Rank-Nullity Equation holds

Answer

$$\begin{bmatrix}1 & 1 & 2\\2 & -1 & 4\\1 & 0 & 2\end{bmatrix} \xrightarrow{RREF} \begin{bmatrix}1 & 0 & 2\\0 & 1 & 0\\0 & 0 & 0\end{bmatrix}$$

(i) We keep only the columns in C which became pivots in RREF.

Thus $\mathbf{Col}(C)$ has basis $\left\{\begin{bmatrix}1\\2\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\0\end{bmatrix}\right\}$

(ii) $\begin{cases}x_1 + 2x_3 = 0\\x_2 = 0\\x_3\ \text{free}\end{cases} \Longrightarrow \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = x_3\begin{bmatrix}-2\\0\\1\end{bmatrix}$

Thus $\mathbf{Null}(C)$ has basis $\left\{\begin{bmatrix}-2\\0\\1\end{bmatrix}\right\}$

(iii) We can see that $\mathbf{Rank}(C) = \mathbf{dim}\,(\mathbf{Col}(C)) = 2$ and $\mathbf{Nullity}(C) = \mathbf{dim}\,(\mathbf{Null}(C)) = 1$:

$\mathbf{Rank}(C) + \mathbf{Nullity}(C) = 2 + 1 = 3 = \mathbf{Number\ of\ columns\ in}\ C$

? Exercise 1.7.4.2*d*

For the matrix $D = \begin{bmatrix}1 & 1 & 1 & 1\\1 & 1 & 1 & 1\\1 & 1 & 1 & 1\\1 & 1 & 1 & 1\end{bmatrix}$

(i) Find a basis for $\mathbf{Col}(D)$

(ii) Find a basis for $\mathbf{Null}(D)$

(iii) Verify the Rank-Nullity Equation holds

Answer

$$\begin{bmatrix}1 & 1 & 1 & 1\\1 & 1 & 1 & 1\\1 & 1 & 1 & 1\\1 & 1 & 1 & 1\end{bmatrix} \xrightarrow{RREF} \begin{bmatrix}1 & 1 & 1 & 1\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}$$

(i) We keep only the columns in D which became pivots in RREF.

Thus $\text{Col}(D)$ has basis $\left\{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}\right\}$

(ii) $\begin{cases} x_1+x_2+x_3+x_4=0 \\ x_2,x_3,x_4 \text{ free} \end{cases} \Longrightarrow \begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} = x_2\begin{bmatrix}-1\\1\\0\\0\end{bmatrix} + x_3\begin{bmatrix}-1\\0\\1\\0\end{bmatrix} + x_4\begin{bmatrix}-1\\0\\0\\1\end{bmatrix}$

Thus $\text{Null}(D)$ has basis $\left\{\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\0\\1\end{bmatrix}\right\}$

(iii) We can see that $\text{Rank}(D) = \dim(\text{Col}(D)) = 1$ and $\text{Nullity}(D) = \dim(\text{Null}(D)) = 3$:

$\text{Rank}(D) + \text{Nullity}(D) = 1+3 = 4 = \text{Number of columns in } D$

? Exercise 1.7.4.2e

For the matrix $E = \begin{bmatrix}0&-1&0&0&0\\0&0&\sqrt{3}&0&0\\0&0&0&e&0\\0&0&0&0&\pi\end{bmatrix}$

(i) Find a basis for $\text{Col}(E)$

(ii) Find a basis for $\text{Null}(E)$

(iii) Verify the Rank-Nullity Equation holds

Answer

$$\begin{bmatrix}0&-1&0&0&0\\0&0&\sqrt{3}&0&0\\0&0&0&e&0\\0&0&0&0&\pi\end{bmatrix} \xrightarrow{RREF} \begin{bmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\end{bmatrix}$$

(i) We keep only the columns in E which became pivots in RREF.

Thus $\text{Col}(E)$ has basis $\left\{\begin{bmatrix}-1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\\sqrt{3}\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\e\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\\pi\end{bmatrix}\right\}$

(ii) $\begin{cases} x_2=0 \\ x_3=0 \\ x_4=0 \\ x_5=0 \\ x_1 \text{ free} \end{cases} \Longrightarrow \begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix} = x_1\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix}$

Thus $\text{Null}(E)$ has basis $\left\{\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix}\right\}$

(iii) We can see that $\text{Rank}(E) = \dim(\text{Col}(E)) = 4$ and $\text{Nullity}(E) = \dim(\text{Null}(E)) = 1$:

$\text{Rank}(E) + \text{Nullity}(E) = 4+1 = 5 = \text{Number of columns in } E$

Exercise 1.7.4.3

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) A set of n linearly independent vectors in $\mathbb{R}^n$ is a spanning set of $\mathbb{R}^n$.

(b) A spanning set of $\mathbb{R}^n$ containing n vectors is a basis for $\mathbb{R}^n$.

(c) A set containing more than n vectors from $\mathbb{R}^n$ is linearly dependent.

(d) A set containing more than n vectors from $\mathbb{R}^n$ is a spanning set.

(e) If the columns of a matrix are linearly independent, then the rows are linearly independent.

(f) If the n columns of a matrix are a spanning set of $\mathbb{R}^n$, then the rows of that matrix form a basis for $\mathbb{R}^n$.

(g) If the rows of a matrix form a spanning set for $\mathbb{R}^n$, then the columns of that matrix are linearly independent.

(h) If a matrix has n columns, then the rank of that matrix is at most n.

(i) If a matrix has m rows, then the rank of that matrix is at most m.

(j) If A is a 3×4 matrix, then the solution set to $A\vec{x} = \vec{0}$ contains non-trivial vectors.

(k) If A is a 4×6 matrix, and the solution set of $A\vec{x} = \vec{0}$ is spanned by two linearly independent vectors, then the column space of A is equal to all of $\mathbb{R}^4$.

Answer

(a) **Always True**

A pivot in every column of an $n \times n$ matrix must mean pivot in every row, so its columns must span all of $\mathbb{R}^n$.

(b) **Always True**

A pivot in every row of an $n \times n$ matrix must have a pivot in every column, and thus its columns must be linearly independent.

(c) **Always True**

A matrix wider than it is tall will always have free variables.

(d) **Sometimes True Sometimes False**

This statement is true when there are at least n linearly independent vectors inside, but can be false if all of the vectors were $\vec{0}$ for instance.

(e) **Sometimes True Sometimes False**

This is true for all square matrices, but if the matrix is taller than it is wide, the rows can never be linearly independent which makes it false.

(f) **Always True**

No free variables in a square matrix means a pivot in every row, which means no row is in the span of the others, or else it would turn into a row of all zeros.

(g) **Always True**

There must be at least n pivot rows to span $\mathbb{R}^n$, and there can only be n columns so they must each have a pivot.

(h) **Always True**

Since $\text{Nullity}(A) \geq 0$ we have $\text{Rank}(A) = n - \text{Nullity}(A) \leq n$.

(i) **Always True**

You cannot have more pivots in echelon form than you have rows, by definition a pivot is the *first nonzero entry in a row.*

(j) **Always True**

Since $\mathbf{Rank}(A) \leq 3$ by (i) above, we know that $\mathbf{Nullity}(A) = 4 - \mathbf{Rank}(A) \geq 1$.

(k) **Always True**

Since $\mathbf{Nullity}(A) = 2$, we have $\mathbf{Rank}(A) = 6 - \mathbf{Nullity}(A) = 4$, so the four linearly independent columns of A must span all of $\mathbb{R}^4$.

1.7.4: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

2: Linear Transformations and Matrix Algebra

2.1: Matrix Transformations

2.1: Matrices as Functions

2.1: Matrix Transformations

2.1: Practice Set

2.2: Linear Transformations

2.2: Linear Transformations

2.2: Matrix Representations

2.2: Practice Set

2.3: Matrix Algebra

2.3: Sums of Transformations and Matrices

2.3: Transformation Composition and Matrix Multiplication

2.3: Matrix Inverses

2.3: Practice Set

2.4: Invertibility

2.4: Invertibility of Linear Transformations

2.4: Invertibility of Matrices

2.4: Practice Set

2.5: Determinants- Definition

2.5: Definition of Determinant

2.5: Properties of the Determinant

2.5: Practice Set

2.6: Cofactor Expansions

2.6: Cofactor Expansion

2.6: Cramer's Rule and Adjugates

2.6: Practice Set

2.7: Determinants and Volumes

2.7: Determinants for Area and Volume

2.7: Practice Set

This page titled 2: Linear Transformations and Matrix Algebra is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

2.1: Matrix Transformations

Objectives

1. Learn to view a matrix geometrically as a function.
2. Learn examples of matrix transformations: reflection, dilation, rotation, shear, projection.
3. Understand the vocabulary surrounding transformations: domain, codomain, range.
4. Understand the domain, codomain, and range of a matrix transformation.
5. Understand the definitions of one-to-one and onto transformations.
6. *Pictures:* common matrix transformations, examples of matrix transformations that are/are not one-to-one and/or onto.
7. *Vocabulary words:* ***transformation*** / ***function***, ***domain***, ***codomain***, ***range***, ***identity transformation***, ***matrix transformation,*** ***one-to-one***, ***onto***

In this section, we learn to understand matrices geometrically as *functions*, or *transformations*. We briefly discuss transformations in general, then specialize to matrix transformations, which are transformations that come from matrices.

This page titled 2.1: Matrix Transformations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: Matrix Transformations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.1: Matrices as Functions

Informally, a *function* is a rule that takes inputs and produces outputs. For instance, $f(x) = x^2$ is a function that accepts one number x as its input, and outputs the square of that number: $f(2) = 4$. In this subsection, we interpret matrices as functions.

Let A be a matrix with m rows and n columns. Consider the matrix equation $\vec{b} = A\vec{x}$ (we write it this way instead of $A\vec{x} = \vec{b}$ to remind the reader of the notation $y = f(x)$). If we vary $\vec{x}$, then $\vec{b}$ will also vary; in this way, we think of A as a function with independent variable (input) $\vec{x}$ and dependent variable (output) $\vec{b}$.

- The independent variable (the input) is $\vec{x}$, which is a vector in $\mathbb{R}^n$.
- The dependent variable (the output) is $\vec{b}$, which is a vector in $\mathbb{R}^m$.

The range of this function is the set of all possible output vectors $\vec{b}$ such that $A\vec{x} = \vec{b}$ has some solution; this is the same as the column space of A.

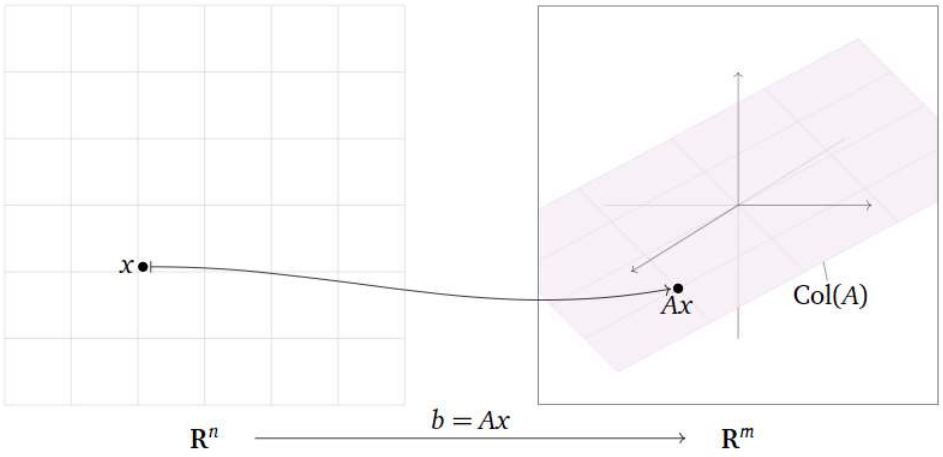


Figure 2.1.1

Example 2.1.1: Projection onto the xy-plane

Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Describe the function $\vec{b} = A\vec{x}$ geometrically.

Solution

In the equation $A\vec{x} = \vec{b}$, the input vector $\vec{x}$ and the output vector $\vec{b}$ are both in $\mathbb{R}^3$. First we multiply A by a vector to see what it does:

$$A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$$

Multiplication by A simply sets the z-coordinate equal to zero: $\boxed{\text{it projects vertically onto the } xy\text{-plane}}$

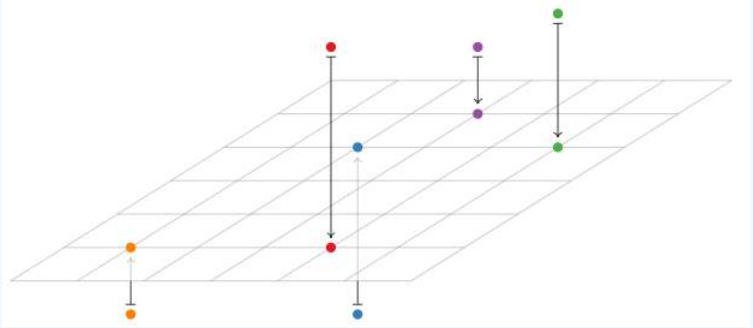


Figure 2.1.2

Since A is in reduced row echelon form, the column space of A is easily seen to be

$$\text{Col}\,(A) = \text{Span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$$

This is, of course, the entire xy-plane. In fact we can see that for any vector in this plane $\begin{bmatrix}a\\b\\0\end{bmatrix}$, we have infinitely many points in the domain that maps to it:

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}a\\b\\0\end{bmatrix}$$

Example 2.1.2: Reflection

Let

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

Describe the function $\vec{b} = A\vec{x}$ geometrically.

Solution

In the equation $A\vec{x} = \vec{b}$, the input vector $\vec{x}$ and the output vector $\vec{b}$ are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

$$A\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}-1&0\\0&1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}-x\\y\end{bmatrix}$$

Multiplication by A negates the x-coordinate: it reflects over the y-axis

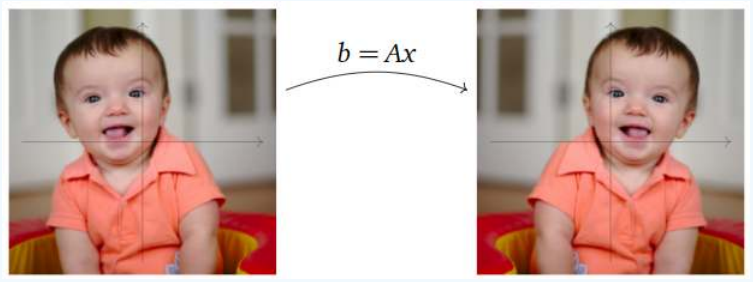


Figure 2.1.3

Since A is in reduced row echelon form, the column space of A is easily seen to be

$$\text{Col}\,(A) = \text{Span}\left\{\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$$

This is, of course, the entirety of $\mathbb{R}^2$ plane. Once more we can see that for any vector in this plane $\begin{bmatrix} a \\ b \end{bmatrix}$, we have at least one point in the domain that maps to it:

$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

Example 2.1.3: Dilation

Let

$$A = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix}$$

Describe the function $\vec{b} = A\vec{x}$ geometrically.

Solution

In the equation $A\vec{x} = \vec{b}$, the input vector $\vec{x}$ and the output vector $\vec{b}$ are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

$$A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.5x \\ 1.5y \end{bmatrix} = 1.5\begin{bmatrix} x \\ y \end{bmatrix}$$

Multiplication by A is the same as scalar multiplication by 1.5: it scales or dilates the plane by a factor of 1.5

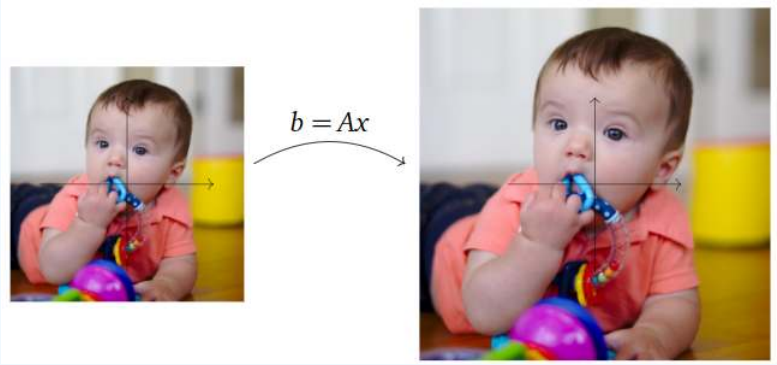


Figure 2.1.4

Since A is in echelon form, the column space of A is easily seen to be

$$\text{Col}\,(A) = \text{Span}\left\{\begin{bmatrix} 1.5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1.5 \end{bmatrix}\right\}$$

This is, of course, the entirety of $\mathbb{R}^2$-plane. Once more we can see that for any vector in this plane $\begin{bmatrix} a \\ b \end{bmatrix}$, we have at least one point in the domain that maps to it:

$$\begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix}\begin{bmatrix} \frac{2}{3}a \\ \frac{2}{3}b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

Example 2.1.4: Identity

Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Describe the function $\vec{b} = A\vec{x}$ geometrically.

Solution

In the equation $A\vec{x} = \vec{b}$, the input vector $\vec{x}$ and the output vector $\vec{b}$ are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

$$A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

Multiplication by A does not change the input vector at all: it is the identity transformation which does nothing

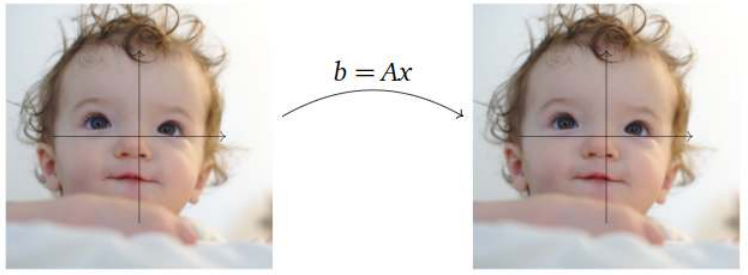


Figure 2.1.5

It is simple to see that once more, the range of this function matches the column space of A, which is all of $\mathbb{R}^2$

Example 2.1.5: Rotation

Let

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

Describe the function $\vec{b} = A\vec{x}$ geometrically.

Solution

In the equation $A\vec{x} = \vec{b}$, the input vector $\vec{x}$ and the output vector $\vec{b}$ are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

$$A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix}$$

We substitute a few test points in order to understand the geometry of the transformation:

$$A\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$
$$A\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$
$$A\begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

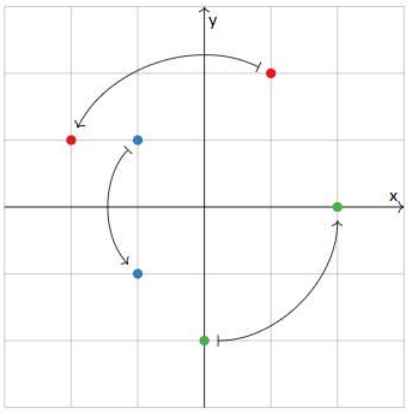


Figure 2.1.6

Multiplication by A is counterclockwise rotation by 90°

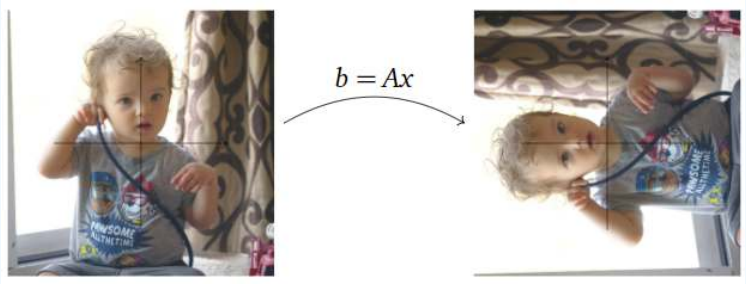


Figure 2.1.7

It is simple to see that once more, the range of this function matches the column space of A, which is all of $\mathbb{R}^2$

Example 2.1.6: Shear

Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Describe the function $\vec{b} = A\vec{x}$ geometrically.

Solution

In the equation $A\vec{x} = \vec{b}$, the input vector $\vec{x}$ and the output vector $\vec{b}$ are both in $\mathbb{R}^2$. First we multiply A by a vector to see what it does:

$$A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ y \end{bmatrix}$$

Multiplication by A adds the y-coordinate to the x-coordinate; this is called a shear in the x-direction

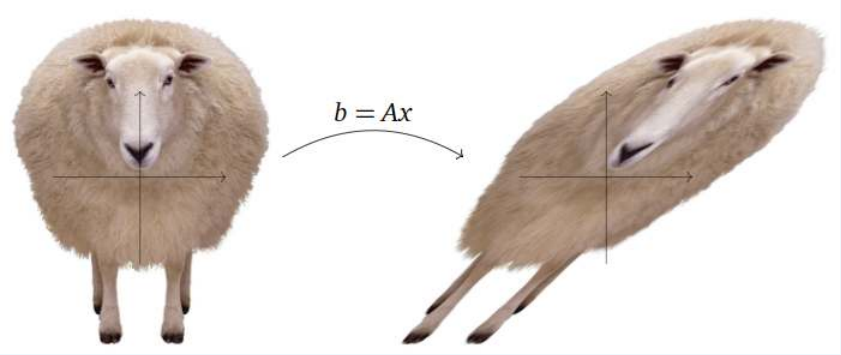


Figure 2.1.8

Since A is in echelon form, the column space of A is easily seen to be

$$\text{Col}\,(A) = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

This is, of course, the entirety of $\mathbb{R}^2$-plane. Once more we can see that for any vector in this plane $\begin{bmatrix} a \\ b \end{bmatrix}$, we have at least one point in the domain that maps to it:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a-b \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

2.1: Matrices as Functions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.1: Matrix Transformations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.1: Matrix Transformations

Now we specialize the general notions and vocabulary above to the functions defined by matrices that we considered in the start of this chapter.

Definition: Matrix Transformation

Let A be an $m \times n$ matrix. The ***matrix transformation*** associated to A is the transformation

$$T\colon \mathbb{R}^n \to \mathbb{R}^m \quad \text{defined by} \quad T(\vec{x}) = A\vec{x}$$

This is the transformation that takes a vector $\vec{x}$ in $\mathbb{R}^n$ to the vector $A\vec{x}$ in $\mathbb{R}^m$.

If A has n columns, then it only makes sense to multiply A by vectors with n entries. This is why the domain of $T(\vec{x}) = A\vec{x}$ is $\mathbb{R}^n$. If A has m rows, then $A\vec{x}$ has m entries for any vector $\vec{x}$ in $\mathbb{R}^n$; this is why the codomain of $T(\vec{x}) = A\vec{x}$ is $\mathbb{R}^m$.

The definition of a matrix transformation T tells us how to evaluate T on any given vector: we multiply the input vector by a matrix. For instance, let

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

and let $T(\vec{x}) = A\vec{x}$ be the associated matrix transformation. Then

$$T\left(\begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}\right) = A\begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix} = \begin{bmatrix} -14 \\ -32 \end{bmatrix}$$

Suppose that A has columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$. If we multiply A by a general vector $\vec{x}$, we get

$$A\vec{x} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n$$

This is just a general linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$. Therefore, the outputs of $T(\vec{x}) = A\vec{x}$ are exactly the linear combinations of the columns of A: the *range* of T is the column space of A.

Note 2.1.1

Let A be an $m \times n$ matrix, and let $T(\vec{x}) = A\vec{x}$ be the associated matrix transformation.

- The *domain* of T is $\mathbb{R}^n$, where n is the number of *columns* of A.
- The *codomain* of T is $\mathbb{R}^m$, where m is the number of *rows* of A.
- The *range* of T is the *column space* of A.

Definition: One-to-one transformations

A transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is ***one-to-one*** if, for every vector $\vec{b}$ in $\mathbb{R}^m$, the equation $T(\vec{x}) = \vec{b}$ has *at most one* solution $\vec{x}$ in $\mathbb{R}^n$.

Remark

Another word for *one-to-one* is *injective*.

Here are some equivalent ways of saying that T is one-to-one:

- For every vector $\vec{b}$ in $\mathbb{R}^m$, the equation $T(\vec{x}) = \vec{b}$ has *zero or one* solution $\vec{x}$ in $\mathbb{R}^n$.
- Different inputs of T have different outputs.

- If $T(\vec{u}) = T(\vec{v})$ then $\vec{u} = \vec{v}$.

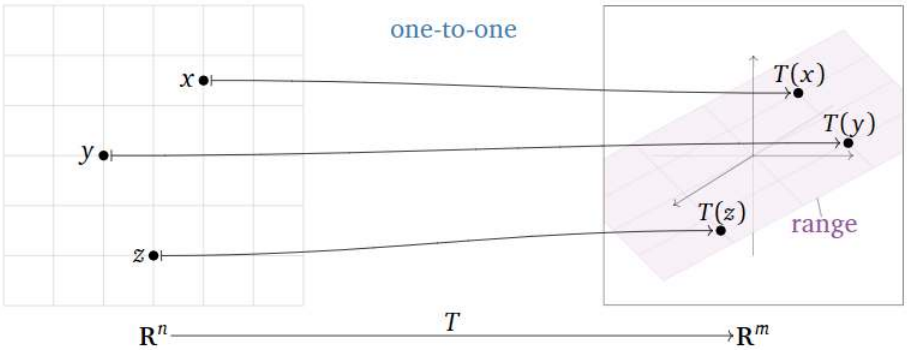


Figure 2.1.12

Here are some equivalent ways of saying that T is *not* one-to-one:

- There exists some vector $\vec{b}$ in $\mathbb{R}^m$ such that the equation $T(\vec{x}) = \vec{b}$ has *more than one* solution $\vec{x}$ in $\mathbb{R}^n$.
- There are two different inputs of T with the same output.
- There exist vectors $\vec{u}, \vec{v}$ such that $\vec{u} \neq \vec{v}$ but $T(\vec{u}) = T(\vec{v})$.

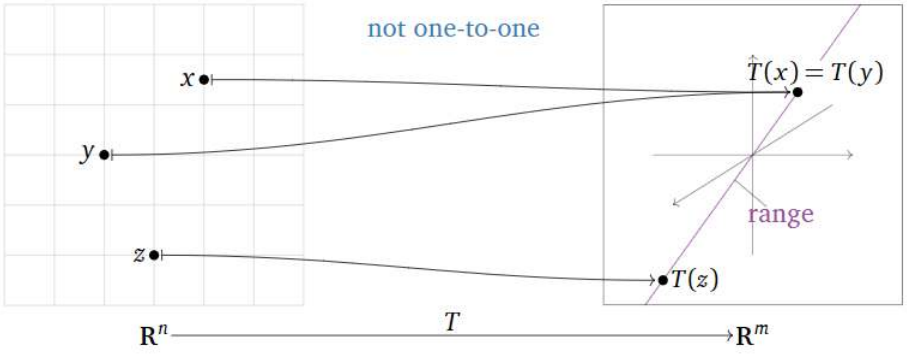


Figure 2.1.13

Illustration 2.1.1: Functions of one variable

The function $\sin\colon \mathbb{R} \to \mathbb{R}$ is not one-to-one. Indeed, $\sin(0) = \sin(\pi) = 0$, so the inputs 0 and π have the same output 0. In fact, the equation $\sin(x) = 0$ has infinitely many solutions $\ldots, -2\pi, -\pi, 0, \pi, 2\pi, \ldots$.

The function $\exp\colon \mathbb{R} \to \mathbb{R}$ defined by $\exp(x) = e^x$ is one-to-one. Indeed, if $T(x) = T(y)$, then $e^x = e^y$, so $\ln(e^x) = \ln(e^y)$, and hence $x = y$. The equation $T(x) = C$ has one solution $x = \ln(C)$ if $C > 0$, and it has zero solutions if $C \leq 0$.

The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$ is one-to-one. Indeed, if $f(x) = f(y)$ then $x^3 = y^3$; taking cube roots gives $x = y$. In other words, the only solution of $f(x) = C$ is $x = \sqrt[3]{C}$.

The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3 - x$ is not one-to-one. Indeed, $f(0) = f(1) = f(-1) = 0$, so the inputs $0, 1, -1$ all have the same output 0. The solutions of the equation $x^3 - x = 0$ are exactly the roots of $f(x) = x(x-1)(x+1)$, and this equation has three roots.

The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is not one-to-one. Indeed, $f(1) = 1 = f(-1)$, so the inputs 1 and -1 have the same outputs. The function $g\colon \mathbb{R} \to \mathbb{R}$ defined by $g(x) = |x|$ is not one-to-one for the same reason.

Theorem 2.1.1: One-to-One Matrix Transformations

Let A be an $m \times n$ matrix, and let $T(\vec{x}) = A\vec{x}$ be the associated matrix transformation. The following statements are equivalent:

1. T is one-to-one.

2. For every $\vec{b}$ in $\mathbb{R}^m$, the equation $T(\vec{x}) = \vec{b}$ has at most one solution.
3. For every $\vec{b}$ in $\mathbb{R}^m$, the equation $A\vec{x} = \vec{b}$ has a unique solution or is inconsistent.
4. $A\vec{x} = \vec{0}$ has *only* the trivial solution.
5. The columns of A are linearly independent.
6. A has a pivot in every column.
7. The range of T has dimension n.

Proof

Statements 1, 2, and 3 are translations of each other. The equivalence of 3 and 4 follows from: if $A\vec{x} = \vec{0}$ has only one solution, then $A\vec{x} = \vec{b}$ has only one solution as well, or it is inconsistent. The equivalence of 4, 5, and 6 is a consequence of how we check for linear independence, and the equivalence of 6 and 7 follows from the fact that the rank of a matrix is equal to the number of columns with pivots.

Recall that *equivalent* means that, for a given matrix, either all of the statements are true simultaneously, or they are all false.

Definition: Onto transformations

A transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is ***onto*** if, for every vector $\vec{b}$ in $\mathbb{R}^m$, the equation $T(\vec{x}) = \vec{b}$ has *at least one* solution $\vec{x}$ in $\mathbb{R}^n$.

Remark

Another word for *onto* is *surjective*.

Here are some equivalent ways of saying that T is onto:

- The range of T is equal to the codomain of T.
- Every vector in the codomain is the output of some input vector.

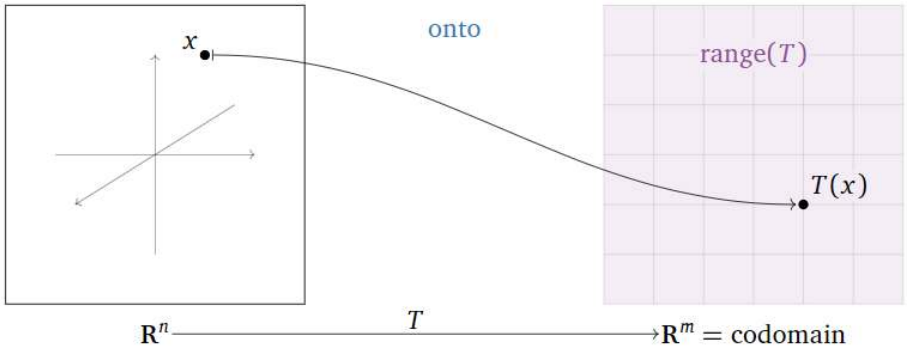


Figure 2.1.14

Here are some equivalent ways of saying that T is *not* onto:

- The range of T is smaller than the codomain of T.
- There exists a vector $\vec{b}$ in $\mathbb{R}^m$ such that the equation $T(\vec{x}) = \vec{b}$ does not have a solution.
- There is a vector in the codomain that is not the output of any input vector.

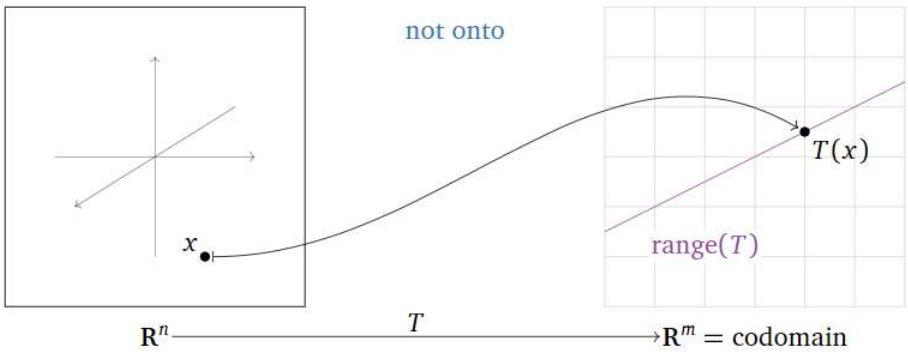


Figure 2.1.15

Illustration 2.1.2: Functions of one variable

The function $\sin\colon \mathbb{R} \to \mathbb{R}$ is not onto. Indeed, taking $b = 2$, the equation $\sin(x) = 2$ has no solution. The range of $\sin$ is the closed interval $[-1, 1]$, which is smaller than the codomain $\mathbb{R}$.

The function $\exp\colon \mathbb{R} \to \mathbb{R}$ defined by $\exp(x) = e^x$ is not onto. Indeed, taking $b = -1$, the equation $\exp(x) = e^x = -1$ has no solution. The range of $\exp$ is the set $(0, \infty)$ of all *positive* real numbers.

The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3$ is onto. Indeed, the equation $f(x) = x^3 = b$ always has the solution $x = \sqrt[3]{b}$.

The function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^3 - x$ is onto. Indeed, the solutions of the equation $f(x) = x^3 - x = b$ are the roots of the polynomial $x^3 - x - b$; as this is a cubic polynomial, it has at least one real root.

Theorem 2.1.2: Onto Matrix Transformations

Let A be an $m \times n$ matrix, and let $T(\vec{x}) = A\vec{x}$ be the associated matrix transformation. The following statements are equivalent:

1. T is onto.
2. $T(\vec{x}) = \vec{b}$ has at least one solution for every $\vec{b}$ in $\mathbb{R}^m$.
3. $A\vec{x} = \vec{b}$ is consistent for every $\vec{b}$ in $\mathbb{R}^m$.
4. The columns of A span $\mathbb{R}^m$.
5. A has a pivot in every row.
6. The range of T has dimension m.

Proof

Statements 1, 2, and 3 are translations of each other. The equivalence of 3, 4, 5, and 6 follows from our earlier work with matrix equations.

Example 2.1.10: Projection onto the xy-plane: reprise

Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and let $T(\vec{x}) = A\vec{x}$. What are the domain, the codomain, and the range of T?

Is T one-to-one? Is T onto?

Solution

We saw earlier that geometrically, the transformation T projects a vector directly "down" onto the xy-plane in $\mathbb{R}^3$.

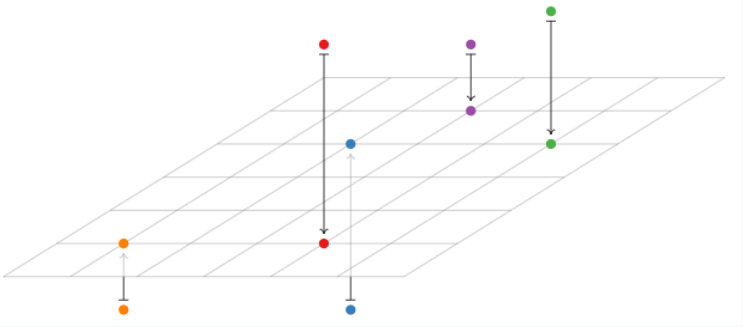


Figure 2.1.16

The inputs and outputs have three entries, so the domain and codomain are both $\mathbb{R}^3$. The possible outputs all lie on the xy-plane, and every point on the xy-plane is an output of T (with itself as the input), so the range of T is the xy-plane. This agrees with the column space of A as discussed when we previously saw this matrix above.

Be careful not to confuse the codomain with the range here. The range is a plane, but it is a plane *in* $\mathbb{R}^3$, so the codomain is still $\mathbb{R}^3$. The outputs of T all have three entries; the last entry is simply always zero. This means that we can find at least one vector $\vec{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ which is not in the range of T.

Clearly it $\boxed{\text{cannot be an onto transformation}}$

Notice that the *reason* that our range ends up "smaller" than our domain/codomain, is that we have many different input vectors that yield the same output. For example, it is easy to see that $T\left(\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}\right) = T\left(\begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}\right)$. Because of these repeats, our function is $\boxed{\text{not a one-to-one transformation}}$

In the case of an $n \times n$ square matrix, the domain and codomain of $T(\vec{x}) = A\vec{x}$ are both $\mathbb{R}^n$. In this situation, one can regard T as *operating on* $\mathbb{R}^n$: it moves the vectors around in the same space.

Illustration 2.1.3: Matrix transformations of $\mathbb{R}^2$

In the first subsection we discussed the transformations defined by several 2×2 matrices, namely:

$$\begin{aligned} \text{Reflection:} \quad & A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \\ \text{Dilation:} \quad & A = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix} \\ \text{Identity:} \quad & A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ \text{Rotation:} \quad & A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\ \text{Shear:} \quad & A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

In each case, the associated matrix transformation $T(\vec{x}) = A\vec{x}$ has domain and codomain equal to $\mathbb{R}^2$. The range is also $\mathbb{R}^2$, as can be seen geometrically (what is the input for a given output?), or using the fact that the columns of A are not collinear (so they form a basis for $\mathbb{R}^2$).

It is worth noticing that none of the matrix transformations above repeat their outputs (so each is a one-to-one transformation). Each vector in the domain maps to one unique output in the codomain (that's not special, just the definition of being a function), and each vector in the codomain is mapped to by one unique input in the domain (so each is also an onto transformation).

Example 2.1.11: Matrix transformation example

Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$, and let $T(\vec{x}) = A\vec{x}$, so $T: \mathbb{R}^2 \to \mathbb{R}^3$ is a matrix transformation.

(a) Evaluate $T(\vec{u})$ for $\vec{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

(b) Let

$$\vec{b} = \begin{bmatrix} 7 \\ 5 \\ 7 \end{bmatrix}$$

Find a vector $\vec{v}$ in $\mathbb{R}^2$ such that $T(\vec{v}) = \vec{b}$. Is there more than one?

(c) Is T a one-to-one transformation?

(d) Is T an onto transformation?

Note: all of the above questions are intrinsic to the transformation T: they make sense to ask whether or not T is a matrix transformation. In this case, as T is in fact a matrix transformation, all of these questions will translate into questions about the corresponding matrix A.

Solution

(a) We evaluate $T(\vec{u})$ by substituting the definition of T in terms of matrix multiplication:

$$T\left(\begin{bmatrix} 3 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \boxed{\begin{bmatrix} 7 \\ 4 \\ 7 \end{bmatrix}}$$

(b) We want to find a vector $\vec{v}$ such that $\vec{b} = T(\vec{v}) = A\vec{v}$. In other words, we want to solve the matrix equation $A\vec{v} = \vec{b}$. We form an augmented matrix and row reduce:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix} \vec{v} = \begin{bmatrix} 7 \\ 5 \\ 7 \end{bmatrix} \xrightarrow{\text{augmented matrix}} \left[\begin{array}{cc|c} 1 & 1 & 7 \\ 0 & 1 & 5 \\ 1 & 1 & 7 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{array}\right].$$

This gives $x = 2$ and $y = 5$, so that there is a unique vector $\boxed{\vec{v} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}}$ such that $T(\vec{v}) = \vec{b}$.

(c) *Translation:* is there any vector $\vec{w}$ in $\mathbb{R}^3$ such that the solution set of $A\vec{v} = \vec{w}$ has more than one vector in it? The solution set of $A\vec{x} = \vec{w}$, if non-empty, is a translate of the solution set of $A\vec{v} = \vec{b}$ above, which has one vector in it. (Remember the solution sets for a given answer column $\vec{b}$ are all translations of the null space of A.) It follows that the solution set of $A\vec{v} = \vec{w}$ can have at most one vector. Therefore this function T is indeed $\boxed{\text{a one-to-one transformation}}$

(d) *Translation:* Can we find a vector $\vec{w}$ such that the matrix equation $A\vec{v} = \vec{w}$ is not consistent? Notice that if we take $\vec{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, (or any other vector with $x \neq z$), then the matrix equation $A\vec{v} = \vec{w}$ translates into the system of equations

$$\begin{cases} x + y = 1 \\ \quad\; y = 2 \\ x + y = 3 \end{cases}$$

which is clearly inconsistent. This shows that the function is $\boxed{\text{not an onto transformation}}$

Example 2.1.12: Non-matrix transformation example

Define a transformation $T: \mathbb{R}^2 \to \mathbb{R}^3$ by the formula

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} e^x \\ \cos(y) \\ e^x \end{bmatrix}$$

(a) Evaluate $T(\vec{u})$ for $\vec{u} = \begin{bmatrix} 1 \\ \pi \end{bmatrix}$

(b) Let

$$\vec{b} = \begin{bmatrix} 7 \\ 1 \\ 7 \end{bmatrix}$$

Find a vector $\vec{v}$ in $\mathbb{R}^2$ such that $T(\vec{v}) = \vec{b}$. Is there more than one?

(c) Is T a one-to-one transformation?

(d) Is T an onto transformation?

Note: we asked (almost) the exact same questions about a matrix transformation in the previous example. The point of this example is to illustrate the fact that the questions make sense for a transformation that has no hope of coming from a matrix. In this case, these questions do not translate into questions about a matrix; they have to be answered in some other way.

Solution

(a) We evaluate $T(\vec{u})$ using the defining formula:

$$T\left(\begin{bmatrix} 1 \\ \pi \end{bmatrix}\right) = \begin{bmatrix} e^1 \\ \cos(\pi) \\ e^1 \end{bmatrix} = \boxed{\begin{bmatrix} e \\ -1 \\ e \end{bmatrix}}$$

(b) We want to find an x and y that give $\begin{bmatrix} e^x \\ \cos(y) \\ e^x \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ 7 \end{bmatrix}$.

Solving $e^x = 7 \implies x = \ln(7)$ and $\cos(y) = 1 \implies y = 2\pi n$ for any integer n.

$$T\left(\begin{bmatrix} \ln(7) \\ 2\pi \end{bmatrix}\right) = \begin{bmatrix} e^{\ln(7)} \\ \cos(2\pi) \\ e^{\ln(7)} \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ 7 \end{bmatrix}$$

$$T\left(\begin{bmatrix} \ln(7) \\ 4\pi \end{bmatrix}\right) = \begin{bmatrix} e^{\ln(7)} \\ \cos(4\pi) \\ e^{\ln(7)} \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ 7 \end{bmatrix}$$

Clearly $\boxed{\text{there are infinitely many such vectors } \vec{v} \text{ in } \mathbb{R}^2 \text{ such that } T(\vec{v}) = \vec{b}}$

(c) The vector $\vec{b}$ from the previous part shows that T is clearly $\boxed{\text{not a one-to-one transformation}}$

(d) Since $\cos(y)$ is always between -1 and 1, the vector

$$\vec{w} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

is not in the range of T. (Alternatively, any vector in $\mathbb{R}^3$ such that $x \neq z$ will not be in the range of T.) Either way, this

transformation $\boxed{\text{is not onto/surjective}}$

2.1: Matrix Transformations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.1: Matrix Transformations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **3.2: One-to-one and Onto Transformations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.1.3: Practice Set

Exercise 2.1.3.1a

Consider the matrix $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 2 & 2 & 2 \end{bmatrix}$ and the matrix transformation $T(\vec{x}) = A\vec{x}$

(i) What is the domain and codomain of T?

(ii) Is T one-to-one?

(iii) Is T onto?

Answer

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

(i) Since A is 3×3, we know $T : \mathbb{R}^3 \to \mathbb{R}^3$ so the domain is $\mathbb{R}^3$ and codomain is $\mathbb{R}^3$

(ii) Since there is no pivot in the third column, T is not one-to-one.

For example, $T\left(\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = T\left(\begin{bmatrix} 2 \\ -4 \\ 2 \end{bmatrix}\right)$

(iii) Since there is no pivot in the third row, T is not onto.

For example, $A\vec{x} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ has no solution $\vec{x} \in \mathbb{R}^3$

Exercise 2.1.3.1b

Consider the matrix $B = \begin{bmatrix} 1 & 2 & 3 & 2 \\ 0 & 1 & 2 & 1 \\ 2 & 2 & 2 & 0 \end{bmatrix}$ and the matrix transformation $T(\vec{x}) = B\vec{x}$

(i) What is the domain and codomain of T?

(ii) Is T one-to-one?

(iii) Is T onto?

Answer

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

(i) Since B is 3×4, we know $T : \mathbb{R}^4 \to \mathbb{R}^3$ so the domain is $\mathbb{R}^4$ and codomain is $\mathbb{R}^3$

(ii) Since there is no pivot in the third column, T is not one-to-one.

For example, $T\left(\begin{bmatrix} 3 \\ -6 \\ 3 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = T\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right)$

(iii) Since there is a pivot in every row, T is onto.

Exercise 2.1.3.1c

Consider the matrix $C = \begin{bmatrix} 4 & 4 & 4 \\ 1 & 1 & 0 \\ 0 & 5 & 5 \\ 0 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$ and the matrix transformation $T(\vec{x}) = C\vec{x}$

(i) What is the domain and codomain of T?

(ii) Is T one-to-one?

(iii) Is T onto?

Answer

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

(i) Since C is 5×3, we know $T : \mathbb{R}^3 \to \mathbb{R}^5$ so the domain is $\mathbb{R}^3$ and codomain is $\mathbb{R}^5$

(ii) Since there is a pivot in every column, T is one-to-one.

(iii) Since there is no pivot in two of the rows, T is not onto.

For example, $C\vec{x} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ has no solution $\vec{x} \in \mathbb{R}^3$

Exercise 2.1.3.1d

Consider the matrix $D = \begin{bmatrix} 5 & 6 & 7 & 8 & 9 \\ 0 & 1 & 2 & 3 & 4 \\ 10 & 10 & 10 & 10 & 10 \end{bmatrix}$ and the matrix transformation $T(\vec{x}) = D\vec{x}$

(i) What is the domain and codomain of T?

(ii) Is T one-to-one?

(iii) Is T onto?

Answer

$$\begin{bmatrix} 5 & 6 & 7 & 8 & 9 \\ 0 & 1 & 2 & 3 & 4 \\ 10 & 10 & 10 & 10 & 10 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & -1 & -2 & -3 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(i) Since D is 3×5, we know $T : \mathbb{R}^5 \to \mathbb{R}^3$ so the domain is $\mathbb{R}^5$ and codomain is $\mathbb{R}^3$

(ii) Since there is no pivot in three of the columns, T is not one-to-one.

For example, $T\left(\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = T\left(\begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}\right)$

(iii) Since there is no pivot in the third row, T is not onto.

For example, $D\vec{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ has no solution $\vec{x} \in \mathbb{R}^5$

Exercise 2.1.3.1e

Consider the matrix $E = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 2 & 0 & 3 \\ 2 & 0 & 1 & -4 \\ 1 & 3 & 0 & 2 \end{bmatrix}$ and the matrix transformation $T(\vec{x}) = E\vec{x}$

(i) What is the domain and codomain of T?

(ii) Is T one-to-one?

(iii) Is T onto?

Answer

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

(i) Since E is 4×4, we know $T : \mathbb{R}^4 \to \mathbb{R}^4$ so the domain is $\mathbb{R}^4$ and codomain is $\mathbb{R}^4$

(ii) Since there is a pivot in every column, T is one-to-one.

(iii) Since there is a pivot in every row, T is onto.

Exercise 2.1.3.2a

Consider the matrix $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 3 & h \end{bmatrix}$ and the matrix transformation $T(\vec{x}) = A\vec{x}$

(i) What value(s) of h make T a one-to-one transformation?

(ii) What value(s) of h make T an onto transformation?

Answer

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 3 & h \end{bmatrix} \xrightarrow{\text{echelon form}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & h+1 \end{bmatrix}$$

(i) If $h + 1 \neq 0$ there will be a pivot in every column, so T is **one-to-one when $h \neq -1$**

(ii) If $h + 1 \neq 0$ there will be a pivot in every row, so T is **onto when $h \neq -1$**

Exercise 2.1.3.2b

Consider the matrix $B = \begin{bmatrix} 1 & 2 & 0 & h \\ 0 & 1 & 2 & h \\ 1 & 3 & 4 & h \end{bmatrix}$ and the matrix transformation $T(\vec{x}) = B\vec{x}$

(i) What value(s) of h make T a one-to-one transformation?

(ii) What value(s) of h make T an onto transformation?

Answer

$$\begin{bmatrix} 1 & 2 & 0 & h \\ 0 & 1 & 2 & h \\ 1 & 3 & 4 & h \end{bmatrix} \xrightarrow{\text{echelon form}} \begin{bmatrix} 1 & 0 & 0 & -3h \\ 0 & 1 & 0 & 2h \\ 0 & 0 & 1 & -\frac{1}{2}h \end{bmatrix}$$

(i) No matter what h is, there will not be a pivot in the fourth column, so T is **never one-to-one.**

(ii) No matter what h is, there will be a pivot in every row, so T is **always onto.**

Exercise 2.1.3.2c

Consider the matrix $C = \begin{bmatrix} 1 & 1 & -11 \\ 2 & 0 & 6 \\ h & h & 7 \\ 1 & 0 & 6 \end{bmatrix}$ and the matrix transformation $T(\vec{x}) = C\vec{x}$

(i) What value(s) of h make T a one-to-one transformation?

(ii) What value(s) of h make T an onto transformation?

Answer

$$\begin{bmatrix} 1 & 1 & -11 \\ 2 & 0 & 6 \\ h & h & 7 \\ 1 & 0 & 6 \end{bmatrix} \xrightarrow{\text{echelon form}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

(i) No matter what h is, there will be a pivot in every column, so T is **always one-to-one.**

(ii) No matter what h is, there will never be a pivot in the fourth row, so T is **never onto.**

Exercise 2.1.3.2d

Consider the matrix $D = \begin{bmatrix} h & h & 0 \\ 1 & 2 & 1 \\ 3 & 4 & 1 \end{bmatrix}$ and the matrix transformation $T(\vec{x}) = D\vec{x}$

(i) What value(s) of h make T a one-to-one transformation?

(ii) What value(s) of h make T an onto transformation?

Answer

$$\begin{bmatrix} h & h & 0 \\ 1 & 2 & 1 \\ 3 & 4 & 1 \end{bmatrix} \xrightarrow{\text{echelon form}} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

(i) No matter what h is, there will never be a pivot in the third column, so T is **never one-to-one.**

(ii) No matter what h is, there will never be a pivot in the third row, so T is **never onto.**

Exercise 2.1.3.3a

Create a 2×2 matrix A for which the matrix transformation $T(\vec{x}) = A\vec{x}$ is onto but not one-to-one (or explain why it is impossible to do so)

Answer

No such matrix can exist, since T ONTO only occurs with a pivot in every row, which forces a pivot in every column, and thus makes T ONE-TO-ONE.

Exercise 2.1.3.3b

Create a 2×2 matrix A for which the matrix transformation $T(\vec{x}) = A\vec{x}$ is one-to-one but not onto (or explain why it is impossible to do so)

Answer

No such matrix can exist, since T ONE-TO-ONE only occurs with a pivot in every column, which forces a pivot in every row, and thus makes T ONTO.

Exercise 2.1.3.3c

Create a 2×3 matrix A for which the matrix transformation $T(\vec{x}) = A\vec{x}$ is onto but not one-to-one (or explain why it is impossible to do so)

Answer

We need a matrix with a pivot in every row, but not every column, like $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

Exercise 2.1.3.3d

Create a 2×3 matrix A for which the matrix transformation $T(\vec{x}) = A\vec{x}$ is one-to-one but not onto (or explain why it is impossible to do so)

Answer

There is no matrix which is 2×3 with a pivot in every column, so it is impossible to be one-to-one. (regardless of onto or not)

Exercise 2.1.3.3e

Create a 3×2 matrix A for which the matrix transformation $T(\vec{x}) = A\vec{x}$ is onto but not one-to-one (or explain why it is impossible to do so)

Answer

There is no matrix which is 3×2 with a pivot in every row, so it is impossible to be onto. (regardless of one-to-one or not)

Exercise 2.1.3.3f

Create a 3×2 matrix A for which the matrix transformation $T(\vec{x}) = A\vec{x}$ is one-to-one but not onto (or explain why it is impossible to do so)

Answer

We need a matrix with a pivot in every column, but not every row, like $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

Exercise 2.1.3.4

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If A is 3×4 and $T(\vec{x}) = A\vec{x}$ is onto, then it is also one-to-one.

(b) If A is 4×4 and $T(\vec{x}) = A\vec{x}$ is onto, then it is also one-to-one.

(c) If A is 4×4 and $T(\vec{x}) = A\vec{x}$ is one-to-one, then it is also onto.

(d) If A is 5×4 and $T(\vec{x}) = A\vec{x}$ is one-to-one, then it is also onto.

(e) If A is $m \times n$ and $T(\vec{x}) = A\vec{x}$, then $T(\vec{u} + \vec{v}) \neq T(\vec{u}) + T(\vec{v})$ for some $\vec{u}, \vec{v} \in \mathbb{R}^n$.

(f) If A is $m \times n$ and $T(\vec{x}) = A\vec{x}$, then $T(r\vec{u}) = rT(\vec{u})$ for all $\vec{u} \in \mathbb{R}^n$ and $r \in \mathbb{R}$.

Answer

(a) **Always False**

It is impossible to have a pivot in every column of a 3×4 matrix.

(b) **Always True**

If there is a pivot in every row, then there must be a pivot in every column, so T must be one-to-one.

(c) **Always True**

If there is a pivot in every column, then there must be a pivot in every row, so T must be onto.

(d) **Always False**

It is impossible to have a pivot in every row of a 5×4 matrix.

(e) **Always False**

$T(\vec{u} + \vec{v}) = A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v} = T(\vec{u}) + T(\vec{v})$ always holds.

(f) **Always True**

$T(r\vec{u}) = A(r\vec{u}) = r(A\vec{u}) = rT(\vec{u})$ always holds.

2.1.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.2: Linear Transformations

Objectives

1. Learn how to verify that a transformation is linear, or prove that a transformation is not linear.
2. Understand the relationship between linear transformations and matrix transformations.
3. *Recipe:* compute the matrix of a linear transformation.
4. *Theorem:* linear transformations and matrix transformations.
5. *Notation:* the ***standard coordinate vectors*** $\vec{e}_1, \vec{e}_2, \ldots$.
6. *Vocabulary words:* ***linear transformation***, ***standard matrix***, ***identity matrix***.

In the last section, we studied the geometry of matrices by regarding them as functions, i.e., by considering the associated *matrix transformations*. We defined some vocabulary (domain, codomain, range), and asked a number of natural questions about a transformation. For a matrix transformation, these translate into questions about matrices, which we have many tools to answer.

In this section, we make a change in perspective. Suppose that we are given a *transformation* that we would like to study. If we can prove that our transformation is a matrix transformation, then we can use linear algebra to study it. This raises two important questions:

1. How can we tell if a transformation is a matrix transformation?
2. If our transformation is a matrix transformation, how do we find its matrix?

For example, we saw previously that the matrix transformation

$$T\colon \mathbb{R}^2 \to \mathbb{R}^2 \qquad T(\vec{x}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \vec{x}$$

is a counterclockwise rotation of the plane by 90°. However, we could have *defined* T in this way:

$$T\colon \mathbb{R}^2 \to \mathbb{R}^2 \qquad T(\vec{x}) = \text{the counterclockwise rotation of } \vec{x} \text{ by } 90^\circ.$$

Given this definition, it is not at all obvious that T is a matrix transformation, or what matrix it is associated to.

This page titled 2.2: Linear Transformations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **3.3: Linear Transformations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.2.1: Linear Transformations

In this section, we introduce the class of transformations that come from matrices.

Definition: Linear Transformation

A ***linear transformation*** is a transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ satisfying

$$T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$$
$$T(c\vec{u}) = cT(\vec{u})$$

for all vectors $\vec{u}, \vec{v}$ in $\mathbb{R}^n$ and all scalars c.

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a matrix transformation: $T(\vec{x}) = A\vec{x}$ for an $m \times n$ matrix A. Clearly

$$T(\vec{u}+\vec{v}) = A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v} = T(\vec{u}) + T(\vec{v})$$
$$T(c\vec{u}) = A(c\vec{u}) = cA\vec{u} = cT(\vec{u})$$

for all vectors $\vec{u}, \vec{v}$ in $\mathbb{R}^n$ and all scalars c. Since a matrix transformation satisfies the two defining properties, it is a linear transformation

We will see soon that the opposite is true: every linear transformation is a matrix transformation; we just haven't computed its matrix yet.

Fact 2.2.1.1: Facts about Linear Transformations

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then:

1. $T(\vec{0}) = \vec{0}$.
2. For any vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ in $\mathbb{R}^n$ and scalars $c_1, c_2, \ldots, c_k$, we have

$$T(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k) = c_1T(\vec{v}_1) + c_2T(\vec{v}_2) + \cdots + c_kT(\vec{v}_k)$$

Proof

1. Since $\vec{0} = -\vec{0}$, we have

$$T(\vec{0}) = T(-\vec{0}) = -T(\vec{0})$$

The only vector $\vec{w}$ such that $\vec{w} = -\vec{w}$ is the zero vector.

2. Let us suppose for simplicity that $k = 2$. Then

$$\begin{aligned} T(c_1\vec{v}_1 + c_2\vec{v}_2) &= T(c_1\vec{v}_1) + T(c_2\vec{v}_2) && \text{first property} \\ &= c_1T(\vec{v}_1) + c_2T(\vec{v}_2) && \text{second property} \end{aligned}$$

In engineering, the second fact is called the *superposition principle*; it should remind you of the distributive property. For example, $T(c\vec{u} + d\vec{v}) = cT(\vec{u}) + dT(\vec{v})$ for any vectors $\vec{u}, \vec{v}$ and any scalars c, d. To restate the first fact:

Note 2.2.1.1

A linear transformation necessarily takes the zero vector to the zero vector.

Example 2.2.1.1: A non-linear transformation

Define $T: \mathbb{R} \to \mathbb{R}$ by $T(x) = x + 1$

Is T a linear transformation?

Solution

We have $T(0) = 0 + 1 = 1$. Since any linear transformation necessarily takes zero to zero by the above important note **2.2.1.1**, we conclude that $\boxed{T \text{ is not a linear transformation}}$ (even though its graph is a line).

Note: in this case, it was not necessary to check explicitly that T does not satisfy both defining properties, Definition **2.2.1.1**: since $T(0) = 0$ is a consequence of these properties, at least one of them must not be satisfied. (In fact, this T satisfies neither.)

$T(x + y) = (x + y) + 1$ is not equal to $T(x) + T(y) = (x + 1) + (y + 1)$

and $T(cx) = (cx) + 1$ is not equal to $cT(x) = c(x + 1)$

Example 2.2.1.2: Verifying linearity: dilation

Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\vec{x}) = 1.5\vec{x}$

Verify that T is linear.

Solution

We have to check the defining properties of a linear transformation for *all* vectors $\vec{u}, \vec{v}$ and *all* scalars c. In other words, we have to treat $\vec{u}, \vec{v}$, and c as *unknowns*. The only thing we are allowed to use is the definition of T.

$$T(\vec{u} + \vec{v}) = 1.5(\vec{u} + \vec{v}) = 1.5\vec{u} + 1.5\vec{v} = T(\vec{u}) + T(\vec{v})$$
$$T(c\vec{u}) = 1.5(c\vec{u}) = c(1.5\vec{u}) = cT(\vec{u})$$

Since T satisfies both defining properties, $\boxed{T \text{ is linear}}$

Note: we know from our previous subsection that T is a matrix transformation: in fact,

$$T(\vec{x}) = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix} \vec{x}$$

Since a matrix transformation is a linear transformation, this is another proof that T is linear.

Example 2.2.1.3: Verifying linearity: rotation

Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ by

$$T(\vec{x}) = \text{ the vector } \vec{x} \text{ rotated counterclockwise by the angle } \theta$$

Verify that T is linear.

Solution

Since T is defined geometrically, we give a geometric argument. For the first property, $T(\vec{u}) + T(\vec{v})$ is the sum of the vectors obtained by rotating $\vec{u}$ and $\vec{v}$ by θ. On the other side of the equation, $T(\vec{u} + \vec{v})$ is the vector obtained by rotating the sum of the vectors $\vec{u}$ and $\vec{v}$. But it does not matter whether we sum or rotate first, as the following picture shows.

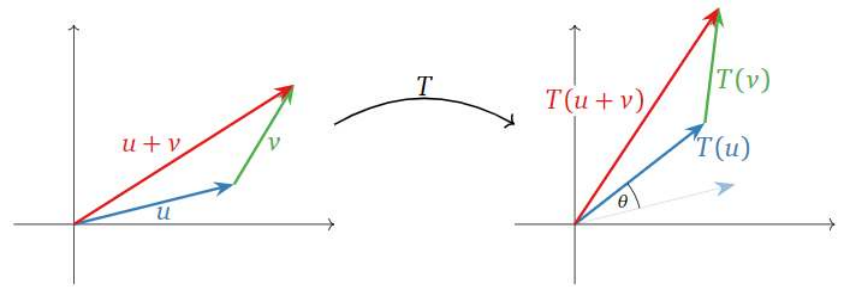


Figure **2.2.1.1**

For the second property, $cT(\vec{u})$ is the vector obtained by rotating $\vec{u}$ by the angle θ, then changing its length by a factor of c (reversing direction of $c < 0$). On the other hand, $T(c\vec{u})$ first changes the length of $\vec{u}$ by a factor of c, then rotates. But it does not

matter in which order we do these two operations.

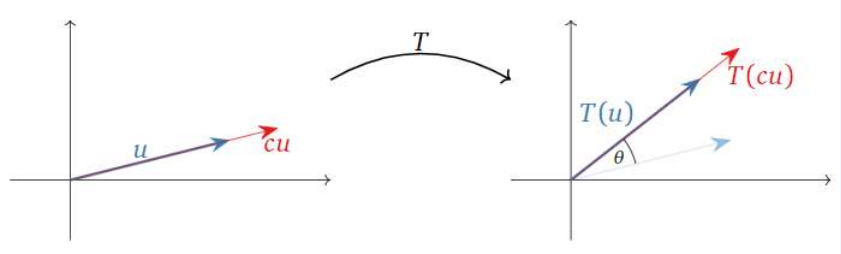


Figure 2.2.1.2

This verifies that **T is a linear transformation**

Note however that it is not at all obvious that or how T can be expressed as multiplication by a matrix.

Example 2.2.1.4: A transformation defined by a formula

Define $T:\mathbb{R}^2 \to \mathbb{R}^3$ by the formula

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x-y \\ y \\ x \end{bmatrix}$$

Verify that T is linear.

Solution

We have to check the defining properties, for *all* vectors $\vec{u},\vec{v}$ and *all* scalars c. In other words, we have to treat $\vec{u},\vec{v}$, and c as *unknowns*; the only thing we are allowed to use is the definition of T. Since T is defined in terms of the coordinates of $\vec{u},\vec{v}$, we need to give those names as well; say $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ and $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. For the first property, we have

$$\begin{aligned} T\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} u_1+v_1 \\ u_2+v_2 \end{bmatrix}\right) &= \begin{bmatrix} 3(u_1+v_1)-(u_2+v_2) \\ u_2+v_2 \\ u_1+v_1 \end{bmatrix} \\ &= \begin{bmatrix} (3u_1-u_2)+(3v_1-v_2) \\ u_2+v_2 \\ u_1+v_1 \end{bmatrix} \\ &= \begin{bmatrix} 3u_1-u_2 \\ u_2 \\ u_1 \end{bmatrix} + \begin{bmatrix} 3v_1-v_2 \\ v_2 \\ v_1 \end{bmatrix} = T\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) + T\left(\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}\right) \end{aligned}$$

For the second property,

$$\begin{aligned} T\left(c\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} cu_1 \\ cu_2 \end{bmatrix}\right) &= \begin{bmatrix} 3(cu_1)-(cu_2) \\ cu_2 \\ cu_1 \end{bmatrix} \\ &= \begin{bmatrix} c(3u_1-u_2) \\ cu_2 \\ cu_1 \end{bmatrix} = c\begin{bmatrix} 3u_1-u_2 \\ u_2 \\ u_1 \end{bmatrix} = cT\left(\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right) \end{aligned}$$

Since T satisfies the defining properties, **T is a linear transformation**

Note: we will see shortly that T is a matrix transformation satisfying

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 3 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3x-y \\ y \\ x \end{bmatrix}$$

One can show that, if a transformation is defined by formulas in the coordinates as in the above example, then the transformation is linear if and only if each coordinate is a linear expression in the variables with no constant term.

Illustration 2.2.1.1: A translation

Define $T: \mathbb{R}^3 \to \mathbb{R}^3$ by

$$T(\vec{x}) = \vec{x} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

This kind of transformation is called a ***translation***. As in a previous Example 2.2.1.1, this T is not linear, because $T(\vec{0})$ is not the zero vector. It is however both a one-to-one and onto transformation (just not a linear one!)

Example 2.2.1.5: More non-linear transformations

Verify that the following transformations from $\mathbb{R}^2$ to $\mathbb{R}^2$ are not linear:

$$T_1\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} |x| \\ y \end{bmatrix} \quad T_2\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} xy \\ y \end{bmatrix} \quad T_3\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x+1 \\ x-2y \end{bmatrix}$$

Solution

In order to verify that a transformation T is *not* linear, we have to show that T does not satisfy *at least one* of the two defining properties in Definition 2.2.1.1. For the first, the negation of the statement "$T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ for all vectors $\vec{u}, \vec{v}$" is "there exists at least one pair of vectors $\vec{u}, \vec{v}$ such that $T(\vec{u}+\vec{v}) \neq T(\vec{u}) + T(\vec{v})$." In other words, it suffices to find *one example* of a pair of vectors $\vec{u}, \vec{v}$ such that $T(\vec{u}+\vec{v}) \neq T(\vec{u}) + T(\vec{v})$. Likewise, for the second, the negation of the statement "$T(c\vec{u}) = cT(\vec{u})$ for all vectors $\vec{u}$ and all scalars c" is "there exists some vector $\vec{u}$ and some scalar c such that $T(c\vec{u}) \neq cT(\vec{u})$." In other words, it suffices to find *one* vector $\vec{u}$ and *one* scalar c such that $T(c\vec{u}) \neq cT(\vec{u})$.

For the first transformation, we note that

$$T_1\left(-3\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = T_1\left(\begin{bmatrix} -3 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} |-3| \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

but that

$$-3T_1\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = -3\begin{bmatrix} |1| \\ 0 \end{bmatrix} = -3\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 0 \end{bmatrix}$$

Therefore, T_1 does not satisfy the second property (it also fails the first property for a suitable choice of vectors) and is **not a linear transformation**

For the second transformation, we note that

$$T_2\left(2\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = T_2\left(\begin{bmatrix} 2 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 2 \cdot 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

but that

$$2T_2\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = 2\begin{bmatrix} 1 \cdot 1 \\ 1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

Therefore, T_2 does not satisfy the second property (it also fails the first property for a suitable choice of vectors) and is **not a linear transformation**

For the third transformation, we observe that

$$T_3\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2(0)+1 \\ 0-2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Since T_3 does not take the zero vector to the zero vector, it is **not a linear transformation**

When deciding whether a transformation T is linear, generally the first thing to do is to check whether $T(\vec{0}) = \vec{0}$; if not, T is automatically not linear. Note however that the non-linear transformations T_1 and T_2 of the above example do take the zero vector to the zero vector. So you cannot use this shortcut to prove a transformation is linear!

2.2.1: Linear Transformations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.3: Linear Transformations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.2.2: Matrix Representations

In the next subsection, we will present the relationship between linear transformations and matrix transformations. Before doing so, we need the following important notation.

Definition: Standard Coordinate Vectors

The ***standard coordinate vectors in*** $\mathbb{R}^n$ are the n vectors

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \cdots, \quad \vec{e}_{n-1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

The ith entry of $\vec{e}_i$ is equal to 1, and the other entries are zero.

From now on, for the rest of the book, we will use the symbols $\vec{e}_1, \vec{e}_2, \ldots$ to denote the standard coordinate vectors.

(In physics and other applications that are restricted to $\mathbb{R}^3$, you will often see $\vec{e}_1 = \vec{i}, \vec{e}_2 = \vec{j}$, and $\vec{e}_3 = \vec{k}$. This notation is not particularly useful in linear algebra and will not be used.)

There is an ambiguity in this notation: one has to know from context that $\vec{e}_1$ is meant to have n entries. That is, the vectors

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

may both be denoted $\vec{e}_1$, depending on whether we are discussing vectors in $\mathbb{R}^2$ or in $\mathbb{R}^3$.

The standard coordinate vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$ are pictured below.

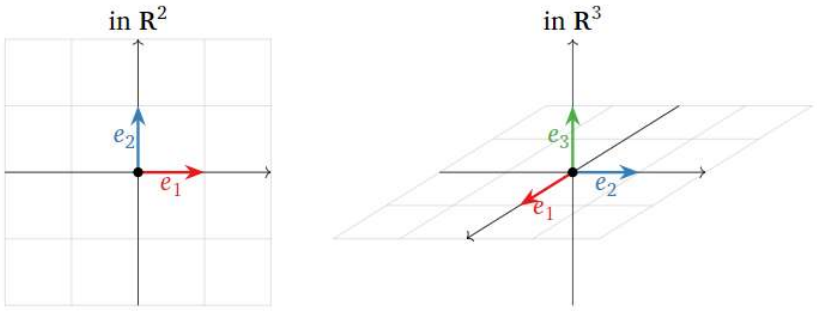


Figure 2.2.2.3

These are the vectors of length 1 that point in the positive directions of each of the axes.

Fact 2.2.2.2: Multiplying a matrix by the standard coordinate vectors

If A is an $m \times n$ matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, then $A\vec{e}_i = \vec{v}_i$ for each $i = 1, 2, \ldots, n$:

$$\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \vec{e}_i = \vec{v}_i$$

In other words, multiplying a matrix by $\vec{e}_i$ simply selects its ith column.

For example,

$$\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}1\\4\\7\end{bmatrix}\quad\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix}=\begin{bmatrix}2\\5\\8\end{bmatrix}\quad\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix}=\begin{bmatrix}3\\6\\9\end{bmatrix}$$

Now we can prove that every linear transformation is a matrix transformation, and we will show how to compute the matrix.

Theorem 2.2.2.1: The Matrix of a Linear Transformation

Let $T:\mathbb{R}^n\to\mathbb{R}^m$ be a linear transformation. Let A be the $m\times n$ matrix

$$A=\begin{bmatrix}|&|&&|\\T(\vec{e}_1)&T(\vec{e}_2)&\cdots&T(\vec{e}_n)\\|&|&&|\end{bmatrix}$$

Then T is the matrix transformation associated with A: that is, $T(\vec{x})=A\vec{x}$.

Proof

We suppose for simplicity that T is a transformation from $\mathbb{R}^3$ to $\mathbb{R}^2$. Let A be the matrix given in the statement of the theorem. Then

$$\begin{aligned}T\left(\begin{bmatrix}x\\y\\z\end{bmatrix}\right)&=T\left(x\begin{bmatrix}1\\0\\0\end{bmatrix}+y\begin{bmatrix}0\\1\\0\end{bmatrix}+z\begin{bmatrix}0\\0\\1\end{bmatrix}\right)\\&=T(x\vec{e}_1+y\vec{e}_2+z\vec{e}_3)\\&=xT(\vec{e}_1)+yT(\vec{e}_2)+zT(\vec{e}_3)\\&=\begin{bmatrix}|&|&|\\T(\vec{e}_1)&T(\vec{e}_2)&T(\vec{e}_3)\\|&|&|\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}\\&=A\begin{bmatrix}x\\y\\z\end{bmatrix}\end{aligned}$$

The matrix A in the above theorem is called the ***standard matrix*** for T. The columns of A are the vectors obtained by evaluating T on the n standard coordinate vectors in $\mathbb{R}^n$. To summarize part of the theorem:

Note 2.2.2.3

Matrix transformations are the same as linear transformations.

Example 2.2.2.6: The matrix of a dilation

Define $T:\mathbb{R}^2\to\mathbb{R}^2$ by $T(\vec{x})=1.5\vec{x}$

Find the standard matrix A for T.

Solution

The columns of A are obtained by evaluating T on the standard coordinate vectors $\vec{e}_1,\vec{e}_2$.

$$\left.\begin{aligned}T(\vec{e}_1)=1.5\vec{e}_1=\begin{bmatrix}1.5\\0\end{bmatrix}\\T(\vec{e}_2)=1.5\vec{e}_2=\begin{bmatrix}0\\1.5\end{bmatrix}\end{aligned}\right\}\Longrightarrow\boxed{A=\begin{bmatrix}1.5&0\\0&1.5\end{bmatrix}}$$

This matches what we saw in our dilation examples earlier.

Example 2.2.2.7: The matrix of a rotation

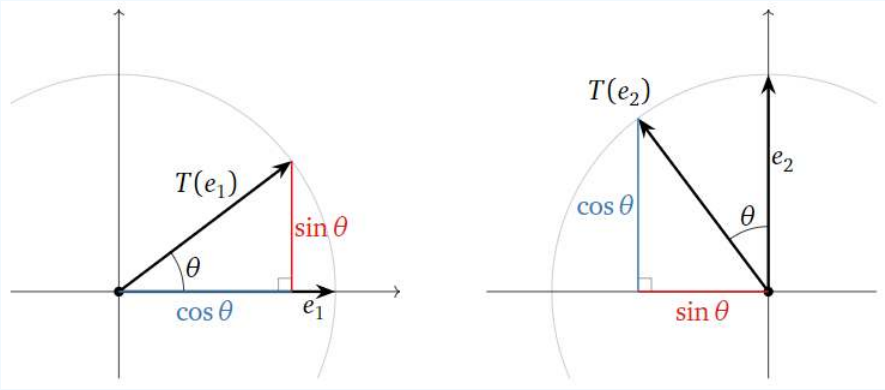
Define $T: \mathbb{R}^2 \to \mathbb{R}^2$ by

$$T(\vec{x}) = \text{ the vector } \vec{x} \text{ rotated counterclockwise by the angle } \theta$$

Find the standard matrix for T.

Solution

The columns of A are obtained by evaluating T on the standard coordinate vectors $\vec{e}_1, \vec{e}_2$. In order to compute the entries of $T(\vec{e}_1)$ and $T(\vec{e}_2)$, we have to do some trigonometry.



We see from the picture that

$$\left.\begin{aligned} T(\vec{e}_1) &= \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \\ T(\vec{e}_2) &= \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \end{aligned}\right\} \Longrightarrow \boxed{A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}$$

We saw in the above example that the matrix for counterclockwise rotation of the plane by an angle of θ is

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Plugging in $\theta = 90°$ gives us the matrix $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ which once again matches with our earlier rotation example!

Example 2.2.2.8: A transformation defined by a formula

Define $T: \mathbb{R}^2 \to \mathbb{R}^3$ by the formula

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 3x - y \\ y \\ x \end{bmatrix}$$

Find the standard matrix for T.

Solution

We substitute the standard coordinate vectors into the formula defining T:

$$\left.\begin{aligned} T(\vec{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) &= \begin{bmatrix} 3(1) - 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \\ T(\vec{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) &= \begin{bmatrix} 3(0) - 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \end{aligned}\right\} \Longrightarrow \boxed{A = \begin{bmatrix} 3 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}}$$

This is how the matrix was determined earlier in Example 2.2.4.

Example 2.2.2.9: A transformation defined in steps

Let $T: \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation that reflects over the xy-plane and then projects onto the yz-plane. What is the standard matrix for T?

Solution

This transformation is described geometrically, in two steps. To find the columns of A, we need to follow the standard coordinate vectors through each of these steps.

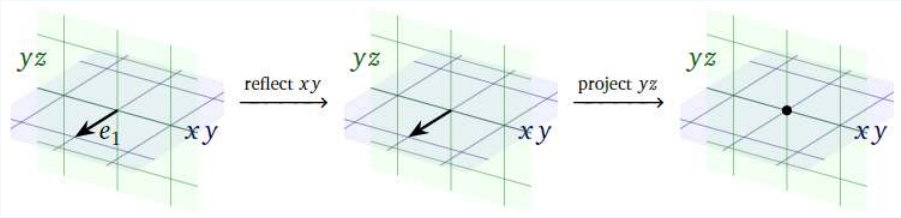


Figure 2.2.2.4

Since $\vec{e}_1$ lies on the xy-plane, reflecting over the xy-plane does not move $\vec{e}_1$. Since $\vec{e}_1$ is perpendicular to the yz-plane, projecting $\vec{e}_1$ onto the yz-plane sends it to zero. Therefore,

$$T(\vec{e}_1) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

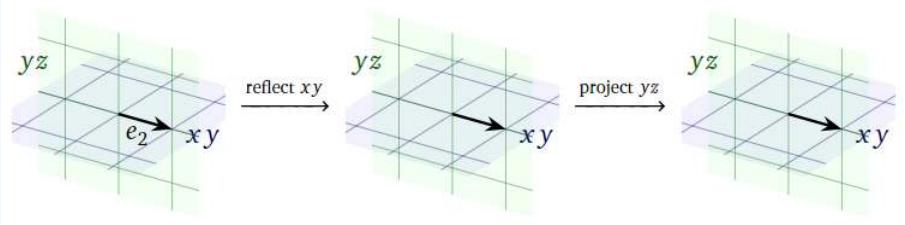


Figure 2.2.2.5

Since $\vec{e}_2$ lies on the xy-plane, reflecting over the xy-plane does not move $\vec{e}_2$. Since $\vec{e}_2$ lies on the yz-plane, projecting onto the yz-plane does not move $\vec{e}_2$ either. Therefore,

$$T(\vec{e}_2) = \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

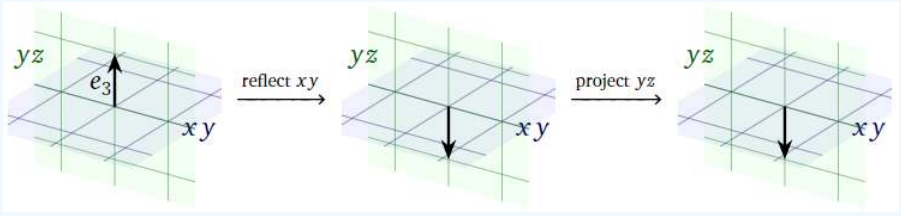


Figure 2.2.2.6

Since $\vec{e}_3$ is perpendicular to the xy-plane, reflecting over the xy-plane takes $\vec{e}_3$ to its negative. Since $-\vec{e}_3$ lies on the yz-plane, projecting onto the yz-plane does not move it. Therefore,

$$T(\vec{e}_3) = -\vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

Now we have computed all three columns of A:

$$\left.\begin{aligned} T(\vec{e}_1) &= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \\ T(\vec{e}_2) &= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\ T(\vec{e}_3) &= \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \end{aligned}\right\} \Longrightarrow \boxed{A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}}$$

Recall that the *identity transformation* is the transformation $\mathcal{I}: \mathbb{R}^n \to \mathbb{R}^n$ defined by $\mathcal{I}(\vec{x}) = \vec{x}$ for every vector $\vec{x}$.

Example 2.2.2.10: The standard matrix of the identity transformation

Verify that the identity transformation $\mathcal{I}: \mathbb{R}^n \to \mathbb{R}^n$ is linear, and compute its standard matrix.

Solution

We verify the two defining properties of linear transformations. Let $\vec{u}, \vec{v}$ be vectors in $\mathbb{R}^n$. Then

$$\mathcal{I}(\vec{u} + \vec{v}) = \vec{u} + \vec{v} = \mathcal{I}(\vec{u}) + \mathcal{I}(\vec{v})$$

If c is a scalar, then

$$\mathcal{I}(c\vec{u}) = c\vec{u} = c\mathcal{I}(\vec{u})$$

Since $\mathcal{I}$ satisfies the two defining properties, it is a linear transformation.

Now that we know that $\mathcal{I}$ is linear, it makes sense to compute its standard matrix. For each standard coordinate vector $\vec{e}_i$, we have $\mathcal{I}(\vec{e}_i) = \vec{e}_i$. In other words, the columns of the standard matrix of $\mathcal{I}$ are the standard coordinate vectors, so the standard matrix is the identity matrix

$$\boxed{I = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}}$$

Definition: Identity Matrix

The $n \times n$ ***identity matrix*** is the matrix I whose columns are the n standard coordinate vectors in $\mathbb{R}^n$:

$$I = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

We computed earlier that the matrix of the identity transform is the identity matrix: for every $\vec{x}$ in $\mathbb{R}^n$,

$$\vec{x} = \mathcal{I}(\vec{x}) = I\vec{x}$$

Therefore, $I\vec{x} = \vec{x}$ for all vectors $\vec{x}$: the product of the identity matrix and a vector is the same vector.

2.2.2: Matrix Representations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.3: Linear Transformations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.2: Practice Set

Exercise 2.2.1a

Let $T:\mathbb{R}^2\to\mathbb{R}^3$ be defined by $T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}2x+5y\\x-3y\\y\end{bmatrix}$

(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x})=A\vec{x}$? If so find A.

(iii) Is T one-to-one and/or onto?

Answer

(i) $T\left(\begin{bmatrix}x+a\\y+b\end{bmatrix}\right)=\begin{bmatrix}2(x+a)+5(y+b)\\(x+a)-3(y+b)\\(y+b)\end{bmatrix}=\begin{bmatrix}2x+5y\\x-3y\\y\end{bmatrix}+\begin{bmatrix}2a+5b\\a-3b\\b\end{bmatrix}=T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)+T\left(\begin{bmatrix}a\\b\end{bmatrix}\right)$

and $T\left(\begin{bmatrix}rx\\ry\end{bmatrix}\right)=\begin{bmatrix}2rx+5ry\\rx-3ry\\ry\end{bmatrix}=r\begin{bmatrix}2x+5y\\x-3y\\y\end{bmatrix}=rT\left(\begin{bmatrix}x\\y\end{bmatrix}\right)$

Therefore T **is a linear transformation.**

(ii) $T(\vec{e}_1)=T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}2\cdot1+5\cdot0\\1-3\cdot0\\0\end{bmatrix}=\begin{bmatrix}2\\1\\0\end{bmatrix}$ and $T(\vec{e}_2)=T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}2\cdot0+5\cdot1\\0-3\cdot1\\1\end{bmatrix}=\begin{bmatrix}5\\-3\\1\end{bmatrix}$

So the matrix representation of T is $A=\begin{bmatrix}2&5\\1&-3\\0&1\end{bmatrix}$

(iii) $\begin{bmatrix}2&5\\1&-3\\0&1\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0\\0&1\\0&0\end{bmatrix}$

Since there is a pivot in every column, T **is one-to-one.** Since there is no pivot in the third row T **is not onto.**

Exercise 2.2.1b

Let $T:\mathbb{R}^2\to\mathbb{R}^4$ be defined by $T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}x-y\\x-1\\0\\y+2\end{bmatrix}$

(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x})=A\vec{x}$? If so find A.

(iii) Is T one-to-one and/or onto?

Answer

(i) $T\left(\begin{bmatrix}x+a\\y+b\end{bmatrix}\right)=\begin{bmatrix}(x+a)-(y+b)\\(x+a)-1\\0\\(y+b)+2\end{bmatrix}$ but $T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)+T\left(\begin{bmatrix}a\\b\end{bmatrix}\right)=\begin{bmatrix}x-y\\x-1\\0\\y+2\end{bmatrix}+\begin{bmatrix}a-b\\a-1\\0\\b+2\end{bmatrix}=\begin{bmatrix}(x+a)-(y+b)\\(x+a)-2\\0\\(y+b)+4\end{bmatrix}$

and $T\left(\begin{bmatrix}rx\\ry\end{bmatrix}\right)=\begin{bmatrix}rx-ry\\rx-1\\0\\ry+2\end{bmatrix}$ but $rT\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=r\begin{bmatrix}x-y\\x-1\\0\\y+2\end{bmatrix}=\begin{bmatrix}rx-ry\\rx-r\\0\\ry+2r\end{bmatrix}$

Neither of these match, so T **is not a linear transformation.**

(ii) All matrix transformations must be linear transformations, so **there is no matrix** A satisfying $T(\vec{x}) = A\vec{x}$

(iii) We cannot use a matrix, but can still check if T is one-to-one the hard way:

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) \implies \begin{bmatrix} x-y \\ x-1 \\ 0 \\ y+2 \end{bmatrix} = \begin{bmatrix} a-b \\ a-1 \\ 0 \\ b+2 \end{bmatrix}$$

Since $x-1 = a-1 \implies x = a$ and $y+2 = b+2 \implies y = b$ we know T **is one-to-one.**

It is very obvious that T **is not onto**, since there is no solution to $T(\vec{x}) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$

Exercise 2.2.1c

Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} y-z \\ x+z \end{bmatrix}$

(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x}) = A\vec{x}$? If so find A.

(iii) Is T one-to-one and/or onto?

Answer

(i) $T\left(\begin{bmatrix} x+a \\ y+b \\ z+c \end{bmatrix}\right) = \begin{bmatrix} (y+b)-(z+c) \\ (x+a)+(z+c) \end{bmatrix} = \begin{bmatrix} y-z \\ x+z \end{bmatrix} + \begin{bmatrix} b-c \\ a+c \end{bmatrix} = T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) + T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right)$

and $T\left(\begin{bmatrix} rx \\ ry \\ rz \end{bmatrix}\right) = \begin{bmatrix} ry-rz \\ rx+rz \end{bmatrix} = r\begin{bmatrix} y-z \\ x+z \end{bmatrix} = rT\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)$

Therefore T **is a linear transformation.**

(ii) $T(\vec{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0-0 \\ 1+0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $T(\vec{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1-0 \\ 0+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $T(\vec{e}_3) = T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0-1 \\ 0+1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

So the matrix representation of T is $A = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$

(iii) $\begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$

Since there is no pivot in the third column, T **is not one-to-one.** Since there is a pivot in every row T **is onto.**

Exercise 2.2.1d

Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x+y \\ x+y+z \end{bmatrix}$

(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x}) = A\vec{x}$? If so find A.

(iii) Is T one-to-one and/or onto?

Answer

(i) $T\left(\begin{bmatrix} x+a \\ y+b \\ z+c \end{bmatrix}\right) = \begin{bmatrix} (x+a)+(y+b) \\ (x+a)+(y+b) \\ (x+a)+(y+b)+(z+c) \end{bmatrix} = \begin{bmatrix} x+y \\ x+y \\ x+y+z \end{bmatrix} + \begin{bmatrix} a+b \\ a+b \\ a+b+c \end{bmatrix} = T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) + T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right)$

and $T\left(\begin{bmatrix} rx \\ ry \\ rz \end{bmatrix}\right) = \begin{bmatrix} rx+ry \\ rx+ry \\ rx+ry+rz \end{bmatrix} = r\begin{bmatrix} x+y \\ x+y \\ x+y+z \end{bmatrix} = rT\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)$

Therefore T **is a linear transformation.**

(ii) $T(\vec{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1+0 \\ 1+0 \\ 1+0+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $T(\vec{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0+1 \\ 0+1 \\ 0+1+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $T(\vec{e}_3) = T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0+0 \\ 0+0 \\ 0+0+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

So the matrix representation of T is $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Since there is no pivot in the second column, T **is not one-to-one.** Since there is no pivot in the third row T **is not onto.**

Exercise 2.2.1e

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ 2x+2y \\ x+y \end{bmatrix}$

(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x}) = A\vec{x}$? If so find A.

(iii) Is T one-to-one and/or onto?

Answer

(i) $T\left(\begin{bmatrix} x+a \\ y+b \end{bmatrix}\right) = \begin{bmatrix} (x+a)+(y+b) \\ 2(x+a)+2(y+b) \\ (x+a)+(y+b) \end{bmatrix} = \begin{bmatrix} x+y \\ 2x+2y \\ x+y \end{bmatrix} + \begin{bmatrix} a+b \\ 2a+2b \\ a+b \end{bmatrix} = T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) + T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right)$

and $T\left(\begin{bmatrix} rx \\ ry \end{bmatrix}\right) = \begin{bmatrix} rx+ry \\ 2rx+2ry \\ rx+ry \end{bmatrix} = r\begin{bmatrix} x+y \\ 2x+2y \\ x+y \end{bmatrix} = rT\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)$

Therefore T **is a linear transformation.**

(ii) $T(\vec{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1+0 \\ 2\cdot 1+2\cdot 0 \\ 1+0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ and $T(\vec{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0+1 \\ 2\cdot 0+2\cdot 1 \\ 0+1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

So the matrix representation of T is $A = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 1 & 1 \end{bmatrix}$

(iii) $\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 1 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

Since there is no pivot in the second column, T **is not one-to-one.** Since there is no pivot in the bottom rows T **is not onto.**

Exercise 2.2.2a

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(\vec{v}) = \vec{v} + \vec{e}_1 + \vec{e}_2$

(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x}) = A\vec{x}$? If so find A.

Answer

(i) $T(\vec{u} + \vec{v}) = (\vec{u} + \vec{v}) + \vec{e}_1 + \vec{e}_2$ but $T(\vec{u}) + T(\vec{v}) = (\vec{u} + \vec{e}_1 + \vec{e}_2) + (\vec{v} + \vec{e}_1 + \vec{e}_2)$

These are clearly not equal, so T is not a linear transformation.

It also fails to satisfy $T(r\vec{v}) = rT(\vec{v})$ and $T(\vec{0}) = \vec{0}$, which are also reasons it is not a linear transformation.

(ii) All matrix transformations must be linear transformations, so **there is no matrix** A satisfying $T(\vec{x}) = A\vec{x}$

Exercise 2.2.2b

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by $T(\vec{v}) = \vec{0}$ for all $\vec{v}$

(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x}) = A\vec{x}$? If so find A.

Answer

(i) $T(\vec{u} + \vec{v}) = \vec{0} = \vec{0} + \vec{0} = T(\vec{u}) + T(\vec{v})$

$T(r\vec{v}) = \vec{0} = r\vec{0} = rT(\vec{v})$ so this function is indeed a linear transformation.

(ii) We can see that $T(\vec{e}_1) = \vec{0}$ and $T(\vec{e}_2) = \vec{0}$, so the standard matrix representation of T is $A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

Exercise 2.2.2c

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(\vec{v})$ reflects $\vec{v}$ across the line $y = x$, then reflects that result across the line $y = -x$.

(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x}) = A\vec{x}$? If so find A.

Answer

(i) As complicated as it sounds, the two reflections that T is made of are both easy to determine:

Reflecting the vector $\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ across $y = x$ turns it into $\vec{v}' = \begin{bmatrix} b \\ a \end{bmatrix}$

Reflecting the vector $\vec{v}' = \begin{bmatrix} b \\ a \end{bmatrix}$ across $y = -x$ turns it into $\vec{v}'' = \begin{bmatrix} -a \\ -b \end{bmatrix}$

Therefore $T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} -a \\ -b \end{bmatrix}$, and it is easy to verify this function is indeed a linear transformation.

(ii) We can see that $T(\vec{e}_1) = -\vec{e}_1$ and $T(\vec{e}_2) = -\vec{e}_2$, so the standard matrix representation of T is $A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

Exercise 2.2.3a

Let $T:\mathbb{R}^2\to\mathbb{R}^3$ be a linear transformation satisfying $T\left(\begin{bmatrix}2\\0\end{bmatrix}\right)=\begin{bmatrix}3\\4\\5\end{bmatrix}$ and $T\left(\begin{bmatrix}0\\3\end{bmatrix}\right)=\begin{bmatrix}3\\5\\4\end{bmatrix}$.

Find $T\left(\begin{bmatrix}6\\-6\end{bmatrix}\right)$

Answer

First we find out what linear combination of known input vectors it takes to make our desired input:

$\left[\begin{array}{cc|c}2&0&6\\0&3&-6\end{array}\right]\xrightarrow{RREF}\left[\begin{array}{cc|c}1&0&3\\0&1&-2\end{array}\right]$ therefore $\begin{bmatrix}6\\-6\end{bmatrix}=3\begin{bmatrix}2\\0\end{bmatrix}-2\begin{bmatrix}0\\3\end{bmatrix}$

So $T\left(\begin{bmatrix}6\\-6\end{bmatrix}\right)=3T\left(\begin{bmatrix}2\\0\end{bmatrix}\right)-2T\left(\begin{bmatrix}0\\3\end{bmatrix}\right)=3\begin{bmatrix}3\\4\\5\end{bmatrix}-2\begin{bmatrix}3\\5\\4\end{bmatrix}=\begin{bmatrix}3\\2\\7\end{bmatrix}$

Exercise 2.2.3b

Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be a linear transformation satisfying $T\left(\begin{bmatrix}5\\3\end{bmatrix}\right)=\begin{bmatrix}0\\1\end{bmatrix}$ and $T\left(\begin{bmatrix}3\\5\end{bmatrix}\right)=\begin{bmatrix}1\\0\end{bmatrix}$.

Find $T\left(\begin{bmatrix}7\\1\end{bmatrix}\right)$

Answer

First we find out what linear combination of known input vectors it takes to make our desired input:

$\left[\begin{array}{cc|c}5&3&7\\3&5&1\end{array}\right]\xrightarrow{RREF}\left[\begin{array}{cc|c}1&0&2\\0&1&-1\end{array}\right]$ therefore $\begin{bmatrix}7\\1\end{bmatrix}=2\begin{bmatrix}5\\3\end{bmatrix}-1\begin{bmatrix}3\\5\end{bmatrix}$

So $T\left(\begin{bmatrix}7\\1\end{bmatrix}\right)=2T\left(\begin{bmatrix}5\\3\end{bmatrix}\right)-1T\left(\begin{bmatrix}3\\5\end{bmatrix}\right)=2\begin{bmatrix}0\\1\end{bmatrix}-1\begin{bmatrix}1\\0\end{bmatrix}=\begin{bmatrix}-1\\2\end{bmatrix}$

Exercise 2.2.3c

Let $T:\mathbb{R}^3\to\mathbb{R}^2$ be a linear transformation satisfying $T\left(\begin{bmatrix}1\\0\\1\end{bmatrix}\right)=\begin{bmatrix}2\\-3\end{bmatrix}$ and $T\left(\begin{bmatrix}1\\1\\0\end{bmatrix}\right)=\begin{bmatrix}3\\4\end{bmatrix}$ and $T\left(\begin{bmatrix}0\\1\\3\end{bmatrix}\right)=\begin{bmatrix}0\\0\end{bmatrix}$.

Find $T\left(\begin{bmatrix}16\\-1\\-11\end{bmatrix}\right)$

Answer

First we find out what linear combination of known input vectors it takes to make our desired input:

$\left[\begin{array}{ccc|c}1&1&0&16\\0&1&1&-1\\1&0&3&-11\end{array}\right]\xrightarrow{RREF}\left[\begin{array}{ccc|c}1&0&0&10\\0&1&0&6\\0&0&1&-7\end{array}\right]$ therefore $\begin{bmatrix}16\\-1\\-11\end{bmatrix}=10\begin{bmatrix}1\\0\\1\end{bmatrix}+6\begin{bmatrix}1\\1\\0\end{bmatrix}-7\begin{bmatrix}0\\1\\3\end{bmatrix}$

So $T\left(\begin{bmatrix}16\\-1\\-11\end{bmatrix}\right)=10T\left(\begin{bmatrix}1\\0\\1\end{bmatrix}\right)+6T\left(\begin{bmatrix}1\\1\\0\end{bmatrix}\right)-7T\left(\begin{bmatrix}0\\1\\3\end{bmatrix}\right)=10\begin{bmatrix}2\\-3\end{bmatrix}+6\begin{bmatrix}3\\4\end{bmatrix}-7\begin{bmatrix}0\\0\end{bmatrix}=\begin{bmatrix}38\\-6\end{bmatrix}$

Exercise 2.2.3d

Let $T:\mathbb{R}^3\to\mathbb{R}^3$ be a linear transformation satisfying $T\left(\begin{bmatrix}1\\-2\\2\end{bmatrix}\right)=\begin{bmatrix}-1\\2\\-2\end{bmatrix}$ and $T\left(\begin{bmatrix}1\\-4\\-1\end{bmatrix}\right)=\begin{bmatrix}-1\\4\\1\end{bmatrix}$ and $T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right)=\begin{bmatrix}0\\0\\-1\end{bmatrix}$.

Find $T\left(\begin{bmatrix}3\\-4\\1\end{bmatrix}\right)$

Answer

First we find out what linear combination of known input vectors it takes to make our desired input:

$$\left[\begin{array}{ccc|c}1&1&0&3\\-2&-4&0&-4\\2&-1&1&1\end{array}\right]\xrightarrow{RREF}\left[\begin{array}{ccc|c}1&0&0&4\\0&1&0&-1\\0&0&1&-8\end{array}\right]\text{ therefore }\begin{bmatrix}3\\-4\\1\end{bmatrix}=4\begin{bmatrix}1\\-2\\2\end{bmatrix}-1\begin{bmatrix}1\\-4\\-1\end{bmatrix}-8\begin{bmatrix}0\\0\\1\end{bmatrix}$$

$$\text{So } T\left(\begin{bmatrix}3\\-4\\1\end{bmatrix}\right)=4T\left(\begin{bmatrix}1\\-2\\2\end{bmatrix}\right)-1T\left(\begin{bmatrix}1\\-4\\-1\end{bmatrix}\right)-8T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right)=4\begin{bmatrix}-1\\2\\-2\end{bmatrix}-1\begin{bmatrix}-1\\4\\1\end{bmatrix}-8\begin{bmatrix}0\\0\\-1\end{bmatrix}=\begin{bmatrix}-3\\4\\-1\end{bmatrix}$$

Exercise 2.2.4a

Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be defined by $T(\vec{v})=$ the projection of $\vec{v}$ onto $y=mx$

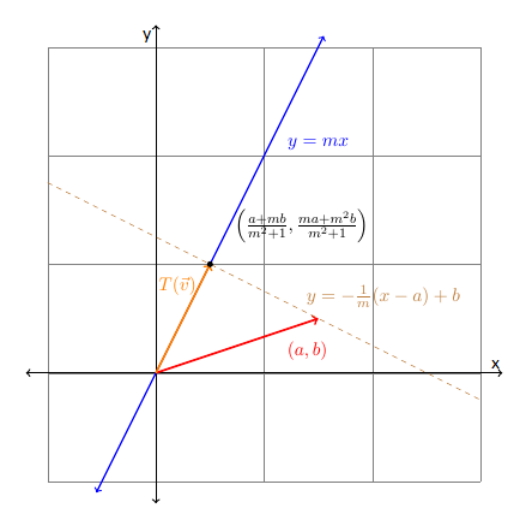
(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x})=A\vec{x}$? If so find A.

Answer

We will answer part (ii), which will automatically give us the answer to part (i).

We start by visualizing $T(\vec{v})$ for $\vec{v}=\begin{bmatrix}a\\b\end{bmatrix}$:



By using the perpendicular line to $y=mx$ that contains the point (a,b), we find its intersection with the line $y=mx$:

$$mx=-\frac{1}{m}(x-a)+b\implies x=\frac{a+mb}{m^2+1}\implies y=\frac{ma+m^2b}{m^2+1}$$

https://math.libretexts.org/@go/page/186412

So our output vector $T(\vec{v}) = \begin{bmatrix} \frac{a+mb}{m^2+1} \\ \frac{ma+m^2b}{m^2+1} \end{bmatrix}$

We find $T(\vec{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \frac{1}{m^2+1} \\ \frac{m}{m^2+1} \end{bmatrix}$ and $T(\vec{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \frac{m}{m^2+1} \\ \frac{m^2}{m^2+1} \end{bmatrix}$

Thus the only possible matrix A that could satisfy $T(\vec{x}) = A\vec{x}$ is $A = \begin{bmatrix} \frac{1}{1+m^2} & \frac{m}{1+m^2} \\ \frac{m}{1+m^2} & \frac{m^2}{1+m^2} \end{bmatrix}$

We check $\begin{bmatrix} \frac{1}{1+m^2} & \frac{m}{1+m^2} \\ \frac{m}{1+m^2} & \frac{m^2}{1+m^2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \frac{a+mb}{m^2+1} \\ \frac{ma+m^2b}{m^2+1} \end{bmatrix} = T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right)$ does hold.

Therefore T is a linear transformation with matrix representation $A = \begin{bmatrix} \frac{1}{1+m^2} & \frac{m}{1+m^2} \\ \frac{m}{1+m^2} & \frac{m^2}{1+m^2} \end{bmatrix}$

Exercise 2.2.4b

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(\vec{v})$ = the reflection of $\vec{v}$ across $y = mx$

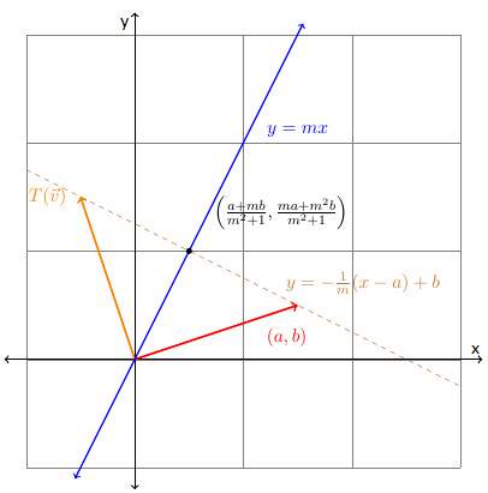
(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x}) = A\vec{x}$? If so find A.

Answer

We will answer part (ii), which will automatically give us the answer to part (i).

We start by visualizing $T(\vec{v})$ for $\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}$:



By using the perpendicular line to $y = mx$ that contains the point (a, b), we find its intersection with the line $y = mx$:

$$mx = -\frac{1}{m}(x-a)+b \implies x = \frac{a+mb}{m^2+1} \implies y = \frac{ma+m^2b}{m^2+1}$$

The difference between this intersection and the initial point (a, b) is the vector $\vec{u} = \begin{bmatrix} \frac{a+mb}{m^2+1} - a \\ \frac{ma+m^2b}{m^2+1} - b \end{bmatrix}$

So our output vector $T(\vec{v}) = \vec{v} + 2\vec{u} = \begin{bmatrix} \frac{(1-m^2)a+2mb}{m^2+1} \\ \frac{2ma+(m^2-1)b}{m^2+1} \end{bmatrix}$

We find $T(\vec{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \frac{1-m^2}{m^2+1} \\ \frac{2m}{m^2+1} \end{bmatrix}$ and $T(\vec{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \frac{2m}{m^2+1} \\ \frac{m^2-1}{m^2+1} \end{bmatrix}$

Thus the only possible matrix A that could satisfy $T(\vec{x}) = A\vec{x}$ is $A = \begin{bmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} \end{bmatrix}$

We check $\begin{bmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \frac{(1-m^2)a+2mb}{m^2+1} \\ \frac{2ma+(m^2-1)b}{m^2+1} \end{bmatrix} = T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right)$ does hold.

Therefore T is a linear transformation with matrix representation $A = \begin{bmatrix} \frac{1-m^2}{1+m^2} & \frac{2m}{1+m^2} \\ \frac{2m}{1+m^2} & \frac{m^2-1}{1+m^2} \end{bmatrix}$

Exercise 2.2.4c

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $T(\vec{v})$ = the rotation of $\vec{v}$ clockwise around the origin by angle θ

(i) Is T a linear transformation?

(ii) Is there a matrix A such that $T(\vec{x}) = A\vec{x}$? If so find A.

Answer

(i) We have seen that rotation by any angle counterclockwise is a linear transformation. Clockwise rotation by θ is just counterclockwise rotation by the angle $2\pi - \theta$, or equivalently $-\theta$. Therefore it is also a linear transformation.

(ii) The vector $\vec{e}_1$ rotated by θ in the clockwise direction moves to $T(\vec{e}_1) = \begin{bmatrix} \cos(-\theta) \\ \sin(-\theta) \end{bmatrix} = \begin{bmatrix} \cos(\theta) \\ -\sin(\theta) \end{bmatrix}$

The vector $\vec{e}_2$ rotated by θ in the clockwise direction moves to $T(\vec{e}_2) = \begin{bmatrix} \cos\left(\frac{\pi}{2} - \theta\right) \\ \sin\left(\frac{\pi}{2} - \theta\right) \end{bmatrix} = \begin{bmatrix} \sin(\theta) \\ \cos(\theta) \end{bmatrix}$

Thus the matrix A that satisfies $T(\vec{x}) = A\vec{x}$ is $A = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$

Exercise 2.2.5

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If $T : \mathbb{R}^4 \to \mathbb{R}^7$ is a linear transformation, then there is some 4×7 matrix A such that $T(\vec{x}) = A\vec{x}$

(b) If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then the set $\{\vec{x} \in \mathbb{R}^n : T(\vec{x}) = \vec{0}\}$ is a subspace of $\mathbb{R}^n$

(c) If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation, then the set $\{\vec{x} \in \mathbb{R}^n : T(\vec{x}) = \vec{x}\}$ is a subspace of $\mathbb{R}^n$

(d) If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then the set $\{\vec{y} \in \mathbb{R}^m : T(\vec{x}) = \vec{y}\}$ is a subspace of $\mathbb{R}^m$

(e) If $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation that satisfies $T\left(\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ then $T\left(\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$

(f) If $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation that satisfies $T\left(\begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}\right) = T\left(\begin{bmatrix} 9 \\ 7 \\ 5 \end{bmatrix}\right)$ then $T\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

(g) If $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation that satisfies $T\left(\begin{bmatrix} -4 \\ 8 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}$ then $T\left(\begin{bmatrix} 2 \\ -4 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Answer

(a) **Always False**

The matrix A would have to be 7×4 to represent T.

(b) **Always True**

We know $T(\vec{0}) = \vec{0}$ always holds, and if $T(\vec{u}) = \vec{0}$ and $T(\vec{v}) = \vec{0}$ then $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) = \vec{0}$ and $T(r\vec{u}) = rT(\vec{u}) = r\vec{0} = \vec{0}$ must hold.

(c) **Always True**

We know $T(\vec{0}) = \vec{0}$ always holds, and if $T(\vec{u}) = \vec{u}$ and $T(\vec{v}) = \vec{v}$ then $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v}) = \vec{u} + \vec{v}$ and $T(r\vec{u}) = rT(\vec{u}) = r\vec{u}$ must hold.

(d) **Always True**

We know $T(\vec{0}) = \vec{0}$ always holds, and if $T(\vec{a}) = \vec{u}$ and $T(\vec{b}) = \vec{v}$ then $T(\vec{a} + \vec{b}) = T(\vec{a}) + T(\vec{b}) = \vec{u} + \vec{v}$ and $T(r\vec{a}) = rT(\vec{a}) = r\vec{u}$ must hold.

(e) **Sometimes True Sometimes False**

We can define many such T which make this true (for example the identity transformation with $T(\vec{x}) = \vec{x}$). To make this false, consider $T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = \begin{bmatrix} a \\ 0 \\ 3a \end{bmatrix}$.

(f) **Always True**

Notice $T\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right) = T\left(\frac{1}{4}\begin{bmatrix} 9 \\ 7 \\ 5 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}\right) = \frac{1}{4}T\left(\begin{bmatrix} 9 \\ 7 \\ 5 \end{bmatrix}\right) - \frac{1}{4}T\left(\begin{bmatrix} 5 \\ 7 \\ 9 \end{bmatrix}\right) = \vec{0}.$

(g) **Always False**

$T\left(\begin{bmatrix} 2 \\ -4 \end{bmatrix}\right) = T\left(-\frac{1}{2}\begin{bmatrix} -4 \\ 8 \end{bmatrix}\right) = -\frac{1}{2}T\left(\begin{bmatrix} -4 \\ 8 \end{bmatrix}\right) = -\frac{1}{2}\begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ -3 \end{bmatrix}.$

2.2: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.3: Matrix Algebra

Objectives

1. Understand compositions of transformations.
2. Understand the relationship between matrix products and compositions of matrix transformations.
3. Become comfortable doing basic algebra involving matrices.
4. *Recipes:* matrix multiplication (two ways), compute the inverse matrix.
5. *Picture:* composition of transformations.
6. *Vocabulary words:* ***composition, inverse matrix***, ***inverse transformation***.

In this section, we discuss addition and scalar multiplication of transformations and of matrices. We then study compositions of transformations. As we will see, composition is a way of chaining transformations together. The composition of matrix transformations corresponds to a notion of *multiplying* two matrices together. We finish with what passes for *division* within matrices as well.

This page titled 2.3: Matrix Algebra is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **3.4: Matrix Multiplication** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.3.1: Sums of Transformations and Matrices

In this subsection we describe the straightforward operations that one can perform on transformations: addition and scalar multiplication. We then translate these operations into the language of matrices.

Definition: Transformation Sum

Let $T, U: \mathbb{R}^n \to \mathbb{R}^m$ be two transformations. Their ***sum*** is the transformation $T + U: \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$(T + U)(\vec{x}) = T(\vec{x}) + U(\vec{x})$$

Note that addition of transformations is only defined when both transformations have the same domain and codomain.

To emphasize, the sum of two transformations $T, U: \mathbb{R}^n \to \mathbb{R}^m$ is another transformation called $T + U$; its value on an input vector $\vec{x}$ is the sum of the outputs of T and U.

Illustration 2.3.1.1: Functions of one variable

Define $f: \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$ and $g: \mathbb{R} \to \mathbb{R}$ by $g(x) = x^3$. The sum $f + g: \mathbb{R} \to \mathbb{R}$ is the transformation defined by the rule

$$(f + g)(x) = f(x) + g(x) = x^2 + x^3$$

For instance, $(f + g)(-2) = (-2)^2 + (-2)^3 = -4$

We want to define matrix addition and scalar multiplication in a way that allows this relationship between linear transformations to be maintained.

Theorem 2.3.1.1

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ and $U: \mathbb{R}^n \to \mathbb{R}^m$ be linear transformations, and let A and B be their standard matrices, respectively, so A and B are each an $m \times n$ matrix. Then $T + U: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, and its standard matrix is formed by adding the matching entries of A and B.

Proof

First we verify that $T + U$ is linear. Let $\vec{x}, \vec{y}$ be vectors in $\mathbb{R}^n$. Then

$$\begin{aligned}(T + U)(\vec{x} + \vec{y}) &= T(\vec{x} + \vec{y}) + U(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y}) + U(\vec{x}) + U(\vec{y}) \\ &= T(\vec{x}) + U(\vec{x}) + T(\vec{y}) + U(\vec{y}) = (T + U)(\vec{x}) + (T + U)(\vec{y}).\end{aligned}$$

If c is a scalar, then

$$(T + U)(c\vec{x}) = T(c\vec{x}) + U(c\vec{x}) = cT(\vec{x}) + cU(\vec{x}) = c(T + U)(\vec{x})$$

Since $T + U$ satisfies the two defining properties, it is a linear transformation.

Now that we know that $T + U$ is linear, it makes sense to compute its standard matrix. Let C be the standard matrix of $T + U$, so $T(\vec{x}) = A\vec{x}$, $U(\vec{x}) = B\vec{x}$, and $(T + U)(\vec{x}) = C\vec{x}$.

The first column of A is $A\vec{e}_1$, the first column of B is $B\vec{e}_1$, and we must have the first column of C be:

$$(T + U)(\vec{e}_1) = T(\vec{e}_1) + U(\vec{e}_1) = A\vec{e}_1 + B\vec{e}_1$$

By definition then, the first column of the sum $C = A + B$ is the sum of the first column of A with the first column of B.

The same argument as applied to each of the ith standard coordinate vector $\vec{e}_i$ shows that each column of $C = A + B$ is given by the sum of that column in the matrix A with that column in the matrix B.

The theorem justifies our choice of definition of the matrix sum. This is the one and only reason that matrix products are defined in this way (not because it is easy to do, that is just a nice bonus). To rephrase:

Definition: Matrix Sum

The ***sum*** of two $m \times n$ matrices is the matrix obtained by summing the entries of A and B individually:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{bmatrix}$$

In other words, the i,j entry of $A+B$ is the sum of the i,j entries of A and B. Note that addition of matrices is only defined when both matrices have the same size.

Similarly, the product of T with a scalar c is another transformation called cT; its value on an input vector $\vec{x}$ is the vector $c \cdot T(\vec{x})$.

Definition: Transformation Scalar Product

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a transformation, and let c be a scalar. The ***scalar product*** of c with T is the transformation $cT: \mathbb{R}^n \to \mathbb{R}^m$ defined by

$$(cT)(\vec{x}) = c \cdot T(\vec{x}).$$

Illustration 2.3.1.2: Functions of one variable

Define $\exp: \mathbb{R} \to \mathbb{R}$ by $\exp(x) = e^x$. The product $2\exp: \mathbb{R} \to \mathbb{R}$ is the transformation defined by the rule

$$(2\exp)(x) = 2 \cdot \exp(x) = 2e^x$$

For instance, $(2\exp)(1) = 2 \cdot \exp(1) = 2e$

Theorem 2.3.1.2

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, c be a scalar, and A be the $m \times n$ standard matrix for T. Then $cT: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, and its standard matrix is formed by multiplying each entry of A by c.

Proof

First we verify that cT is linear. Let $\vec{x}, \vec{y}$ be vectors in $\mathbb{R}^n$. Then

$$\begin{aligned} (cT)(\vec{x}+\vec{y}) &= cT(\vec{x}+\vec{y}) = c\left(T(\vec{x}) + T(\vec{y})\right) \\ &= cT(\vec{x}) + cT(\vec{y}) = (cT)(\vec{x}) + (cT)(\vec{y}). \end{aligned} \tag{2.3.1.1}$$

If k is a scalar, then

$$(cT)(k\vec{x}) = c\left(T(k\vec{x})\right) = c\left(kT(\vec{x})\right) = k(cT)(\vec{x})$$

Since cT satisfies the two defining properties, it is a linear transformation.

Now that we know that cT is linear, it makes sense to compute its standard matrix. Let D be the standard matrix of cT, so $T(\vec{x}) = A\vec{x}$, and $(cT)(\vec{x}) = D\vec{x}$.

The first column of A is $A\vec{e}_1$, and we must have the first column of D be:

$$(cT)(\vec{e}_1) = cT(\vec{e}_1) = c\left(A\vec{e}_1\right)$$

By definition then, the first column of the scalar product $D = cA$ is the scalar product of the first column of A by c.

The same argument as applied to the ith standard coordinate vector $\vec{e}_i$ shows that each column of $D = cA$ is given by the scalar multiple of that column in the matrix A by c.

This leads to the analogous operations for matrices.

Definition: Matrix Scalar Product

The ***scalar product*** of a scalar c with a matrix A is obtained by scaling all entries of A by c:

$$c\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & ca_{13} \\ ca_{21} & ca_{22} & ca_{23} \end{bmatrix}$$

In other words, the i, j entry of cA is c times the i, j entry of A.

By defining the sum and scalar product of matrices in this way, we ensure that the link between linear transformations and matrices remains intact. In other words:

Fact 2.3.1.1

Let $T, U: \mathbb{R}^n \to \mathbb{R}^m$ be linear transformations with standard matrices A, B, respectively, and let c be a scalar.

- The standard matrix for $T + U$ is $A + B$.
- The standard matrix for cT is cA.

Note 2.3.1.1: Properties of addition and scalar multiplication for transformations

Let $S, T, U: \mathbb{R}^n \to \mathbb{R}^m$ be transformations and let c, d be scalars. The following properties are easily verified:

$$\begin{aligned} T + U &= U + T & S + (T + U) &= (S + T) + U \\ c(T + U) &= cT + cU & (c + d)T &= cT + dT \\ c(dT) &= (cd)T & T + 0 &= T \end{aligned}$$

In one of the above properties, we used 0 to denote the transformation $\mathbb{R}^n \to \mathbb{R}^m$ that is zero on every input vector: $0(\vec{x}) = \vec{0}$ for all $\vec{x}$. This is called the ***zero transformation***.

Note 2.3.1.2: Properties of addition and scalar multiplication for matrices

Let A, B, C be $m \times n$ matrices and let c, d be scalars. Then:

$$\begin{aligned} A + B &= B + A & C + (A + B) &= (C + A) + B \\ c(A + B) &= cA + cB & (c + d)A &= cA + dA \\ c(dA) &= (cd)A & A + 0 &= A \end{aligned}$$

In one of the above properties, we used $\mathbf{0}$ to denote the $m \times n$ matrix whose entries are all zero. This is the standard matrix of the zero transformation, and is called the ***zero matrix***.

2.3.1: Sums of Transformations and Matrices is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.4: Matrix Multiplication** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.3.2: Transformation Composition and Matrix Multiplication

Composition means the same thing in linear algebra as it does in Calculus. Here is the definition.

Definition: Transformation Composition

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ and $U: \mathbb{R}^p \to \mathbb{R}^n$ be transformations. Their ***composition*** is the transformation $T \circ U: \mathbb{R}^p \to \mathbb{R}^m$ defined by

$$(T \circ U)(\vec{x}) = T(U(\vec{x}))$$

Composing two transformations means chaining them together: $T \circ U$ is the transformation that first applies U, then applies T (note the order of operations). More precisely, to evaluate $T \circ U$ on an input vector $\vec{x}$, first you evaluate $U(\vec{x})$, then you take this output vector of U and use it as an input vector of T: that is, $(T \circ U)(\vec{x}) = T(U(\vec{x}))$. Of course, this only makes sense when the outputs of U are valid inputs of T, that is: when the range of U is contained in the domain of T.

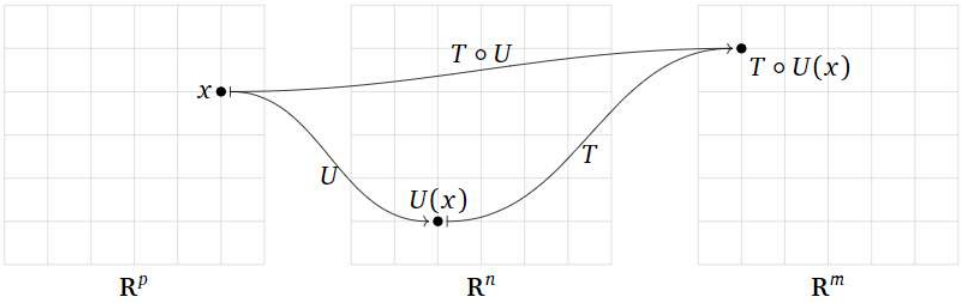


Figure 2.3.2.1

Here is a picture of the composition $T \circ U$ as a "machine" that first runs U, then takes its output and feeds it into T;

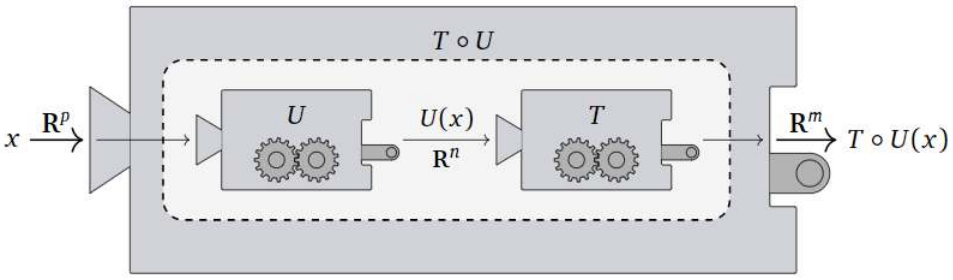


Figure 2.3.2.2

Note 2.3.2.3: Domain and Codomain of a Composition

- In order for $T \circ U$ to be defined, the codomain of U must equal the domain of T.
- The domain of $T \circ U$ is the domain of U.
- The codomain of $T \circ U$ is the codomain of T.

Illustration 2.3.2.3: Functions of one variable

Define $f: \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$ and $g: \mathbb{R} \to \mathbb{R}$ by $g(x) = x^3$. The composition $f \circ g: \mathbb{R} \to \mathbb{R}$ is the transformation defined by the rule

$$(f \circ g)(x) = f(g(x)) = f(x^3) = (x^3)^2 = x^6$$

For instance, $f \circ g(-2) = f(-8) = 64$

Illustration 2.3.2.4: A composition of matrix transformations

Define $T:\mathbb{R}^3 \to \mathbb{R}^2$ and $U:\mathbb{R}^2 \to \mathbb{R}^3$ by

$$T(\vec{x}) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \vec{x} \quad \text{and} \quad U(\vec{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \vec{x}$$

Their composition is the transformation $T \circ U:\mathbb{R}^2 \to \mathbb{R}^2$; and we can see that:

$$(T \circ U)\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = T\left(U\left(\begin{bmatrix} a \\ b \end{bmatrix}\right)\right) = T\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}\right)$$

$$= T\left(\begin{bmatrix} a \\ b \\ a \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ a \end{bmatrix} = \begin{bmatrix} a+b \\ a+b \end{bmatrix}$$

This shows that $T \circ U$ is a matrix transformation with representation matrix $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

Note 2.3.2.4: Properties of composition

Let S, T, U be transformations and let c be a scalar. Suppose that $T:\mathbb{R}^n \to \mathbb{R}^m$, and that in each of the following identities, the domains and the codomains are compatible when necessary for the composition to be defined. The following properties are easily verified:

$$S \circ (T+U) = S \circ T + S \circ U \qquad (S+T) \circ U = S \circ U + T \circ U$$
$$c(T \circ U) = (cT) \circ U \qquad c(T \circ U) = T \circ (cU) \quad \text{if } T \text{ is linear}$$
$$T \circ \mathcal{I} = T \qquad \mathcal{I} \circ T = T$$
$$S \circ (T \circ U) = (S \circ T) \circ U$$

The final property is called ***associativity***. Unwrapping both sides, it says:

$$S \circ (T \circ U)(\vec{x}) = S(T \circ U(\vec{x})) = S(T(U(\vec{x}))) = S \circ T(U(\vec{x})) = (S \circ T) \circ U(\vec{x})$$

In other words, both $S \circ (T \circ U)$ and $(S \circ T) \circ U$ are the transformation defined by first applying U, then T, then S.

Illustration 2.3.2.5: Functions of one variable

Define $f:\mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$ and $g:\mathbb{R} \to \mathbb{R}$ by $g(x) = e^x$. The composition $f \circ g:\mathbb{R} \to \mathbb{R}$ is the transformation defined by the rule

$$f \circ g(x) = f(g(x)) = f(e^x) = (e^x)^2 = e^{2x}$$

The composition $g \circ f:\mathbb{R} \to \mathbb{R}$ is the transformation defined by the rule

$$g \circ f(x) = g(f(x)) = g(x^2) = e^{x^2}$$

Note that $e^{x^2} \neq e^{2x}$ in general; for instance, if $x = 1$ then $e^{x^2} = e$ and $e^{2x} = e^2$. Thus $f \circ g$ is not equal to $g \circ f$, and we can already see with functions of one variable that composition of functions is not commutative.

Note 2.3.2.5

Composition of transformations is *not* commutative in general. That is, in general, $T \circ U \neq U \circ T$, even when both compositions are defined.

Theorem 2.3.2.3

Let $T:\mathbb{R}^n \to \mathbb{R}^m$ and $U:\mathbb{R}^p \to \mathbb{R}^n$ be linear transformations, and let A and B be their standard matrices, respectively, so A is an $m \times n$ matrix and B is an $n \times p$ matrix. Then $T \circ U:\mathbb{R}^p \to \mathbb{R}^m$ is a linear transformation, and its standard matrix is formed by multiplying each column of B by the matrix A.

Proof

First we verify that $T \circ U$ is linear. Let $\vec{u}, \vec{v}$ be vectors in $\mathbb{R}^p$. Then

$$\begin{aligned} T \circ U(\vec{u}+\vec{v}) &= T(U(\vec{u}+\vec{v})) = T(U(\vec{u})+U(\vec{v})) \\ &= T(U(\vec{u})) + T(U(\vec{v})) = T \circ U(\vec{u}) + T \circ U(\vec{v}) \end{aligned}$$

If c is a scalar, then

$$T \circ U(c\vec{v}) = T(U(c\vec{v})) = T(cU(\vec{v})) = cT(U(\vec{v})) = cT \circ U(\vec{v})$$

Since $T \circ U$ satisfies the two defining properties, it is a linear transformation.

Now that we know that $T \circ U$ is linear, it makes sense to compute its standard matrix. Let E be the standard matrix of $T \circ U$, so $T(\vec{x}) = A\vec{x}$, $U(\vec{x}) = B\vec{x}$, and $T \circ U(\vec{x}) = E\vec{x}$. The first column of E is $E\vec{e}_1$, and the first column of B is $\vec{b}_1 = B\vec{e}_1$. We have

$$T \circ U(\vec{e}_1) = T(U(\vec{e}_1)) = T(B\vec{e}_1) = T(\vec{b}_1) = A\vec{b}_1$$

By definition, the first column of the matrix representation E is the product of A with the first column of B,

The same argument as applied to each of the ith standard coordinate vectors $\vec{e}_i$ shows that E has ith column defined as the product of A with the ith column vector in B.

Note that we could alternatively say that

$$E\vec{x} = T \circ U(\vec{x}) = A(B\vec{x}) = (AB)\vec{x}$$

It follows that E should be how we define the matrix product AB.

The theorem justifies our choice of definition of the matrix product. This is the one and only reason that matrix products are defined in this (terrible) way. To rephrase:

Definition: Entry of the Matrix

Let A be an $m \times n$ matrix. We will generally write a_{ij} for the entry in the ith row and the jth column. It is called the i,j ***entry*** of the matrix.

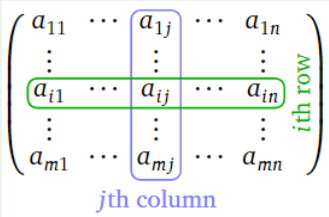


Figure 2.3.2.3

Definition: Matrix Multiplication

Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix. Denote the columns of B by $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p$:

$$B = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_p \\ | & | & & | \end{bmatrix}$$

The *product* AB is the $m \times p$ matrix with columns $A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_p$:

$$AB = \begin{bmatrix} | & | & & | \\ A\vec{v}_1 & A\vec{v}_2 & \cdots & A\vec{v}_p \\ | & | & & | \end{bmatrix}$$

In other words, matrix multiplication is defined column-by-column, or "distributes over the columns of B."

Illustration 2.3.2.6

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \left[\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right]$$
$$= \left[\begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right] = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

In order for the vectors $A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_p$ to be defined, the numbers of rows of B has to equal the number of columns of A.

Note 2.3.2.6: The Sizes of the Matrices in the Matrix Product

- In order for AB to be defined, the number of rows of B has to equal the number of columns of A.
- The product of an $m \times n$ matrix and an $n \times p$ matrix is an $m \times p$ matrix.

If B has only one column, then AB also has one column. A matrix with one column is the same as a vector, so the definition of the matrix product generalizes the definition of the matrix-vector product we defined previously.

If A is a square matrix, then we can multiply it by itself; we define its ***powers*** to be

$$A^2 = AA \qquad A^3 = AAA \qquad \text{etc.}$$

Recall from Chapter 1 that the product of a row vector $\vec{a}$ and a column vector $\vec{x}$ is the scalar

$$\vec{a} \cdot \vec{x} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1x_1 + a_2x_2 + \cdots + a_nx_n$$

The following procedure for finding the matrix product is much better adapted to computations by hand.

Recipe: The Row-Column Rule for Matrix Multiplication

Let A be an $m \times n$ matrix, let B be an $n \times p$ matrix, and let $C = AB$. Then the ij entry of C is the ith row of A times the jth column of B:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Here is a diagram:

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{k1} & \cdots & b_{kj} & \cdots & b_{kp} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{ip} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mp} \end{pmatrix}$$

ith row; jth column; ij entry

Figure 2.3.2.4

Illustration 2.3.2.7

The row-column rule allows us to compute the product matrix one entry at a time:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} 1\cdot 1 + 2\cdot 2 + 3\cdot 3 & \square \\ \square & \square \end{bmatrix} = \begin{bmatrix} 14 & \square \\ \square & \square \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} \square & \square \\ 4\cdot 1 + 5\cdot 2 + 6\cdot 3 & \square \end{bmatrix} = \begin{bmatrix} 14 & \square \\ 32 & \square \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} \square & 1\cdot(-3) + 2\cdot(-2) + 3\cdot(-1) \\ \square & \square \end{bmatrix} = \begin{bmatrix} 14 & -10 \\ 32 & \square \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}\begin{bmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} \square & \square \\ 4\cdot(-3) + 5\cdot(-2) + 6\cdot(-1) & \square \end{bmatrix} = \begin{bmatrix} 14 & -10 \\ 32 & -28 \end{bmatrix}$$

Although matrix multiplication satisfies many of the properties one would expect (see the end of the section), one must be careful when doing matrix arithmetic, as there are several properties that are not satisfied in general.

Note 2.3.2.7: Matrix Multiplication Caveats

- Matrix multiplication is not commutative: AB is not usually equal to BA, even when both products are defined and have the same size. (See this demonstrated in the example below.)
- Matrix multiplication does not satisfy the cancellation law: $AB = AC$ does not imply $B = C$, even when $A \neq 0$. For example,

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix}$$

- It is possible for $AB = 0$, even when $A \neq 0$ and $B \neq 0$. For example,

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

While matrix multiplication is not commutative in general there are examples of matrices A and B with $AB = BA$. For example, this always works when A is the zero matrix, or when $A = B$. The reader is encouraged to find other examples.

Illustration 2.3.2.8: Non-commutative multiplication of matrices

Consider the products of the matrices $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

The matrix AB is $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

Whereas the matrix BA is $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$

Clearly we have $AB \neq BA$. We see matrix multiplication is not always commutative.

Note 2.3.2.8: Order of Operations

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ and $U: \mathbb{R}^p \to \mathbb{R}^n$ be linear transformations, and let A and B be their standard matrices, respectively. Recall that $T \circ U(\vec{x})$ is the vector obtained by first applying U to $\vec{x}$, and then T.

On the matrix side, the standard matrix of $T \circ U$ is the product AB, so $T \circ U(\vec{x}) = (AB)\vec{x}$. By associativity of matrix multiplication, we have $(AB)\vec{x} = A(B\vec{x})$, so the product $(AB)\vec{x}$ can be computed by first multiplying $\vec{x}$ by B, *then* multiplying

the product by A.

Therefore, matrix multiplication happens in the same order as composition of transformations. In other words, *both matrices and transformations are written in the order opposite from the order in which they act.* But matrix multiplication and composition of transformations are written in the same order as each other: the matrix for $T \circ U$ is AB.

Illustration 2.3.2.9: Composition of rotations

We have seen that the standard matrix for the counterclockwise rotation of the plane by an angle of θ is

$$A = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be counterclockwise rotation by $45°$, and let $U: \mathbb{R}^2 \to \mathbb{R}^2$ be counterclockwise rotation by $90°$. The matrices A and B for T and U are, respectively,

$$A = \begin{bmatrix} \cos(45°) & -\sin(45°) \\ \sin(45°) & \cos(45°) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$B = \begin{bmatrix} \cos(90°) & -\sin(90°) \\ \sin(90°) & \cos(90°) \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

The standard matrix of the composition $T \circ U$ is

$$AB = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

This is consistent with the fact that $T \circ U$ is counterclockwise rotation by $90° + 45° = 135°$: we have

$$\begin{bmatrix} \cos(135°) & -\sin(135°) \\ \sin(135°) & \cos(135°) \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Challenge

Derive the trigonometric identities

$$\sin(\alpha \pm \beta) = \sin(\alpha)\cos(\beta) \pm \cos(\alpha)\sin(\beta)$$

and

$$\cos(\alpha \pm \beta) = \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta)$$

as in the previous example.

Recall that the *identity matrix* is the $n \times n$ matrix I whose columns are the standard coordinate vectors in $\mathbb{R}^n$. The identity matrix is the standard matrix of the identity transformation: that is, $\vec{x} = \mathcal{I}(\vec{x}) = I\vec{x}$ for all vectors $\vec{x}$ in $\mathbb{R}^n$. For any linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ we have

$$T \circ \mathcal{I} = T$$

and by the same token we have for any $m \times n$ matrix A we have

$$AI = A$$

Similarly, we have $\mathcal{I} \circ T = T$ and $IA = A$ for $T: \mathbb{R}^m \to \mathbb{R}^n$ and A any $n \times m$ matrix.

We can also combine addition and scalar multiplication of matrices with multiplication of matrices. Since matrix multiplication corresponds to composition of transformations, the following properties are consequences of the corresponding properties of transformations.

Note 2.3.2.9: Properties of matrix multiplication

Let A, B, C be matrices and let c be a scalar. Suppose that A is an $m \times n$ matrix, and that in each of the following identities, the sizes of B and C are compatible when necessary for the product to be defined. Then:

$$\begin{aligned} C(A+B) &= CA + CB & (A+B)C &= AC + BC \\ c(AB) &= (cA)B & c(AB) &= A(cB) \\ AI &= A & IA &= A \\ (AB)C &= A(BC) \end{aligned}$$

Most of the above properties are easily verified directly from the definitions. The *associativity* property $(AB)C = A(BC)$, however, is not (try it!). It is much easier to prove by relating matrix multiplication to composition of transformations, and using the obvious fact that composition of transformations is associative.

2.3.2: Transformation Composition and Matrix Multiplication is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.4: Matrix Multiplication** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.3.3: Matrix Inverses

In this subsection, we learn to "divide" by a matrix. One has to take care when "dividing by matrices", however, because not every matrix has an inverse, and the order of matrix multiplication is important.

The *reciprocal* or *inverse* of a nonzero number a is the number b which is characterized by the property that $ab = 1$. For instance, the inverse of 7 is $\frac{1}{7}$. We use this formulation to define the inverse of a matrix.

Definition: Invertible

Let A be an $n \times n$ (square) matrix. We say that A is ***invertible*** if there is an $n \times n$ matrix B such that

$$AB = I \quad \text{and} \quad BA = I$$

In this case, the matrix B is called the ***inverse*** of A, and we write $B = A^{-1}$.

We have to require $AB = I$ *and* $BA = I$ because in general matrix multiplication is not commutative. However, we will show in the next section that if A and B are $n \times n$ matrices such that $AB = I$, then automatically $BA = I$.

Example 2.3.3.1

Determine if the matrices

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

are inverses.

Solution

We will check that $AB = I$ and that $BA = I$.

$$AB = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$BA = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Therefore, $\boxed{A \text{ is invertible, with inverse } B}$

Remark

There exist non-square matrices whose product is the identity. Indeed, if

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

then $AB = I$. However, $BA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq I$, so B does not deserve to be called the inverse of A.

One can show in the future that if A is an $n \times m$ matrix for $n \neq m$, then there is no $m \times n$ matrix B such that $AB = I$ and $BA = I$. For this reason, we restrict ourselves to *square* matrices when we discuss matrix invertibility.

Fact 2.3.3.2: Facts about Invertible Matrices

Let A and B be invertible $n \times n$ matrices.

1. A^{-1} is invertible, and its inverse is $(A^{-1})^{-1} = A$.
2. AB is invertible, and its inverse is $(AB)^{-1} = B^{-1}A^{-1}$ (note the order).

Proof

1. The equations $AA^{-1} = I$ and $A^{-1}A = I$ at the same time exhibit A^{-1} as the inverse of A and A as the inverse of A^{-1}.
2. We compute

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$$

Here we used the associativity of matrix multiplication and the fact that $IB = B$. This shows that $B^{-1}A^{-1}$ is the inverse of AB.

Why is the inverse of AB not equal to $A^{-1}B^{-1}$? If it were, then we would have

$$I = (AB)(A^{-1}B^{-1}) = ABA^{-1}B^{-1}$$

But there is no reason for $ABA^{-1}B^{-1}$ to equal the identity matrix: one cannot switch the order of A^{-1} and B, so there is nothing to cancel in this expression. In fact, if $I = (AB)(A^{-1}B^{-1})$, then we can multiply both sides on the right by BA to conclude that $AB = BA$. In other words, $(AB)^{-1} = A^{-1}B^{-1}$ if and only if $AB = BA$.

More generally, the inverse of a product of several invertible matrices is the product of the inverses, in the opposite order. This intuitive fact is commonly referred to as the "socks and shoes" theorem. If you start the day by putting on your socks B, then putting on your shoes A, you will end your day by first taking off your shoes A^{-1}, then taking off your socks B^{-1}, not the other way around! For instance,

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$$

So far we have defined the inverse matrix without giving any strategy for computing it. The following theorem gives a procedure for computing A^{-1} in general.

Theorem 2.3.3.4

Let A be an $n \times n$ matrix, and let $[\,A \mid I\,]$ be the matrix obtained by augmenting A by the identity matrix. If the reduced row echelon form of $[\,A \mid I\,]$ has the form $[\,I \mid B\,]$, then A is invertible and $B = A^{-1}$. Otherwise, A is not invertible.

Proof

First suppose that the reduced row echelon form of $[\,A \mid I\,]$ does not have the form $[\,I \mid B\,]$. This means that fewer than n pivots are contained in the first n columns (the non-augmented part), so A has fewer than n pivots. It follows that $\mathbf{Nul}(A) \neq \{\vec{0}\}$ (the equation $A\vec{x} = \vec{0}$ has a free variable), so there exists a nonzero vector $\vec{v}$ in $\mathbf{Nul}(A)$. Suppose that there were a matrix B such that $BA = I$. Then

$$\vec{v} = I\vec{v} = BA\vec{v} = B\vec{0} = \vec{0}$$

which is impossible as $\vec{v} \neq \vec{0}$. Therefore, A is not invertible.

Now suppose that the reduced row echelon form of $[\,A \mid I\,]$ has the form $[\,I \mid B\,]$. In this case, all pivots are contained in the non-augmented part of the matrix, so the augmented part plays no role in the row reduction: the entries of the augmented part do not influence the choice of row operations used. Hence, row reducing $[\,A \mid I\,]$ is equivalent to solving the n systems of linear equations $A\vec{x}_1 = \vec{e}_1$, $A\vec{x}_2 = \vec{e}_2$, ..., $A\vec{x}_n = \vec{e}_n$, where $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ are the standard coordinate vectors, which are the columns of I.

$$A\vec{x}_1 = \vec{e}_1: \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -4 & 0 & 0 & 1 \end{array}\right]$$

$$A\vec{x}_2 = \vec{e}_2: \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -4 & 0 & 0 & 1 \end{array}\right]$$

$$A\vec{x}_3 = \vec{e}_3: \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -4 & 0 & 0 & 1 \end{array}\right]$$

The columns $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ of the matrix B in the row reduced form are the solutions to these equations:

$$A\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \vec{e}_1: \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & \frac{3}{2} & \frac{1}{2} \end{array}\right]$$

$$A\begin{bmatrix} -6 \\ -2 \\ \frac{3}{2} \end{bmatrix} = \vec{e}_2: \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & \frac{3}{2} & \frac{1}{2} \end{array}\right]$$

$$A\begin{bmatrix} -2 \\ -1 \\ \frac{1}{2} \end{bmatrix} = \vec{e}_3: \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & \frac{3}{2} & \frac{1}{2} \end{array}\right]$$

The product $B\vec{e}_i$ is just the ith column of B, which is $\vec{x}_i$ so

$$\vec{e}_i = A\vec{x}_i = AB\vec{e}_i$$

for all i. By the same fact, the ith column of AB is $\vec{e}_i$, which means that AB is the identity matrix. Thus B is the inverse of A.

Example 2.3.3.2: An invertible matrix

Find the inverse of the matrix

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

Solution

We augment by the identity and row reduce:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -4 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_3 = R_3 + 3R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 3 & 1 \end{array}\right]$$

$$\xrightarrow{\begin{array}{l} R_1 = R_1 - 2R_3 \\ R_2 = R_2 - R_3 \end{array}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 2 & 0 & 3 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 = R_3 \div 2} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & \frac{3}{2} & \frac{1}{2} \end{array}\right]$$

Therefore, the inverse matrix is

$$\boxed{A^{-1} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & -3 & -4 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -6 & -2 \\ 0 & -2 & -1 \\ 0 & \frac{3}{2} & \frac{1}{2} \end{bmatrix}}$$

We check the product to confirm our answer:

$$\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & -3 & -4 \end{bmatrix} \begin{bmatrix} 1 & -6 & -2 \\ 0 & -2 & -1 \\ 0 & \frac{3}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -6 & -2 \\ 0 & -2 & -1 \\ 0 & \frac{3}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & -3 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Example 2.3.3.3: A non-invertible matrix

Is the following matrix invertible?

$$A = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & -3 & -6 \end{bmatrix}$$

Solution

We augment by the identity and row reduce:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -6 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_3 = R_3 + 3R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \end{array}\right]$$

At this point we can stop, because it is clear that the reduced row echelon form will not have I in the non-augmented part: it will have a row of zeros, and therefore the matrix is **not invertible**

2.3.3: Matrix Inverses is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.5: Matrix Inverses** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.3.4: Practice Set

Exercise 2.3.4.1

Consider the matrices $A = \begin{bmatrix} 1 & 2 \\ 5 & 9 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix}$ and evaluate each expression below (if it exists).

(a) $A + B$

(b) $4A$

(c) AB

(d) A^{-1}

(e) B^{-1}

(f) $(A + B)^{-1}$

Answer

(a) $A + B = \begin{bmatrix} 1 & 1 \\ 7 & 10 \end{bmatrix}$

(b) $4A = \begin{bmatrix} 4 & 8 \\ 20 & 36 \end{bmatrix}$

(c) $AB = \begin{bmatrix} 4 & 1 \\ 18 & 4 \end{bmatrix}$

(d) $A^{-1} = \begin{bmatrix} -9 & 2 \\ 5 & -1 \end{bmatrix}$

(e) $B^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -1 & 0 \end{bmatrix}$

(f) $(A + B)^{-1} = \begin{bmatrix} \frac{10}{3} & -\frac{1}{3} \\ -\frac{7}{3} & \frac{1}{3} \end{bmatrix}$

Exercise 2.3.4.2

Consider the matrices $A = \begin{bmatrix} 0 & -1 & 4 \\ 1 & 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -3 \\ 2 & -6 \end{bmatrix}$ and evaluate each expression below (if it exists).

(a) $3B$

(b) AB

(c) BA

(d) A^2

(e) B^3

(f) B^{-1}

Answer

(a) $3B = \begin{bmatrix} 3 & -9 \\ 6 & -18 \end{bmatrix}$

(b) AB does not exist, since A is 2×3 and B is 2×2

(c) $BA = \begin{bmatrix} -3 & -1 & -2 \\ -6 & -2 & -4 \end{bmatrix}$

(d) A^2 does not exist since A is not square and cannot multiply by itself

(e) $B^3 = \begin{bmatrix} 25 & -75 \\ 50 & -150 \end{bmatrix}$

(f) B^{-1} does not exist, since it has RREF $\begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix} \neq I$

Exercise 2.3.4.3

Consider the matrices $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$ and evaluate each expression below (if it exists).

(a) $6B - 5A$

(b) BA

(c) A^{-1}

(d) B^{-1}

(e) $(A - B)^{-1}$

(f) $(AB)^{-1}$

Answer

(a) $6B - 5A = \begin{bmatrix} 1 & -10 & 0 \\ 0 & 7 & 5 \\ -5 & -5 & -18 \end{bmatrix}$

(b) $BA = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & -2 \\ -3 & -3 & 0 \end{bmatrix}$

(c) $A^{-1} = \begin{bmatrix} -1 & 0 & 2 \\ 1 & 0 & -1 \\ 1 & -1 & -1 \end{bmatrix}$

(d) $B^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{3} \end{bmatrix}$

(e) $(A - B)^{-1} = \begin{bmatrix} 1 & 3 & 1 \\ \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & -1 & 0 \end{bmatrix}$

(f) $(AB)^{-1} = \begin{bmatrix} -1 & 0 & 2 \\ \frac{1}{2} & 0 & -\frac{1}{2} \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$

Exercise 2.3.4.4

Consider the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and evaluate each product below.

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} A$

(b) $A \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A$

(d) $A \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix} A$

(f) $A \begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix}$

(g) $\begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix} A$

(h) $A \begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix}$

Answer

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

(b) $A \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A = \begin{bmatrix} c & d \\ a & b \end{bmatrix}$

(d) $A \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}$

(e) $\begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix} A = \begin{bmatrix} a+rc & b+rd \\ c & d \end{bmatrix}$

(f) $A \begin{bmatrix} 1 & r \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & ra+b \\ c & rc+d \end{bmatrix}$

(g) $\begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix} A = \begin{bmatrix} ra & rb \\ c & d \end{bmatrix}$

(h) $A \begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} ra & b \\ rc & d \end{bmatrix}$

? Exercise 2.3.4.5

What relationship between b, c, d must hold in order for $A = \begin{bmatrix} 1 & b \\ c & d \end{bmatrix}$ to have an inverse?

Answer

We row reduce A to echelon form: $\begin{bmatrix} 1 & b \\ c & d \end{bmatrix} \xrightarrow{R_2 = R_2 - cR_1} \begin{bmatrix} 1 & b \\ 0 & d-bc \end{bmatrix}$

Clearly this will only row reduce to I if $d - bc \neq 0$, so our answer is $d \neq bc$

? Exercise 2.3.4.6

Assume that A is a 3×3 matrix satisfying $A \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \vec{e}_1$, $A \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} = \vec{e}_2$, and $A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2\vec{e}_1 + \vec{e}_3$.

What is A^{-1}?

Answer

Define $\vec{v}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$, and $\vec{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, so we can write $A\vec{v}_1 = \vec{e}_1, A\vec{v}_2 = \vec{e}_2$ and $A\vec{v}_3 = 2\vec{e}_1 + \vec{e}_3$

Further define $\vec{w}_1, \vec{w}_2, \vec{w}_3$ to be columns of A^{-1} (if it exists), so that we have

$$AA^{-1} = A\begin{bmatrix} | & | & | \\ \vec{w}_1 & \vec{w}_2 & \vec{w}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} | & | & | \\ A\vec{w}_1 & A\vec{w}_2 & A\vec{w}_3 \\ | & | & | \end{bmatrix} = I$$

To find each $\vec{w}_i$, we just need to solve $A\vec{w}_i = \vec{e}_i$.

The first two of these are given to us! $\vec{w}_1 = \vec{v}_1$ since $A\vec{v}_1 = \vec{e}_1$ and $\vec{w}_2 = \vec{v}_2$ since $A\vec{v}_2 = \vec{e}_2$

The third can be solved by using properties of linear transformations:

$\vec{e}_3 = (2\vec{e}_1 + \vec{e}_3) - 2(\vec{e}_1) = A\vec{v}_3 - 2A\vec{v}_1 = A(\vec{v}_3 - 2\vec{v}_1)$

Therefore $\vec{w}_3 = \vec{v}_3 - 2\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \\ -1 \end{bmatrix}$

Thus we have found $A^{-1} = \begin{bmatrix} 3 & 1 & -5 \\ 1 & 0 & -1 \\ 1 & -2 & -1 \end{bmatrix}$

2.3.4: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.4: Invertibility

Objectives

1. Understand what it means for a square matrix to be invertible.
2. Learn about invertible transformations, and understand the relationship between invertible matrices and invertible transformations.
3. *Recipes:* compute the inverse matrix, solve a linear system by taking inverses.
4. *Picture:* the inverse of a transformation.
5. *Theorem:* the invertible matrix theorem.
6. *Vocabulary words:* ***invertible matrix, inverse transformation***.

In this section, we learn when it is possible to find an inverse of a given linear transformation or matrix.

This page titled 2.4: Invertibility is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **3.5: Matrix Inverses** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **3.6: The Invertible Matrix Theorem** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.4: Invertibility of Linear Transformations

As with matrix multiplication, it is helpful to understand matrix inversion as an operation on linear transformations. Recall that the identity transformation on $\mathbb{R}^n$ is denoted $\mathcal{I}$.

Definition: Invertible

A transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is ***invertible*** if there exists a transformation $U: \mathbb{R}^n \to \mathbb{R}^n$ such that $T \circ U = \mathcal{I}$ and $U \circ T = \mathcal{I}$. In this case, the transformation U is called the ***inverse*** of T, and we write $U = T^{-1}$.

The inverse U of T "undoes" whatever T did. We have

$$T \circ U(\vec{x}) = \vec{x} \quad \text{and} \quad U \circ T(\vec{x}) = \vec{x}$$

for all vectors $\vec{x}$. This means that if you apply T to $\vec{x}$, then you apply U, you get the vector $\vec{x}$ back, and likewise in the other order.

Illustration 2.4.1: Functions of one variable

Define $f: \mathbb{R} \to \mathbb{R}$ by $f(x) = 2x$. This is an invertible transformation, with inverse $g(x) = \dfrac{x}{2}$. Indeed,

$$f \circ g(x) = f(g(x)) = f\left(\frac{x}{2}\right) = 2\left(\frac{x}{2}\right) = x$$

and

$$g \circ f(x) = g(f(x)) = g(2x) = \frac{2x}{2} = x$$

In other words, dividing by 2 undoes the transformation that multiplies by 2.

Define $f: \mathbb{R} \to \mathbb{R}$ by $f(x) = x^3$. This is an invertible transformation, with inverse $g(x) = \sqrt[3]{x}$. Indeed,

$$f \circ g(x) = f(g(x)) = f(\sqrt[3]{x}) = \left(\sqrt[3]{x}\right)^3 = x$$

and

$$g \circ f(x) = g(f(x)) = g(x^3) = \sqrt[3]{x^3} = x$$

In other words, taking the cube root undoes the transformation that takes a number to its cube.

Define $f: \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. This is *not* an invertible function. Indeed, we have $f(2) = 4 = f(-2)$, so there is no way to undo f: the inverse transformation would not know if it should send 2 to 2 or -2. More formally, if $g: \mathbb{R} \to \mathbb{R}$ satisfies $g(f(x)) = x$, then

$$2 = g(f(2)) = g(2) \quad \text{and} \quad -2 = g(f(-2)) = g(2)$$

which is impossible: $g(2)$ is a number, so it cannot be equal to 2 and -2 at the same time.

Define $f: \mathbb{R} \to \mathbb{R}$ by $f(x) = e^x$. This is *not* an invertible function. Indeed, if there were a function $g: \mathbb{R} \to \mathbb{R}$ such that $f \circ g = x$, then we would have

$$-1 = f \circ g(-1) = f(g(-1)) = e^{g(-1)}$$

But e^x is a *positive* number for every x, so this is impossible.

Illustration 2.4.2: Dilation

Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be dilation by a factor of $\dfrac{3}{2}$, that is: $T(\vec{x}) = \dfrac{3}{2}\vec{x}$

We will show that T is an invertible transformation.

Let $U: \mathbb{R}^2 \to \mathbb{R}^2$ be dilation by a factor of $\frac{2}{3}$: that is, $U(\vec{x}) = \frac{2}{3}\vec{x}$. Then

$$T \circ U(\vec{x}) = T\left(\frac{2}{3}\vec{x}\right) = \frac{3}{2} \cdot \frac{2}{3}\vec{x} = \vec{x}$$

and

$$U \circ T(\vec{x}) = U\left(\frac{3}{2}\vec{x}\right) = \frac{2}{3} \cdot \frac{3}{2}\vec{x} = \vec{x}$$

Hence $T \circ U = \mathcal{I}$ and $U \circ T = \mathcal{I}$, so T is invertible, with inverse U. In other words, *shrinking* by a factor of $\frac{2}{3}$ undoes *stretching* by a factor of $\frac{3}{2}$.

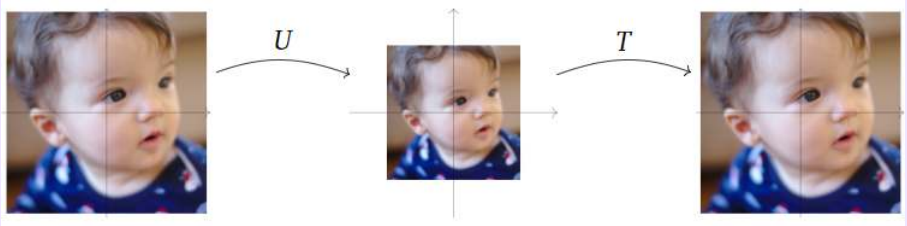


Figure 2.4.1

We can see that the matrix representation of T is $A = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{3}{2} \end{bmatrix}$

Similarly the matrix representation of U is $B = \begin{bmatrix} \frac{2}{3} & 0 \\ 0 & \frac{2}{3} \end{bmatrix}$

It is easy to see that $AB = I$ and $BA = I$, so B is the inverse of A. This hints to us that the matrix representation of T^{-1} may be the inverse of the matrix representation of T.

Illustration 2.4.3: Rotation

Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be counterclockwise rotation by 45°.

We will show that T is an invertible transformation.

Let $U: \mathbb{R}^2 \to \mathbb{R}^2$ be *clockwise* rotation by 45°. Then $T \circ U$ first rotates clockwise by 45°, then counterclockwise by 45°, so the composition rotates by zero degrees: it is the identity transformation. Likewise, $U \circ T$ first rotates counterclockwise, then clockwise by the same amount, so it is the identity transformation. In other words, *clockwise* rotation by 45° undoes *counterclockwise* rotation by 45°.



Figure 2.4.2

This time the matrix representation of T is $A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

And the matrix representation of U is $B = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

Once again we see that $AB = I$ and $BA = I$, so B is the inverse of A as expected.

Illustration 2.4.4: Reflection

Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the reflection over the y-axis.

We will show that T is an invertible transformation.

The transformation T is invertible; in fact, it is equal to its own inverse. Reflecting a vector x over the y-axis twice brings the vector back to where it started, so $T \circ T = \mathcal{I}$.



Figure 2.4.3

The matrix representation of T is $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$

We can see here that $A^2 = AA = I$, so A is its own inverse, just like the transformation.

Illustration 2.4.5: Projection

Let $T: \mathbb{R}^3 \to \mathbb{R}^3$ be the projection onto the xy-plane.

We will show that T is *not* an invertible transformation.

Every vector on the z-axis projects onto the zero vector, so there is no way to undo what T did: the inverse transformation would not know which vector on the z-axis it should send the zero vector to. More formally, suppose there were a transformation $U: \mathbb{R}^3 \to \mathbb{R}^3$ such that $U \circ T = \mathcal{I}$. Then

$$\vec{0} = U \circ T(\vec{0}) = U(T(\vec{0})) = U(\vec{0})$$

and

$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = U \circ T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = U\left(T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right)\right) = U(\vec{0})$$

But $U(\vec{0})$ is a single vector in $\mathbb{R}^3$, so it cannot be equal to $\vec{0}$ and to $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ at the same time.

The matrix representation of T is $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

We attempt to find A^{-1}:

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

But it is already in RREF on the left side, so it is impossible to create I. This means A^{-1} does not exist, and the matrix A is not invertible (just like T)

Proposition 2.4.1

1. A transformation $T: \mathbb{R}^n \to \mathbb{R}^n$ is invertible if and only if it is both one-to-one and onto.
2. If T is already known to be invertible, then $U: \mathbb{R}^n \to \mathbb{R}^n$ is the inverse of T provided that *either* $T \circ U = \mathcal{I}$ or $U \circ T = \mathcal{I}$: it is only necessary to verify one.

Proof

To say that T is one-to-one and onto means that $T(\vec{x}) = \vec{b}$ has *exactly one* solution for every $\vec{b}$ in $\mathbb{R}^n$.

Suppose that T is invertible. Then $T(\vec{x}) = \vec{b}$ always has the unique solution $\vec{x} = T^{-1}(\vec{b})$: indeed, applying T^{-1} to both sides of $T(\vec{x}) = \vec{b}$ gives

$$\vec{x} = T^{-1}(T(\vec{x})) = T^{-1}(\vec{b})$$

and applying T to both sides of $\vec{x} = T^{-1}(\vec{b})$ gives

$$T(\vec{x}) = T(T^{-1}(\vec{b})) = \vec{b}$$

Conversely, suppose that T is one-to-one and onto. Let $\vec{b}$ be a vector in $\mathbb{R}^n$, and let $\vec{x} = U(\vec{b})$ be the unique solution of $T(\vec{x}) = \vec{b}$. Then U defines a transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$. For any $\vec{x}$ in $\mathbb{R}^n$, we have $U(T(\vec{x})) = \vec{x}$, because $\vec{x}$ is the unique solution of the equation $T(\vec{x}) = \vec{b}$ for $\vec{b} = T(\vec{x})$. For any $\vec{b}$ in $\mathbb{R}^n$, we have $T(U(\vec{b})) = \vec{b}$, because $\vec{x} = U(\vec{b})$ is the unique solution of $T(\vec{x}) = \vec{b}$. Therefore, U is the inverse of T, and T is invertible.

Suppose now that T is an invertible transformation, and that U is another transformation such that $T \circ U = \mathcal{I}$. We must show that $U = T^{-1}$, i.e., that $U \circ T = \mathcal{I}$. We compose both sides of the equality $T \circ U = \mathcal{I}$ on the left by T^{-1} and on the right by T to obtain

$$T^{-1} \circ T \circ U \circ T = T^{-1} \circ \mathcal{I} \circ T$$

We have $T^{-1} \circ T = \mathcal{I}$ and $\mathcal{I} \circ U = U$, so the left side of the above equation is $U \circ T$. Likewise, $\mathcal{I} \circ T = T$ and $T^{-1} \circ T = \mathcal{I}$, so our equality simplifies to $U \circ T = \mathcal{I}$, as desired.

If instead we had assumed only that $U \circ T = \mathcal{I}$, then the proof that $T \circ U = \mathcal{I}$ proceeds similarly.

Remark

It makes sense to define the inverse of a transformation $T: \mathbb{R}^n \to \mathbb{R}^m$, for $m \neq n$, to be a transformation $U: \mathbb{R}^m \to \mathbb{R}^n$ such that $T \circ U = \mathcal{I}$ and $U \circ T = \mathcal{I}$. In fact, there exist invertible transformations $T: \mathbb{R}^n \to \mathbb{R}^m$ for any m and n, but they are not linear, or even continuous.

If T is a *linear* transformation, then it can only be invertible when $m = n$, i.e., when its domain is equal to its codomain. Indeed, if $T: \mathbb{R}^n \to \mathbb{R}^m$ is one-to-one, then $n \leq m$ must hold, and if T is onto, then $m \leq n$ must also hold. Therefore, when discussing invertibility we restrict ourselves to the case $m = n$.

As we have been seeing, the matrix for the inverse of a linear transformation is the inverse of the matrix for the transformation:

Theorem 2.4.1

Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation with standard matrix A. Then T is invertible if and only if A is invertible, in which case T^{-1} is linear with standard matrix A^{-1}.

Proof

Suppose that T is invertible. Let $U: \mathbb{R}^n \to \mathbb{R}^n$ be the inverse of T. We claim that U is linear.

Let $\vec{u}, \vec{v}$ be vectors in $\mathbb{R}^n$. Then

$$\vec{u} + \vec{v} = T(U(\vec{u})) + T(U(\vec{v})) = T(U(\vec{u}) + U(\vec{v}))$$

by linearity of T. Applying U to both sides gives

$$U(\vec{u}+\vec{v}) = U\Big(T\left(U(\vec{u})+U(\vec{v})\right)\Big) = U(\vec{u})+U(\vec{v})$$

Let c be a scalar. Then

$$c\vec{u} = cT(U(\vec{u})) = T(cU(\vec{u}))$$

by linearity of T. Applying U to both sides gives

$$U(c\vec{u}) = U\Big(T(cU(\vec{u}))\Big) = cU(\vec{u})$$

Since U satisfies the defining properties, it is a linear transformation.

Now that we know that U is linear, we know that it has a standard matrix B. By the compatibility of matrix multiplication and composition, the matrix for $T \circ U$ is AB. But $T \circ U$ is the identity transformation $\mathcal{I}$, and the standard matrix for $\mathcal{I}$ is I, so $AB = I$. One shows similarly that $BA = I$. Hence A is invertible and $B = A^{-1}$.

Conversely, suppose that A is invertible. Let $B = A^{-1}$, and define $U\colon \mathbb{R}^n \to \mathbb{R}^n$ by $U(\vec{x}) = B\vec{x}$. By the compatibility of matrix multiplication and composition, the matrix for $T \circ U$ is $AB = I$, and the matrix for $U \circ T$ is $BA = I$. Therefore,

$$T \circ U(\vec{x}) = AB\vec{x} = I\vec{x} = \vec{x} \quad \text{and} \quad U \circ T(\vec{x}) = BA\vec{x} = I\vec{x} = \vec{x}$$

which shows that T is invertible with inverse transformation U as desired.

Example 2.4.1: Matrix Transformation example

Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation defined by, $T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} 3a-2b \\ 5b-7a \end{bmatrix}$

Is T invertible? If so, what is T^{-1}?

Solution

Since T is a linear transformation, we know it is a matrix transformation with matrix representation:

$$A = \begin{bmatrix} | & | \\ T(\vec{e}_1) & T(\vec{e}_2) \\ | & | \end{bmatrix} = \begin{bmatrix} 3 & -2 \\ -7 & 5 \end{bmatrix}$$

We know T will have an inverse if A does:

$$[A|I] = \left[\begin{array}{cc|cc} 3 & -2 & 1 & 0 \\ -7 & 5 & 0 & 1 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|cc} 1 & 0 & 5 & 2 \\ 0 & 1 & 7 & 3 \end{array}\right] = [I|A^{-1}]$$

Since A^{-1} exists, we know T^{-1} must also exist and satisfy $T^{-1}(\vec{x}) = A^{-1}\vec{x}$:

$$\boxed{T^{-1}\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} 5 & 2 \\ 7 & 3 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 5a+2b \\ 7a+3b \end{bmatrix}}$$

2.4: Invertibility of Linear Transformations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.5: Matrix Inverses** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.4.2: Invertibility of Matrices

We summarize with a single important theorem containing many equivalent conditions for a matrix to be invertible. This is one of the most important theorems in this textbook. We will append more criteria as we progress in this book.

Theorem 2.4.2.2: Invertible Matrix Theorem

Let A be an $n \times n$ matrix, and let $T: \mathbb{R}^n \to \mathbb{R}^n$ be the matrix transformation $T(\vec{x}) = A\vec{x}$. The following statements are equivalent:

1. A is invertible.
2. A has n pivots.
3. $\text{Nul}(A) = \{\vec{0}\}$.
4. The columns of A are linearly independent.
5. The columns of A span $\mathbb{R}^n$.
6. $A\vec{x} = \vec{b}$ has a unique solution for each $\vec{b}$ in $\mathbb{R}^n$.
7. T is invertible.
8. T is one-to-one.
9. T is onto.

Proof

$(1 \iff 2)$: The matrix A has n pivots if and only if its reduced row echelon form is the identity matrix I. This happens exactly when the procedure to compute the inverse succeeds.

$(2 \iff 3)$: The null space of a matrix is $\{\vec{0}\}$ if and only if the matrix has no free variables, which means that every column is a pivot column, which means A has n pivots.

$(2 \iff 4, 2 \iff 5)$: These follow from how we check for linear independence, since A has n pivots if and only if it has a pivot in every row/column.

$(4 + 5 \iff 6)$: We know $A\vec{x} = \vec{b}$ has at least one solution for every $\vec{b}$ if and only if the columns of A span $\mathbb{R}^n$, and $A\vec{x} = \vec{b}$ has at most one solution for every $\vec{b}$ if and only if the columns of A are linearly independent. Hence $A\vec{x} = \vec{b}$ has exactly one solution for every $\vec{b}$ if and only if its columns are linearly independent and span $\mathbb{R}^n$.

$(1 \iff 7)$:, $(7 \implies 8 + 9)$:, and $(8 \iff 4, 9 \iff 5)$: all follow from our discussion above about invertible transformations.

To reiterate, the invertible matrix theorem means:

Note 2.4.2.1

There are two kinds of *square* matrices:

1. invertible matrices
2. non-invertible matrices

For invertible matrices, all of the statements of the invertible matrix theorem are true.

For non-invertible matrices, all of the statements of the invertible matrix theorem are false.

The reader should be comfortable translating any of the statements in the invertible matrix theorem into a statement about the pivot positions of a matrix.

Note 2.4.2.2: Other Conditions for Invertibility

The following conditions are also equivalent to the invertibility of a square matrix A. They are all simple restatements of conditions in the invertible matrix theorem.

1. The reduced row echelon form of A is the identity matrix I.

2. $A\vec{x} = \vec{0}$ has no solutions other than the trivial one.
3. $\text{nullity}(A) = 0$.
4. The columns of A form a basis for $\mathbb{R}^n$.
5. $A\vec{x} = \vec{b}$ is consistent for all b in $\mathbb{R}^n$.
6. $\text{Col}(A) = \mathbb{R}^n$.
7. $\dim \text{Col}(A) = n$.
8. $\text{rank}(A) = n$.

Now we can show that to check $B = A^{-1}$, it's enough to show $AB = I$ *or* $BA = I$.

Corollary 2.4.2.1: A Left or Right Inverse Suffices

Let A be an $n \times n$ matrix, and suppose that there exists an $n \times n$ matrix B such that $AB = I$ *or* $BA = I$. Then A is invertible and $B = A^{-1}$.

Proof

Suppose that $AB = I$. We claim that $T(\vec{x}) = A\vec{x}$ is onto. Indeed, for any $\vec{b}$ in $\mathbb{R}^n$, we have

$$\vec{b} = I\vec{b} = (AB)\vec{b} = A(B\vec{b})$$

so $T(B\vec{b}) = \vec{b}$, and hence $\vec{b}$ is in the range of T. Therefore, A is invertible by the Invertible Matrix Theorem. Since A is invertible, we have

$$A^{-1} = A^{-1}I = A^{-1}(AB) = (A^{-1}A)B = IB = B$$

so $B = A^{-1}$.

Now suppose that $BA = I$. We claim that $T(\vec{x}) = A\vec{x}$ is one-to-one. Indeed, suppose that $T(\vec{x}) = T(\vec{y})$. Then $A\vec{x} = A\vec{y}$, so $BA\vec{x} = BA\vec{y}$. But $BA = I$, so $I\vec{x} = I\vec{y}$, and hence $\vec{x} = \vec{y}$. Therefore, A is invertible by the Theorem 2.4.2.2. One shows that $B = A^{-1}$ as above.

We conclude with some common situations in which the invertible matrix theorem is useful.

Example 2.4.2.2

Is this matrix invertible?

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & 7 \\ -2 & -4 & 1 \end{bmatrix}$$

Solution

The second column is a multiple of the first. The columns of A are linearly dependent, so A does not satisfy condition 4 of the Invertible Matrix Theorem. Therefore, A is $\boxed{\text{not invertible}}$

Example 2.4.2.3

Let A be an $n \times n$ matrix and let $T(\vec{x}) = A\vec{x}$. Suppose that the range of T is $\mathbb{R}^n$. Show that the columns of A are linearly independent.

Solution

The range of T is the column space of A, so A satisfies condition 5 of the Invertible Matrix Theorem. Therefore, A also satisfies condition 4, which says that the columns of A are linearly independent.

Thus, the matrix is $\boxed{\text{invertible}}$

Example 2.4.2.4

Let A be a 3×3 matrix such that

$$A\begin{bmatrix}1\\7\\0\end{bmatrix} = A\begin{bmatrix}2\\0\\-1\end{bmatrix}$$

Show that the rank of A is at most 2.

Solution

If we set

$$\vec{b} = A\begin{bmatrix}1\\7\\0\end{bmatrix} = A\begin{bmatrix}2\\0\\-1\end{bmatrix}$$

then $A\vec{x} = \vec{b}$ has multiple solutions, so it does not satisfy condition 6 of the Invertible Matrix Theorem. Therefore, it does not satisfy condition 5, so the columns of A do not span $\mathbb{R}^3$. Therefore, the column space has dimension strictly less than 3, the rank is at most 2.

Thus, the matrix is $\boxed{\text{not invertible}}$

Example 2.4.2.5

Suppose that A is an $n \times n$ matrix such that $A\vec{x} = \vec{b}$ is inconsistent for some vector $\vec{b}$. Show that $A\vec{x} = \vec{c}$ has infinitely many solutions for some other vector $\vec{c}$.

Solution

By hypothesis, A does not satisfy condition 6 of the Invertible Matrix Theorem. Therefore, it does not satisfy condition 3, so $\text{Nul}(A)$ is an infinite set. If we take $\vec{c} = \vec{0}$, then the equation $A\vec{x} = \vec{c}$ has infinitely many solutions.

Thus, the matrix is $\boxed{\text{not invertible}}$

Next, we learn to solve $A\vec{x} = \vec{b}$ by using an inverse.

Theorem 2.4.2.3

Let A be an invertible $n \times n$ matrix, and let $\vec{b}$ be a vector in $\mathbb{R}^n$. Then the matrix equation $A\vec{x} = \vec{b}$ has exactly one solution:

$$\vec{x} = A^{-1}\vec{b}$$

Proof

We calculate:

$$\begin{aligned} A\vec{x} = \vec{b} &\implies A^{-1}(A\vec{x}) = A^{-1}\vec{b} \\ &\implies (A^{-1}A)\vec{x} = A^{-1}\vec{b} \\ &\implies I\vec{x} = A^{-1}\vec{b} \\ &\implies \vec{x} = A^{-1}\vec{b} \end{aligned}$$

Here we used associativity of matrix multiplication, and the fact that $I\vec{x} = \vec{x}$ for any vector $\vec{x}$.

Example 2.4.2.6: Solving a 2×2 system using inverses

Solve the matrix equation

$$\begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Solution

By the Theorem above, the only solution of our linear system is

$$\vec{x} = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{5} & -\frac{3}{5} \\ \frac{1}{5} & \frac{1}{5} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \boxed{\begin{bmatrix} -\frac{1}{5} \\ \frac{2}{5} \end{bmatrix}}$$

Example 2.4.2.7: Solving a 3×3 system using inverses

Solve the system of equations

$$\begin{cases} 2x_1 + 3x_2 + 2x_3 = 1 \\ x_1 + 3x_3 = 1 \\ 2x_1 + 2x_2 + 3x_3 = 1 \end{cases}$$

Solution

First, we write our system as a matrix equation $A\vec{x} = \vec{b}$, where

$$A = \begin{bmatrix} 2 & 3 & 2 \\ 1 & 0 & 3 \\ 2 & 2 & 3 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Next we find the inverse of A by augmenting and row reducing:

$$\left[\begin{array}{ccc|ccc} 2 & 3 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 2 & 3 & 2 & 1 & 0 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_3 = R_3 - 2R_1]{R_2 = R_2 - 2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 3 & -4 & 1 & -2 & 0 \\ 0 & 2 & -3 & 0 & -2 & 1 \end{array}\right]$$

$$\xrightarrow{R_2 = R_2 - R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 2 & -3 & 0 & -2 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 = R_3 - 2R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & -1 & -2 & -2 & 3 \end{array}\right]$$

$$\xrightarrow{R_3 = -R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 & 2 & -3 \end{array}\right]$$

$$\xrightarrow[R_2 = R_2 + R_3]{R_1 = R_1 - 3R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -6 & -5 & 9 \\ 0 & 1 & 0 & 3 & 2 & -4 \\ 0 & 0 & 1 & 2 & 2 & -3 \end{array}\right]$$

Therefore, the only solution of our linear system is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 2 \\ 1 & 0 & 3 \\ 2 & 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -6 & -5 & 9 \\ 3 & 2 & -4 \\ 2 & 2 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \boxed{\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}}$$

The advantage of solving a linear system using inverses is that it becomes much faster to solve the matrix equation $A\vec{x} = \vec{b}$ for other, or even unknown, values of $\vec{b}$. For instance, in the above example, the solution of the system of equations

$$\begin{cases} 2x_1 + 3x_2 + 2x_3 = b_1 \\ x_1 + 3x_3 = b_2 \\ 2x_1 + 2x_2 + 3x_3 = b_3 \end{cases}$$

where b_1, b_2, b_3 are unknowns, is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 2 \\ 1 & 0 & 3 \\ 2 & 2 & 3 \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} -6 & -5 & 9 \\ 3 & 2 & -4 \\ 2 & 2 & -3 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} -6b_1 - 5b_2 + 9b_3 \\ 3b_1 + 2b_2 - 4b_3 \\ 2b_1 + 2b_2 - 3b_3 \end{bmatrix}$$

2.4.2: Invertibility of Matrices is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **3.6: The Invertible Matrix Theorem** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **3.5: Matrix Inverses** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.4: Practice Set

Exercise 2.4.1a

Find the inverse of the transformation $T:\mathbb{R}^2\to\mathbb{R}^2$ defined by $T\left(\begin{bmatrix}a\\b\end{bmatrix}\right)=\begin{bmatrix}a+b\\2a+3b\end{bmatrix}$ or explain why it does not exist.

Answer

First we find $T(\vec{e}_1)=\begin{bmatrix}1\\2\end{bmatrix}$ and $T(\vec{e}_2)=\begin{bmatrix}1\\3\end{bmatrix}$

Therefore the matrix representation of T is $A=\begin{bmatrix}1&1\\2&3\end{bmatrix}$

Now we find the inverse of this matrix:

$$[A|I]=\left[\begin{array}{cc|cc}1&1&1&0\\2&3&0&1\end{array}\right]\xrightarrow{RREF}\left[\begin{array}{cc|cc}1&0&3&-1\\0&1&-2&1\end{array}\right]=[I|A^{-1}]$$

Therefore $A^{-1}=\begin{bmatrix}3&-1\\-2&1\end{bmatrix}$ which gives

$$T^{-1}\left(\begin{bmatrix}a\\b\end{bmatrix}\right)=\begin{bmatrix}3&-1\\-2&1\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}=\begin{bmatrix}3a-b\\-2a+b\end{bmatrix}$$

Exercise 2.4.1b

Find the inverse of the transformation $T:\mathbb{R}^2\to\mathbb{R}^2$ defined by $T\left(\begin{bmatrix}a\\b\end{bmatrix}\right)=\begin{bmatrix}b\\6b\end{bmatrix}$ or explain why it does not exist.

Answer

First we find $T(\vec{e}_1)=\begin{bmatrix}0\\0\end{bmatrix}$ and $T(\vec{e}_2)=\begin{bmatrix}1\\6\end{bmatrix}$

Therefore the matrix representation of T is $A=\begin{bmatrix}0&1\\0&6\end{bmatrix}$

Now we find the inverse of this matrix:

$$[A|I]=\left[\begin{array}{cc|cc}0&1&1&0\\0&6&0&1\end{array}\right]\xrightarrow{RREF}\left[\begin{array}{cc|cc}0&1&0&\frac{1}{6}\\0&0&1&-\frac{1}{6}\end{array}\right]$$

We see A is not invertible (we could have seen that without finding RREF, since the columns are obviously dependent)

Therefore T is not invertible.

Exercise 2.4.1c

Find the inverse of the transformation $T:\mathbb{R}^3\to\mathbb{R}^3$ defined by $T\left(\begin{bmatrix}a\\b\\c\end{bmatrix}\right)=\begin{bmatrix}a-b-2c\\8b+13c\\3b+5c\end{bmatrix}$ or explain why it does not exist.

Answer

First we find $T(\vec{e}_1)=\begin{bmatrix}1\\0\\0\end{bmatrix}$ and $T(\vec{e}_2)=\begin{bmatrix}-1\\8\\3\end{bmatrix}$ and $T(\vec{e}_3)=\begin{bmatrix}-2\\13\\5\end{bmatrix}$

Therefore the matrix representation of T is $A=\begin{bmatrix}1&-1&-2\\0&8&13\\0&3&5\end{bmatrix}$

Now we find the inverse of this matrix:

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & -1 & -2 & 1 & 0 & 0 \\ 0 & 8 & 13 & 0 & 1 & 0 \\ 0 & 3 & 5 & 0 & 0 & 1 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -1 & 3 \\ 0 & 1 & 0 & 0 & 5 & -13 \\ 0 & 0 & 1 & 0 & -3 & 8 \end{array}\right] = [I|A^{-1}]$$

Therefore $A^{-1} = \begin{bmatrix} 1 & -1 & 3 \\ 0 & 5 & -13 \\ 0 & -3 & 8 \end{bmatrix}$ which gives

$$T^{-1}\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = \begin{bmatrix} 1 & -1 & 3 \\ 0 & 5 & -13 \\ 0 & -3 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a-b+3c \\ 5b-13c \\ -3b+8c \end{bmatrix}$$

Exercise 2.4.1d

Find the inverse of the transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by $T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = \begin{bmatrix} 11a+7c \\ b-3c \end{bmatrix}$ or explain why it does not exist.

Answer

First we find $T(\vec{e}_1) = \begin{bmatrix} 11 \\ 0 \end{bmatrix}$ and $T(\vec{e}_2) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $T(\vec{e}_3) = \begin{bmatrix} 7 \\ -3 \end{bmatrix}$

Therefore the matrix representation of T is $A = \begin{bmatrix} 11 & 0 & 7 \\ 0 & 1 & -3 \end{bmatrix}$

This matrix is not square, so it cannot be invertible. Therefore T is not invertible.

Exercise 2.4.1e

Find the inverse of the transformation $T : \mathbb{R}^4 \to \mathbb{R}^4$ defined by $T\left(\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}\right) = \begin{bmatrix} a+b \\ b+2c \\ c+3d \\ d \end{bmatrix}$ or explain why it does not exist.

Answer

First we find $T(\vec{e}_1) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and $T(\vec{e}_2) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and $T(\vec{e}_3) = \begin{bmatrix} 0 \\ 2 \\ 1 \\ 0 \end{bmatrix}$ and $T(\vec{e}_4) = \begin{bmatrix} 0 \\ 0 \\ 3 \\ 1 \end{bmatrix}$

Therefore the matrix representation of T is $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Now we find the inverse of this matrix:

$$[A|I] = \left[\begin{array}{cccc|cccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -1 & 2 & -6 \\ 0 & 1 & 0 & 0 & 0 & 1 & -2 & 6 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] = [I|A^{-1}]$$

Therefore $A^{-1} = \begin{bmatrix} 1 & -1 & 2 & -6 \\ 0 & 1 & -2 & 6 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ which gives

$$T^{-1}\left(\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}\right) = \begin{bmatrix} 1 & -1 & 2 & -6 \\ 0 & 1 & -2 & 6 \\ 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a-b+2c-6d \\ b-2c+6d \\ c-3d \\ d \end{bmatrix}$$

Exercise 2.4.2

Consider the matrix $A = \begin{bmatrix} 4 & -3 \\ -5 & 4 \end{bmatrix}$

(a) Find A^{-1} if it exists.

(b) Find a matrix X which satisfies $AX = \begin{bmatrix} 7 & 8 \\ 0 & -3 \end{bmatrix}$ or explain why one does not exist.

(c) Find a matrix X which satisfies $XA = \begin{bmatrix} 7 & 8 \\ 0 & -3 \end{bmatrix}$ or explain why one does not exist.

(d) Find a matrix X which satisfies $AX = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ or explain why one does not exist.

(e) Find a matrix X which satisfies $XA = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ or explain why one does not exist.

Answer

(a) $[A|I] = \left[\begin{array}{cc|cc} 4 & -3 & 1 & 0 \\ -5 & 4 & 0 & 1 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|cc} 1 & 0 & 4 & 3 \\ 0 & 1 & 5 & 4 \end{array}\right] = [I|A^{-1}]$

Therefore $A^{-1} = \begin{bmatrix} 4 & 3 \\ 5 & 4 \end{bmatrix}$

(b) $AX = B \longrightarrow A^{-1}AX = A^{-1}B \longrightarrow X = A^{-1}B$ (if it exists)

$X = A^{-1}B = \begin{bmatrix} 4 & 3 \\ 5 & 4 \end{bmatrix}\begin{bmatrix} 7 & 8 \\ 0 & -3 \end{bmatrix} = \begin{bmatrix} 28 & 23 \\ 35 & 28 \end{bmatrix}$

(c) $XA = B \longrightarrow XAA^{-1} = BA^{-1} \longrightarrow X = BA^{-1}$ (if it exists)

$X = BA^{-1} = \begin{bmatrix} 7 & 8 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 5 & 4 \end{bmatrix} = \begin{bmatrix} 68 & 53 \\ -15 & -12 \end{bmatrix}$

(d) $AX = C \longrightarrow A^{-1}AX = A^{-1}C \longrightarrow X = A^{-1}C$ (if it exists)

$X = A^{-1}C = \begin{bmatrix} 4 & 3 \\ 5 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 16 & 23 & 30 \\ 21 & 30 & 39 \end{bmatrix}$

(e) $XA = C \longrightarrow XAA^{-1} = CA^{-1} \longrightarrow X = CA^{-1}$ (if it exists)

Since C is 2×3 and A^{-1} is 2×2 we cannot multiply these matrices in this order, so no such X can exist.

Exercise 2.4.3

Consider the matrix $A = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 7 & 7 \\ 1 & 5 & 6 \end{bmatrix}$

(a) Find A^{-1} if it exists.

(b) Find a matrix X which satisfies $AXA^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ or explain why one does not exist.

(c) Find a matrix X which satisfies $A^{-1}XA = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ or explain why one does not exist.

(d) Find a matrix X which satisfies $A^2X = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ or explain why one does not exist.

(e) Find a matrix X which satisfies $XA + A^2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ or explain why one does not exist.

Answer

(a) $[A|I] = \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 1 & 7 & 7 & 0 & 1 & 0 \\ 1 & 5 & 6 & 0 & 0 & 1 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 7 & 15 & -21 \\ 0 & 1 & 0 & 1 & 3 & -4 \\ 0 & 0 & 1 & -2 & -5 & 7 \end{array}\right] = [I|A^{-1}]$

Therefore $A^{-1} = \left[\begin{array}{ccc} 7 & 15 & -21 \\ 1 & 3 & -4 \\ -2 & -5 & 7 \end{array}\right]$

(b) $AXA^{-1} = B \longrightarrow A^{-1}AXA^{-1}A = A^{-1}BA \longrightarrow X = A^{-1}BA$ (if it exists)

$$X = A^{-1}BA = \left[\begin{array}{ccc} 7 & 15 & -21 \\ 1 & 3 & -4 \\ -2 & -5 & 7 \end{array}\right]\left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right]\left[\begin{array}{ccc} 1 & 0 & 3 \\ 1 & 7 & 7 \\ 1 & 5 & 6 \end{array}\right] = \left[\begin{array}{ccc} 31 & 75 & 135 \\ 6 & 16 & 27 \\ -10 & -25 & -44 \end{array}\right]$$

(c) $A^{-1}XA = B \longrightarrow AA^{-1}XAA^{-1} = ABA^{-1} \longrightarrow X = ABA^{-1}$ (if it exists)

$$X = ABA^{-1} = \left[\begin{array}{ccc} 1 & 0 & 3 \\ 1 & 7 & 7 \\ 1 & 5 & 6 \end{array}\right]\left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right]\left[\begin{array}{ccc} 7 & 15 & -21 \\ 1 & 3 & -4 \\ -2 & -5 & 7 \end{array}\right] = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 35 & 71 & -98 \\ 25 & 50 & -69 \end{array}\right]$$

(d) $A^2X = B \longrightarrow (A^{-1})^2 A^2X = (A^{-1})^2 B \longrightarrow X = (A^{-1})^2 B$ (if it exists)

$$X = (A^{-1})^2 B = \left[\begin{array}{ccc} 7 & 15 & -21 \\ 1 & 3 & -4 \\ -2 & -5 & 7 \end{array}\right]\left[\begin{array}{ccc} 7 & 15 & -21 \\ 1 & 3 & -4 \\ -2 & -5 & 7 \end{array}\right]\left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{ccc} 361 & 255 & -99 \\ 62 & 44 & -17 \\ -113 & -80 & 31 \end{array}\right]$$

(e) $XA + A^2 = B \longrightarrow XA = B - A^2 \longrightarrow XAA^{-1} = (B - A^2)A^{-1} \longrightarrow X = (B - A^2)A^{-1}$ (if it exists)

$$X = (B - A^2)A^{-1} = \left(\left[\begin{array}{ccc} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right] - \left[\begin{array}{ccc} 1 & 0 & 3 \\ 1 & 7 & 7 \\ 1 & 5 & 6 \end{array}\right]^2\right)\left[\begin{array}{ccc} 7 & 15 & -21 \\ 1 & 3 & -4 \\ -2 & -5 & 7 \end{array}\right] = \left[\begin{array}{ccc} 6 & 15 & -24 \\ 5 & 6 & -25 \\ -3 & -10 & 1 \end{array}\right]$$

? Exercise 2.4.4

Determine the value(s) of h which make each matrix invertible.

(a) $A = \left[\begin{array}{cc} 1 & h \\ h & 4 \end{array}\right]$

(b) $B = \left[\begin{array}{ccc} h-1 & 2 & 3 \\ 0 & h-4 & 5 \\ 0 & 0 & h-6 \end{array}\right]$

(c) $C = \left[\begin{array}{ccc} h & 1 & 1 \\ 1 & h & 1 \\ 1 & 1 & h \end{array}\right]$

Answer

(a) $\left[\begin{array}{cc} 1 & h \\ h & 4 \end{array}\right] \xrightarrow{R_2 = R_2 - hR_1} \left[\begin{array}{cc} 1 & h \\ 0 & 4-h^2 \end{array}\right]$

This matrix will have a pivot in every column only when $4 - h^2 \neq 0$, so it will be invertible when $h \neq \pm 2$

(b) $\left[\begin{array}{ccc} h-1 & 2 & 3 \\ 0 & h-4 & 5 \\ 0 & 0 & h-6 \end{array}\right]$ is already in echelon form

There will be a pivot in every column only when $h - 1 \neq 0$ AND $h - 4 \neq 0$ AND $h - 6 \neq 0$, so it is invertible when $h \neq 1, 4, 6$

(c) $\left[\begin{array}{ccc} h & 1 & 1 \\ 1 & h & 1 \\ 1 & 1 & h \end{array}\right] \xrightarrow{R_1 = R_1 - hR_3} \left[\begin{array}{ccc} 0 & 1-h & 1-h^2 \\ 1 & h & 1 \\ 1 & 1 & h \end{array}\right] \xrightarrow{R_2 = R_2 - R_3} \left[\begin{array}{ccc} 0 & 1-h & 1-h^2 \\ 0 & h-1 & 1-h \\ 1 & 1 & h \end{array}\right] \xrightarrow{R_1 = R_1 + R_2} \left[\begin{array}{ccc} 0 & 0 & 2-h-h^2 \\ 0 & h-1 & 1-h \\ 1 & 1 & h \end{array}\right]$

There will be a pivot in every row only when $h - 1 \neq 0$ and $2 - h - h^2 \neq 0$, so it is invertible when $h \neq 1, -2$

Exercise 2.4.5

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If A is a 2×2 matrix which satisfies $A \begin{bmatrix} \pi \\ \pi^2 \end{bmatrix} = \vec{0}$, then the columns of A are linearly dependent.

(b) If A is a 2×2 matrix which satisfies $A \begin{bmatrix} \sqrt{2} \\ \sqrt{3} \end{bmatrix} = \vec{0}$, then the columns of A span all of $\mathbb{R}^2$

(c) If A is a 2×2 matrix which satisfies $A \begin{bmatrix} 1 \\ 7 \end{bmatrix} = A \begin{bmatrix} 3 \\ -1 \end{bmatrix}$, then the columns of A are linearly independent.

(d) If A is a 3×3 matrix which satisfies $A \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \vec{0}$, then A is invertible.

(e) If A is a $n \times n$ invertible matrix, then the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ defined by $T(\vec{x}) = A\vec{x}$ is onto.

(f) If A is a $n \times n$ invertible matrix, then the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ defined by $T(\vec{x}) = A\vec{x}$ is one-to-one.

(g) If the function $T : \mathbb{R}^n \to \mathbb{R}^n$ defined by $T(\vec{x}) = A\vec{x}$ is invertible, then the matrix A has linearly independent columns.

(h) If A, B, and C are invertible $n \times n$ matrices, then ABC is invertible.

(i) If A, B, and C are invertible $n \times n$ matrices satisfying $AB = CB$ then $A = C$.

(j) If A, B, and C are $n \times n$ matrices satisfying $AC = BC$ then $A = B$.

(k) The sum of two invertible linear transformations is an invertible linear transformation.

Answer

(a) **Always True**

The existence of nontrivial solutions to $A\vec{x} = \vec{0}$ means A is not invertible, and has linearly dependent columns.

(b) **Always False**

The existence of nontrivial solutions to $A\vec{x} = \vec{0}$ means A is not invertible, and it cannot be onto.

(c) **Always False**

Since A is not one-to-one, it is not invertible and must have linearly dependent columns.

(d) **Sometimes True Sometimes False**

This is true if there are no other homogeneous solutions, but is false when there are nontrivial homogeneous solutions.

(e) **Always True**

If A is invertible, then T must be invertible and therefore onto.

(f) **Always True**

If A is invertible, then T must be invertible and therefore one-to-one.

(g) **Always True**

If T is invertible, then A must be invertible and must have linearly independent columns.

(h) **Always True**

$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$ since $ABC \cdot C^{-1}B^{-1}A^{-1} = ABIB^{-1}A^{-1} = ABB^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$.

(i) **Always True**

$ABB^{-1} = CBB^{-1} \longrightarrow AI = CI \longrightarrow A = C$.

(j) **Sometimes True Sometimes False**

True when C is invertible, but false when C is not invertible.

(k) **Sometimes True Sometimes False**

True for the example $\mathcal{I} + \mathcal{I}$, which is invertible, and false for the example $\mathcal{I} + (-\mathcal{I})$, which is not invertible.

2.4: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.5: Determinants- Definition

Objectives

1. Learn the definition of the determinant.
2. Learn some ways to eyeball a matrix with zero determinant, and how to compute determinants of upper- and lower-triangular matrices.
3. Learn the basic properties of the determinant, and how to apply them.
4. *Recipe:* compute the determinant using row and column operations.
5. *Theorems:* existence theorem, invertibility property, multiplicativity property, transpose property.
6. *Vocabulary words:* ***diagonal***, ***upper-triangular***, ***lower-triangular***, ***transpose***.
7. *Essential vocabulary word:* ***determinant***.

In this section, we define the determinant, and we present one way to compute it. Then we discuss some of the many wonderful properties the determinant enjoys.

This page titled 2.5: Determinants- Definition is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **4.1: Determinants- Definition** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.5.1: Definition of Determinant

The determinant of a square matrix A is a real number $\det(A)$. It is defined via its behavior with respect to row operations; this means we can use row reduction to compute it. We will give a recursive formula for the determinant in the next section.

Definition: Determinant

The ***determinant*** is a function

$$\det: \{\text{square matrices}\} \to \mathbb{R}$$

satisfying the following properties:

1. Doing a row replacement on A does not change $\det(A)$.
2. Scaling a row of A by a scalar c multiplies the determinant by c.
3. Swapping two rows of a matrix multiplies the determinant by -1.
4. The determinant of the identity matrix I is equal to 1.

A common shorthand notation is to write the determinant of a matrix by using the absolute value on a matrix: $\det(A) = |A|$

In other words, to every square matrix A we assign a number $\det(A)$ in a way that satisfies the above properties.

In each of the first three cases, doing a row operation on a matrix scales the determinant by a *nonzero* number. (Multiplying a row by zero is not a row operation.) Therefore, doing row operations on a square matrix A does not change whether or not the determinant is zero.

The main motivation behind using these particular defining properties is geometric, which we will discuss later. Another motivation for this definition is that it tells us how to compute the determinant: we row reduce and keep track of the changes.

Illustration 2.5.1.1

Let us compute $\det\left(\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}\right)$. Assume the determinant is d. We row reduce, and keep track of how the determinant changes in order to get an equation to solve for it:

$$\begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix} \qquad \det = d$$

$$\xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix} \qquad \det = -1 \cdot d$$

$$\xrightarrow{R_2 = R_2 - 2R_1} \begin{bmatrix} 1 & 4 \\ 0 & -7 \end{bmatrix} \qquad \det = 1 \cdot (-1) \cdot d$$

$$\xrightarrow{R_2 = -\frac{1}{7}R_2} \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix} \qquad \det = -\frac{1}{7} \cdot 1 \cdot (-1) \cdot d$$

$$\xrightarrow{R_1 = R_1 - 4R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \det = 1 \cdot \left(-\frac{1}{7}\right) \cdot 1 \cdot (-1) \cdot d$$

The reduced row echelon form of the matrix is the identity matrix I, so its determinant is 1.

We now solve $1 = 1 \cdot \left(-\frac{1}{7}\right) \cdot 1 \cdot (-1) \cdot d$ and get $d = 7$

Thus the determinant of this matrix is 7, or we can write $\begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = 7$

Example 2.5.1.1

Compute $\begin{vmatrix} 1 & 0 \\ 0 & 3 \end{vmatrix}$

Solution

Let $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$. Since A is obtained from I by multiplying the second row by the constant 3, we have $\det(A) = 3\det(I) = 3 \cdot 1 = \boxed{3}$.

In this case, since we can easily see how to get from I to A, we do not need to set up and solve an equation to get the determinant.

Example 2.5.1.2

Compute $d = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 5 & 1 & 0 \end{vmatrix}$

Solution

First we row reduce, then we compute the determinant in the opposite order:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 5 & 1 & 0 \end{bmatrix} \quad \det = d$$

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \det = -1 \cdot d$$

$$\xrightarrow{R_2 = R_2 - 5R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \det = 1 \cdot (-1) \cdot d$$

The reduced row echelon form is I, which has determinant 1.

We solve $1 \cdot (-1) \cdot d = 1$ to easily find that our determinant is $d = \boxed{-1}$.

Here is the general method for computing determinants using row reduction.

Recipe: Computing Determinants by Row Reducing

Let A be a square matrix. Suppose that you do some number of row operations on A to obtain a matrix B in row echelon form. Then

$$\det(A) = (-1)^r \cdot \frac{(\text{product of the diagonal entries of } B)}{(\text{product of all scaling factors used})}$$

where r is the number of row swaps performed.

In other words, the determinant of A is the product of diagonal entries of the row echelon form B, times a factor of ± 1 coming from the number of row swaps you made, divided by the product of the scaling factors used in the row reduction.

Remark

This is an efficient way of computing the determinant of a large matrix, either by hand or by computer. The computational complexity of row reduction is $O(n^3)$; by contrast, the cofactor expansion algorithm we will learn in the next section has complexity $O(n!) \approx O(n^n\sqrt{n})$, which is much larger. (But, cofactor expansion has other uses to make up for this!)

Example 2.5.1.3

Compute $\begin{vmatrix} 0 & -7 & -4 \\ 2 & 4 & 6 \\ 3 & 7 & -1 \end{vmatrix}$

Solution

We row reduce the matrix, keeping track of the number of row swaps and of the scaling factors used.

$$\begin{bmatrix} 0 & -7 & -4 \\ 2 & 4 & 6 \\ 3 & 7 & -1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 2 & 4 & 6 \\ 0 & -7 & -4 \\ 3 & 7 & -1 \end{bmatrix} \quad \text{row swap}$$

$$\xrightarrow{R_1 = \frac{1}{2} R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -7 & -4 \\ 3 & 7 & -1 \end{bmatrix} \quad \text{scaling factor} = \frac{1}{2}$$

$$\xrightarrow{R_3 = R_3 - 3R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -7 & -4 \\ 0 & 1 & -10 \end{bmatrix}$$

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -10 \\ 0 & -7 & -4 \end{bmatrix} \quad \text{row swap}$$

$$\xrightarrow{R_3 = R_3 + 7R_2} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & -10 \\ 0 & 0 & -74 \end{bmatrix}$$

We made two row swaps and scaled once by a factor of $\frac{1}{2}$, so we see that according to the above shortcut:

$$\begin{vmatrix} 0 & -7 & -4 \\ 2 & 4 & 6 \\ 3 & 7 & -1 \end{vmatrix} = (-1)^2 \cdot \frac{1 \cdot 1 \cdot (-74)}{\frac{1}{2}} = \boxed{-148}$$

Example 2.5.1.4

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

Solution

We row reduce the matrix, keeping track of the number of row swaps and of the scaling factors used.

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & 1 \end{bmatrix} \xrightarrow[R_3 = R_3 - 3R_1]{R_2 = R_2 - 2R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -5 & -5 \\ 0 & -6 & -8 \end{bmatrix}$$

$$\xrightarrow{R_2 = -\frac{1}{5} R_2} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & -6 & -8 \end{bmatrix} \quad \text{scaling factor } = -\frac{1}{5}$$

$$\xrightarrow{R_3 = R_3 + 6R_2} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & -2 \end{bmatrix}$$

We did not make any row swaps, and we scaled once by a factor of $-\frac{1}{5}$, so we see that according to the above shortcut:

$$\begin{vmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & 1 \end{vmatrix} = \frac{1 \cdot 1 \cdot (-2)}{-\frac{1}{5}} = \boxed{10}$$

Illustration 2.5.1.2: The Determinant of a 2×2 Matrix

Let us use the shortcut to compute the determinant of a general 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

- If $a = 0$, then

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} 0 & b \\ c & d \end{vmatrix} = -\begin{vmatrix} c & d \\ 0 & b \end{vmatrix} = -bc$$

- If $a \neq 0$, then

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a \cdot \begin{vmatrix} 1 & \frac{b}{a} \\ c & d \end{vmatrix} = a \cdot \begin{vmatrix} 1 & \frac{b}{a} \\ 0 & d - c \cdot \frac{b}{a} \end{vmatrix}$$
$$= a \cdot 1 \cdot \left(d - \frac{bc}{a} \right) = ad - bc$$

In either case, we recover $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ This is a valuable shortcut to memorize!

If a matrix is already in row echelon form, then you can simply read off the determinant as the product of the diagonal entries. It turns out this is true for a slightly larger class of matrices called *triangular*.

Definition: Diagonal

- The ***diagonal*** entries of a matrix A are the entries $a_{11}, a_{22}, \ldots$:

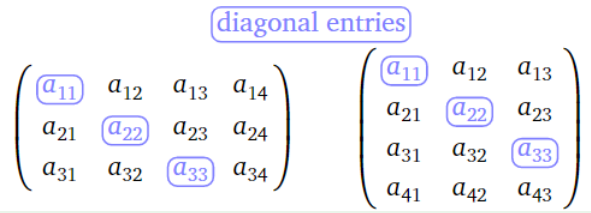


Figure 2.5.1.1

- A square matrix is called ***upper-triangular*** if its nonzero entries all lie above the diagonal, and it is called ***lower-triangular*** if its nonzero entries all lie below the diagonal. It is called ***diagonal*** if all of its nonzero entries lie on the diagonal, i.e., if it is both upper-triangular and lower-triangular.

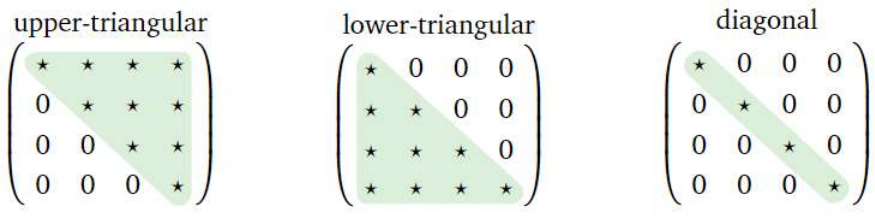


Figure 2.5.1.2

Proposition 2.5.1.1

Let A be an $n \times n$ matrix.

1. If A has a zero row or column, then $\det(A) = 0$.
2. If A is upper-triangular or lower-triangular, then $\det(A)$ is the product of its diagonal entries.

Proof

1. Suppose that A has a zero row. Let B be the matrix obtained by negating the zero row. Then $\det(A) = -\det(B)$ by the second defining property of determinants above. But $A = B$, so $\det(A) = \det(B)$:

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \end{bmatrix} \xrightarrow{R_2 = -R_2} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \end{bmatrix}$$

Putting these together yields $\det(A) = -\det(A)$, so $\det(A) = 0$.
Now suppose that A has a zero column. Then A is not invertible (since it has linearly dependent columns), so its reduced row echelon form has a zero row. Since row operations do not change whether the determinant is zero, we conclude $\det(A) = 0$.

2. First suppose that A is upper-triangular, and that one of the diagonal entries is zero, say $a_{ii} = 0$. We can perform row operations to clear the entries above the nonzero diagonal entries:

$$\begin{bmatrix} a_{11} & \star & \star & \star \\ 0 & a_{22} & \star & \star \\ 0 & 0 & 0 & \star \\ 0 & 0 & 0 & a_{44} \end{bmatrix} \longrightarrow \begin{bmatrix} a_{11} & 0 & \star & 0 \\ 0 & a_{22} & \star & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

In the resulting matrix, the ith row is zero, so $\det(A) = 0$ by the first part.
Still assuming that A is upper-triangular, now suppose that all of the diagonal entries of A are nonzero. Then A can be transformed to the identity matrix by scaling the diagonal entries and then doing row replacements:

$$\underset{\det = abc}{\begin{bmatrix} a & \star & \star \\ 0 & b & \star \\ 0 & 0 & c \end{bmatrix}} \xrightarrow[\longleftarrow]{\text{scale by } \frac{1}{a}, \frac{1}{b}, \frac{1}{c}} \underset{\det = 1}{\begin{bmatrix} 1 & \star & \star \\ 0 & 1 & \star \\ 0 & 0 & 1 \end{bmatrix}} \xrightarrow[\longleftarrow]{\text{row replacements}} \underset{\det = 1}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}$$

Since $\det(I_n) = 1$ and we scaled by the reciprocals of the diagonal entries, this implies $\det(A)$ is the product of the diagonal entries.
The same argument works for lower triangular matrices, except that the row replacements go down instead of up.

Example 2.5.1.5

Compute the following determinants:

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} \qquad \begin{vmatrix} -20 & 0 & 0 \\ \pi & 0 & 0 \\ 100 & 3 & -7 \end{vmatrix} \qquad \begin{vmatrix} 17 & -3 & 4 \\ 0 & 0 & 0 \\ 11 & 1 & e \end{vmatrix}$$

Solution

The first matrix is upper-triangular:

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{vmatrix} = 1 \cdot 4 \cdot 6 = \boxed{24}$$

The second matrix is lower-triangular:

$$\begin{vmatrix} -20 & 0 & 0 \\ \pi & 0 & 0 \\ 100 & 3 & -7 \end{vmatrix} = -20 \cdot 0 \cdot -7 = \boxed{0}$$

The third matrix has a zero row:

$$\begin{vmatrix} 17 & -3 & 4 \\ 0 & 0 & 0 \\ 11 & 1 & e \end{vmatrix} = \boxed{0}$$

A matrix can always be transformed into row echelon form by a series of row operations, and a matrix in row echelon form is upper-triangular. Therefore, we have completely justified the recipe for computing Determinants by Row Reducing.

The determinant is characterized by its defining properties, Definition **2.5.1.1**, since we can compute the determinant of any matrix using row reduction, as in the above Recipe: Computing Determinants by Row Reducing. However, we have not yet proved the existence of a function satisfying the defining properties! Row reducing will compute the determinant *if it exists*, but we cannot use row reduction to prove existence, because we do not yet know that you compute the same number by row reducing in two different ways.

Theorem 2.5.1.1: Existence of the Determinant

There exists one and only one function from the set of square matrices to the real numbers, that satisfies the four defining properties of the determinant in Definition **2.5.1.1**.

We will prove this theorem in the next section by exhibiting a recursive formula for the determinant. Again, the real content of the existence theorem is:

Note 2.5.1.1

No matter which row operations you do to row reduce, you will always compute the same value for the determinant.

2.5.1: Definition of Determinant is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.1: Determinants- Definition** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.5.2: Properties of the Determinant

In this subsection, we will discuss a number of the amazing properties enjoyed by the determinant: the invertibility property, the transpose property, and the multiplicative property.

Proposition 2.5.2.2: Invertibility Property

A square matrix is invertible if and only if $\det(A) \neq 0$.

Proof

If A is invertible, then it has a pivot in every row and column by the Invertible Matrix Theorem, so its reduced row echelon form is the identity matrix. Since row operations do not change whether the determinant is zero, and since $\det(I) = 1$, this implies $\det(A) \neq 0$. Conversely, if A is not invertible, then it is row equivalent to a matrix with a zero row. Again, row operations do not change whether the determinant is nonzero, so in this case $\det(A) = 0$.

By the invertibility property, a matrix that does not satisfy any of the properties of the Invertible Matrix Theorem has zero determinant.

Corollary 2.5.2.1

Let A be a square matrix. If the rows or columns of A are linearly dependent, then $\det(A) = 0$.

Proof

If the columns of A are linearly dependent, then A is not invertible by condition 4 of the Invertible Matrix Theorem. Suppose now that the rows of A are linearly dependent. If $\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_n$ are the rows of A, then one of the rows is in the span of the others, so we have an equation like

$$\vec{r}_2 = \alpha\vec{r}_1 + \beta\vec{r}_3 + \gamma\vec{r}_4$$

If we perform the following row operations on A:

$$R_2 = R_2 - \alpha R_1; \quad R_2 = R_2 - \beta R_3; \quad R_2 = R_2 - \gamma R_4$$

then the second row of the resulting matrix is zero. Hence A is not invertible in this case either.

In particular, if two rows/columns of A are multiples of each other, then $\det(A) = 0$. We also recover the fact that a matrix with a row or column of zeros has determinant zero.

Illustration 2.5.2.3

The following matrices all have zero determinant:

$$\begin{bmatrix} 0 & 2 & -1 \\ 0 & 5 & 10 \\ 0 & -7 & 3 \end{bmatrix}, \quad \begin{bmatrix} 5 & -15 & 11 \\ 3 & -9 & 2 \\ 2 & -6 & 16 \end{bmatrix}, \quad \begin{bmatrix} 3 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 \\ 4 & 2 & 5 & 12 \\ -1 & 3 & 4 & 8 \end{bmatrix}, \quad \begin{bmatrix} \pi & e & 11 \\ 3\pi & 3e & 33 \\ 12 & -7 & 2 \end{bmatrix}$$

The proof of the transpose and multiplicative properties below makes use of a special class of invertible matrices called elementary matrices.

Definition 2.5.2.3: Elementary Matrices

An ***elementary matrix*** is a matrix obtained by doing one row operation to the identity matrix. There are three kinds of elementary matrices: those arising from row replacement, row scaling, and row swaps:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_2=R_2-2R_1} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1=3R_1} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1\leftrightarrow R_2} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The important property of elementary matrices is the following claim.

Claim: If E is the elementary matrix for a row operation, then EA is the matrix obtained by performing the same row operation on A.

In other words, left-multiplication by an elementary matrix applies a row operation. For example,

$$\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21}-2a_{11} & a_{22}-2a_{12} & a_{23}-2a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 3a_{11} & 3a_{12} & 3a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The proof of the Claim is by direct calculation; we leave it to the reader to generalize the above equalities to $n \times n$ matrices.

As a consequence of the definition of elementary matrices and the four defining properties of a determinant, we have the following observation. Let C be any square matrix.

1. If E is the elementary matrix for a row replacement, then $\det(EC) = \det(C)$. In other words, *left-multiplication by E does not change the determinant.*
2. If E is the elementary matrix for a row scale by a factor of c, then $\det(EC) = c\det(C)$. In other words, *left-multiplication by E scales the determinant by a factor of c.*
3. If E is the elementary matrix for a row swap, then $\det(EC) = -\det(C)$. In other words, *left-multiplication by E negates the determinant.*

Proposition 2.5.2.3: Multiplicativity Property

If A and B are $n \times n$ matrices, then

$$\det(AB) = \det(A)\det(B)$$

Proof

Case 1. If we assume that A is invertible, then after finitely many row operations, it must become I. Let each of these row operations be represented by some elementary matrix E_i for $i = 1, 2, \ldots, m$. In other words:

$$E_m E_{m-1} \cdots E_2 E_1 A = I$$

Since each elementary matrix is invertible, we get $A = E_1^{-1} E_2^{-1} \cdots E_{m-1}^{-1} E_m^{-1}$

Moreover, the inverse of any elementary matrix is also an elementary matrix, so we can write $A = F_1 \cdots F_m$ for some elementary matrices $F_1, F_2, \ldots, F_m$. This finally gives us what we need:

$$\begin{aligned} \det(AB) &= \det(F_1 F_2 \cdots F_m B) = \det(F_1)\det(F_2 \cdots F_m B) \\ &\vdots \\ &= \det(F_1)\det(F_2)\cdots\det(F_m)\det(B) \\ &\vdots \\ &= \det(F_1 F_2 \cdots F_m)\det(B) = \det(A)\det(B) \end{aligned}$$

This proves that our claim holds for all invertible matrices A.

Case 2. If A and B are both not invertible, then $\det(A) = \det(B) = 0$, so we want to prove that $\det(AB) = 0 \cdot 0 = 0$ must hold. This boils down to proving that AB is not invertible.

If B is also not invertible, then it also has a non-trivial null space, so there exists some nonzero vector $\vec{y}$ in $Null(B)$

Clearly, this means $(AB)\vec{y} = A(B\vec{y}) = A\vec{0} = \vec{0}$, so AB has nontrivial null space as well. This proves AB is not invertible and our claim holds.

Case 3. If A is not invertible, and B is invertible, then $\det(A) = 0$, so we want to prove that $\det(AB) = 0 \cdot \det(B) = 0$ must hold. Again this boils down to proving that AB is not invertible.

Since A is not invertible, it has a non-trivial null space. Let $\vec{z}$ be a nonzero element of $Null(A)$. Since B is invertible, we know there is a unique solution to $B\vec{y} = \vec{z}$, and moreover:

$(AB)\vec{y} = A(B\vec{y}) = A\vec{z} = \vec{0}$, so AB has nontrivial null space as well. This proves AB is not invertible and our claim holds.

No matter what combination of invertible or non-invertible matrices you have, we can safely say that $\det(AB) = \det(A)\det(B)$

Recall that taking a power of a square matrix A means taking products of A with itself:

$$A^2 = AA \qquad A^3 = AAA \qquad \text{etc.}$$

If A is invertible, then we define

$$A^{-2} = A^{-1}A^{-1} \qquad A^{-3} = A^{-1}A^{-1}A^{-1} \qquad \text{etc.}$$

For completeness, we set $A^0 = I$ if $A \neq 0$.

Corollary 2.5.2.2

If A is a square matrix, then

$$\det(A^n) = \det(A)^n$$

for all $n \geq 1$. If A is invertible, then the equation holds for all $n \leq 0$ as well; in particular,

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Proof

Using the multiplicativity property, Proposition **2.5.2.3**, we compute

$$\det(A^2) = \det(AA) = \det(A)\det(A) = \det(A)^2$$

and

$$\det(A^3) = \det(AAA) = \det(A)\det(AA) = \det(A)\det(A)\det(A) = \det(A)^3$$

the pattern is clear.

We have

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$$

by the multiplicativity property, Proposition **2.5.2.3** and the fourth defining property, Definition **2.5.2.1**, which shows that $\det(A^{-1}) = \det(A)^{-1}$. Thus

$$\det(A^{-2}) = \det(A^{-1}A^{-1}) = \det(A^{-1})\det(A^{-1}) = \det(A^{-1})^2 = \det(A)^{-2}$$

and so on.

Example 2.5.2.6

Compute $\det(A^{100})$, where

$$A = \begin{bmatrix} 4 & 1 \\ 2 & 1 \end{bmatrix}$$

Solution

We have $\det(A) = 4 - 2 = 2$, so

$$\det(A^{100}) = \det(A)^{100} = 2^{100}$$

Nowhere did we have to compute the 100^{th} power of A! (We will learn an efficient way to do that eventually)

Here is another application of the multiplicative property.

Corollary 2.5.2.3

Let $A_1, A_2, \ldots, A_k$ be $n \times n$ matrices. Then the product $A_1A_2\cdots A_k$ is invertible if and only if each A_i is invertible.

Proof

The determinant of the product is the product of the determinants by the multiplicativity property, Proposition **2.5.2.3**:

$$\det(A_1A_2\cdots A_k) = \det(A_1)\det(A_2)\cdots\det(A_k)$$

By the invertibility property, Proposition **2.5.2.2**, this is nonzero if and only if $A_1A_2\cdots A_k$ is invertible. On the other hand, $\det(A_1)\det(A_2)\cdots\det(A_k)$ is nonzero if and only if each $\det(A_i) \neq 0$, which means each A_i is invertible.

Summary: Magical Properties of the Determinant

1. There is one and only one function $\det: \{n \times n \text{ matrices}\} \to \mathbb{R}$ satisfying the four defining properties of determinant.
2. The determinant of an upper-triangular or lower-triangular matrix is the product of the diagonal entries.
3. A square matrix is invertible if and only if $\det(A) \neq 0$; in this case,

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

4. If A and B are $n \times n$ matrices, then

$$\det(AB) = \det(A)\det(B)$$

2.5.2: Properties of the Determinant is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.1: Determinants- Definition** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.5.3: Practice Set

Exercise 2.5.3.1a

Find the determinant of $A = \begin{bmatrix} 4 & 6 \\ \frac{1}{3} & \frac{3}{2} \end{bmatrix}$ and determine if it is invertible.

Answer

$$A = \begin{bmatrix} 4 & 6 \\ \frac{1}{3} & \frac{3}{2} \end{bmatrix} \xrightarrow{R_2 = R_2 - \frac{1}{12}R_1} \begin{bmatrix} 4 & 6 \\ 0 & 1 \end{bmatrix} = A'$$

Since $\det(A') = 4 \cdot 1 = 4$ and the determinant is unchanged by this operation, we know $\det(A) = \det(A')$

Therefore $\det(A) = 4$, and it is invertible.

Exercise 2.5.3.1b

Find the determinant of $B = \begin{bmatrix} \pi & 6\sqrt{3} \\ -\sqrt{3} & \frac{1}{\pi} \end{bmatrix}$ and determine if it is invertible.

Answer

$$B = \begin{bmatrix} \pi & 6\sqrt{3} \\ -\sqrt{3} & \frac{1}{\pi} \end{bmatrix} \xrightarrow{R_2 = R_2 + \frac{\sqrt{3}}{\pi}R_1} \begin{bmatrix} \pi & 6\sqrt{3} \\ 0 & \frac{19}{\pi} \end{bmatrix} = B'$$

Since $\det(B') = \pi \cdot \frac{19}{\pi} = 19$ and the determinant is unchanged by this operation, we know $\det(B) = \det(B')$

Therefore $\det(B) = 19$, and it is invertible.

Exercise 2.5.3.1c

Find the determinant of $C = \begin{bmatrix} 1 & 2 & 9 \\ 0 & 4 & -1 \\ 0 & 0 & 5 \end{bmatrix}$ and determine if it is invertible.

Answer

$C = \begin{bmatrix} 1 & 2 & 9 \\ 0 & 4 & -1 \\ 0 & 0 & 5 \end{bmatrix}$ is already in echelon form

Therefore $\det(C) = 1 \cdot 4 \cdot 5 = 20$ and it is invertible.

Exercise 2.5.3.1d

Find the determinant of $D = \begin{bmatrix} 1 & -3 & 2 \\ 3 & -4 & 11 \\ 1 & 3 & 1 \end{bmatrix}$ and determine if it is invertible.

Answer

$$D = \begin{bmatrix} 1 & -3 & 2 \\ 3 & -4 & 11 \\ 1 & 3 & 1 \end{bmatrix} \xrightarrow{R_2 = R_2 - 3R_1, R_3 = R_3 - R_1} \begin{bmatrix} 1 & -3 & 2 \\ 0 & 5 & 5 \\ 0 & 6 & -1 \end{bmatrix} \xrightarrow{R_2 = \frac{1}{5}R_2} \begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & 1 \\ 0 & 6 & -1 \end{bmatrix} \xrightarrow{R_3 = R_3 - 6R_2} \begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & -7 \end{bmatrix} = D'$$

Since $\det(D') = 1 \cdot 1 \cdot (-7) = -7$ and the determinant is only changed by $R_2 = \frac{1}{5}R_2$, we know $\frac{1}{5}\det(D) = \det(D')$

Therefore $\det(D) = -35$, and it is invertible.

? Exercise 2.5.3.1e

Find the determinant of $E = \begin{bmatrix} 8 & 4 & -12 \\ 2 & 1 & -3 \\ 6 & 4 & 11 \end{bmatrix}$ and determine if it is invertible.

Answer

$$E = \begin{bmatrix} 8 & 4 & -12 \\ 2 & 1 & -3 \\ 6 & 4 & 11 \end{bmatrix} \xrightarrow{R_1 = \frac{1}{4}R_1} \begin{bmatrix} 2 & 1 & -3 \\ 2 & 1 & -3 \\ 6 & 4 & 11 \end{bmatrix} \xrightarrow{R_2 = R_2 - R_1, R_3 = R_3 - 3R_1} \begin{bmatrix} 2 & 1 & -3 \\ 0 & 0 & 0 \\ 0 & 1 & 20 \end{bmatrix}$$

Since E' has a row of zeros, we know $\det(E') = 0$, and since the determinant is only changed by $R_1 = \frac{1}{4}R_1$, we have $\frac{1}{4}\det(E) = \det(E')$

Therefore $\det(E) = 0$, and it is not invertible.

? Exercise 2.5.3.1f

Find the determinant of $F = \begin{bmatrix} 6 & 6 & -3 \\ 1 & 2 & -1 \\ 7 & 8 & -4 \end{bmatrix}$ and determine if it is invertible.

Answer

$$F = \begin{bmatrix} 6 & 6 & -3 \\ 1 & 2 & -1 \\ 7 & 8 & -4 \end{bmatrix} \xrightarrow{R_1 = R_1 - 6R_2, R_3 = R_3 - 7R_2} \begin{bmatrix} 0 & -6 & 3 \\ 1 & 2 & -1 \\ 0 & -6 & 3 \end{bmatrix} \xrightarrow{R_3 = R_3 - R_1} \begin{bmatrix} 0 & -6 & 3 \\ 1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} = F'$$

Since F' has a row of zeros, we know $\det(F') = 0$, and since the determinant is unchanged by this operation, we know $\det(F) = \det(F')$

Therefore $\det(F) = 0$, and it is not invertible.

? Exercise 2.5.3.1g

Find the determinant of $G = \begin{bmatrix} 1 & 2 & 0 & 5 \\ 7 & 11 & 0 & -2 \\ 2 & 5 & 0 & -4 \\ 1 & 1 & 0 & 1 \end{bmatrix}$ and determine if it is invertible.

Answer

The matrix has a column of all zeros, so the matrix cannot be invertible (linearly dependent columns).

Therefore $\det(G) = 0$ must hold.

? Exercise 2.5.3.1h

Find the determinant of $H = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 1 & \frac{5}{3} \\ \frac{1}{6} & \frac{1}{3} & 1 & \frac{5}{6} \\ 10 & 21 & 30 & 50 \\ 10 & 21 & 30 & 45 \end{bmatrix}$ and determine if it is invertible.

Answer

$$H = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 1 & \frac{5}{3} \\ \frac{1}{6} & \frac{1}{3} & 1 & \frac{5}{6} \\ 10 & 21 & 30 & 50 \\ 10 & 21 & 30 & 45 \end{bmatrix} \xrightarrow{R_1 = 3R_1, R_2 = 6R_2} \begin{bmatrix} 1 & 2 & 3 & 5 \\ 1 & 2 & 6 & 5 \\ 10 & 21 & 30 & 50 \\ 10 & 21 & 30 & 45 \end{bmatrix} \xrightarrow{R_2 = R_2 - R_1, R_3 = R_3 - 10R_1, R_4 = R_4 - 10R_1} \begin{bmatrix} 1 & 2 & 3 & 5 \\ 0 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -5 \end{bmatrix}$$

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & 3 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 1 & 0 & -5 \end{bmatrix} \xrightarrow{R_4 = R_4 - R_2} \begin{bmatrix} 1 & 2 & 3 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -5 \end{bmatrix} = H'$$

Since $\det(H') = 1 \cdot 1 \cdot 3 \cdot (-5) = -15$ and the determinant is only changed by $R_1 = 3R_1, R_2 = 6R_2, R_2 \leftrightarrow R_3$, we know $3 \cdot 6 \cdot (-1)\det(H) = \det(H')$

Therefore $\det(H) = \frac{-15}{-18} = \frac{5}{6}$, and it is invertible.

Exercise 2.5.3.2

If $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has $\det(M) = -2$, evaluate the determinant of each matrix below.

(a) $A = \begin{bmatrix} 2a & 2b \\ -3c & -3d \end{bmatrix}$

(b) $B = \begin{bmatrix} a - 5c & b - 5d \\ c & d \end{bmatrix}$

(c) $C = \begin{bmatrix} -4a + 7c & -4b + 7d \\ 2a - c & 2b - d \end{bmatrix}$

(d) $D = M^3$

(e) $E = M^{-1}$

(f) $F = 2M$

Answer

(a) $M \xrightarrow{R_1 = 2R_1, R_2 = -3R_2} A$ therefore $\det(A) = 2 \cdot (-3) \cdot \det(M) = 12$

(b) $M \xrightarrow{R_1 = R_1 - 5R_2} B$ therefore $\det(B) = \det(M) = -2$

(c) This one is hard to see the row operations from M to C so we go in reverse:

$$C \xrightarrow{R_1 = R_1 + 7R_2} \begin{bmatrix} 10a & 10b \\ 2a - c & 2b - d \end{bmatrix} \xrightarrow{R_1 = \frac{1}{10}R_1} \begin{bmatrix} a & b \\ 2a - c & 2b - d \end{bmatrix} \xrightarrow{R_2 = R_2 - 2R_1} \begin{bmatrix} a & b \\ -c & -d \end{bmatrix} \xrightarrow{R_2 = -1 \cdot R_2} M$$

Therefore $\frac{1}{10} \cdot (-1)\det(C) = \det(M)$, which means $\det(C) = 20$

(d) $\det(D) = \det(M^3) = [\det(M)]^3 = (-2)^3 = -8$

(e) $\det(E) = \det(M^{-1}) = [\det(M)]^{-1} = (-2)^{-1} = -\frac{1}{2}$

(f) $M \xrightarrow{R_1 = 2R_1, R_2 = 2R_2} F$ therefore $\det(F) = 2 \cdot 2 \cdot \det(M) = -8$

Exercise 2.5.3.3

If $M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ has $\det(M) = 10$, evaluate the determinant of each matrix below.

(a) $A = \begin{bmatrix} d & e & f \\ g & h & i \\ a & b & c \end{bmatrix}$

(b) $B = \begin{bmatrix} 2a & 2b & 2c \\ d - 3g & e - 3h & f - 3i \\ a - g & b - h & c - i \end{bmatrix}$

(c) $C = \begin{bmatrix} a + d + 2g & b + e + 2h & c + f + 2i \\ d - a - g & e - b - h & f - c - i \\ 3a + g & 3b + h & 3c + i \end{bmatrix}$

(d) $D = \begin{bmatrix} 2a-2g & 2b-2h & 2c-2i \\ 4a-d & 4b-e & 4c-f \\ g-a & h-b & i-c \end{bmatrix}$

(e) $E = M^5$

(f) $F = \frac{1}{10}M$

Answer

(a) $M \xrightarrow{R_2 \leftrightarrow R_3, R_1 \leftrightarrow R_3} A$ therefore $\det(A) = (-1)\cdot(-1)\cdot\det(M) = 10$

(b) $M \xrightarrow{R_2 = R_2 - 3R_3, R_3 = R_3 - R_1, R_3 = -1\cdot R_3, R_1 = 2R_1} B$ therefore $\det(B) = (-1)\cdot 2 \cdot \det(M) = -20$

(c) This one is hard to see the row operations from M to C so we go in reverse:

$$C \xrightarrow{R_2 = R_2 + R_3} \begin{bmatrix} a+d+2g & b+e+2h & c+f+2i \\ d+2a & e+2b & f+2c \\ 3a+g & 3b+h & 3c+i \end{bmatrix} \xrightarrow{R_1 = R_1 - R_2 - 2R_3} \begin{bmatrix} -7a & -7b & -7c \\ d+2a & e+2b & f+2c \\ 3a+g & 3b+h & 3c+i \end{bmatrix}$$

$$\xrightarrow{R_1 = -\frac{1}{7}R_1} \begin{bmatrix} a & b & c \\ d+2a & e+2b & f+2c \\ 3a+g & 3b+h & 3c+i \end{bmatrix} \xrightarrow{R_2 = R_2 - 2R_1, R_3 = R_3 - 3R_1} M$$

Therefore $-\frac{1}{7}\det(C) = \det(M)$, which means $\det(C) = -70$

(d) Note that $D \xrightarrow{R_1 = R_1 + 2R_3} \begin{bmatrix} 0 & 0 & 0 \\ 4a-d & 4b-e & 4c-f \\ g-a & h-b & i-c \end{bmatrix}$, so $\det(D) = 0$

(This matrix D cannot be made from M, and M cannot be made from D via row operations.)

(e) $\det(E) = \det(M^5) = [\det(M)]^5 = 10^5 = 100000$

(f) $M \xrightarrow{R_1 = \frac{1}{10}R_1, R_2 = \frac{1}{10}R_2, R_3 = \frac{1}{10}R_3} F$ therefore $\det(F) = \frac{1}{10}\cdot\frac{1}{10}\cdot\frac{1}{10}\cdot\det(M) = \frac{1}{100}$

? Exercise 2.5.3.4

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If $\det(A) = 1$, then the RREF of A is I.

(b) If $\det(A) = 0$, then the columns of A are linearly independent.

(c) If $\det(A) = 4$, then $\det(2A) = 32$.

(d) If $\det(A) = 0$, then $\text{Nullity}(A) > 0$.

(e) If A^5 is not invertible, then A is not invertible.

(f) If A and B are non-zero $n \times n$ matrices, then $\det(A+B) = \det(A) + \det(B)$.

(g) If A and B are $n \times n$ matrices and $\det(A) = 1$, then $\det(AB) = \det(B)$.

(h) If A and B are $n \times n$ matrices and $\det(AB) = 1$, then $B = A^{-1}$.

(i) If A satisfies $A^2 = A$, then $\det(A) = 0$ or $\det(A) = 1$.

(j) If A, B and P are $n \times n$ matrices satisfying $A = PBP^{-1}$, then $\det(A) = \det(B)$.

(k) If A is an elementary matrix, then A is invertible and A^{-1} is also an elementary matrix.

Answer

(a) **Always True**

Since $\det(A) \neq 0$, it is invertible, so it must row reduce to I.

(b) **Always False**

Since $\det(A) = 0$, it is not invertible, so its columns must be linearly dependent.

(c) **Sometimes True Sometimes False**

If A is 3×3 then $\det(2A) = 2^3 \det(A)$ so it is true, but if A is not 3×3 this is false.

(d) **Always True**

Since $\det(A) = 0$, it is not invertible, so $A\vec{x} = \vec{0}$ must have nontrivial solutions.

(e) **Always True**

We know $\det(A^5) = (\det(A))^5 = 0$ so $\det(A) = 0$ must hold.

(f) **Sometimes True Sometimes False**

This is true for the example $A = B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, and false for the example $A = B = I$.

(g) **Always True**

We see $\det(AB) = \det(A) \cdot \det(B) = 1 \cdot \det(B)$.

(h) **Sometimes True Sometimes False**

For the true case, consider the example $A = B = I$. For the false case, consider $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & 1 \end{bmatrix}$.

(i) **Always True**

$\det(A) = \det(A^2) = (\det(A))^2$, so let $x = \det(A)$ and solve $x = x^2$, which has solutions $x = 0, 1$.

(j) **Always True**

$\det(A) = \det(PBP^{-1}) = \det(P)\det(B)\det(P^{-1}) = \det(P)\det(B)\frac{1}{\det(P)} = \det(B)$.

(k) **Always True**

To invert $R_i = R_i + kR_j$, we use $R_i = R_i - kR_j$. To invert $R_i = kR_i$, we use $R_i = \frac{1}{k}R_i$. To invert $R_i \leftrightarrow R_j$, we use $R_i \leftrightarrow R_j$. Thus, multiplying by any elementary matrix A can be undone by a row operation, which is multiplication by some elementary matrix A^{-1}.

2.5.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.6: Cofactor Expansions

Objectives

1. Learn to recognize which methods are best suited to compute the determinant of a given matrix.
2. *Recipes:* the determinant of a $n \times n$ matrix for $n \geq 3$, compute the determinant using cofactor expansions.
3. *Vocabulary words:* ***minor***, ***cofactor***.

In this section, we give a recursive formula for the determinant of a matrix, called a *cofactor expansion*. The formula is recursive in that we will compute the determinant of an $n \times n$ matrix *assuming* we already know how to compute the determinant of an $(n-1) \times (n-1)$ matrix. We finish with a supplementary subsection on Cramer's rule and a cofactor formula for the inverse of a matrix.

This page titled 2.6: Cofactor Expansions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **4.2: Cofactor Expansions** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.6.1: Cofactor Expansion

The following observation is useful for theoretical purposes.

We can think of $\det$ as a function of the rows of a matrix:

$$\det(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n) = \det\left(\begin{bmatrix} -\vec{v}_1- \\ -\vec{v}_2- \\ \vdots \\ -\vec{v}_n- \end{bmatrix}\right)$$

Proposition 2.6.1.1: Multilinearity Property

Let i be a whole number between 1 and n, and fix $n-1$ vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{i-1}, \vec{v}_{i+1}, \dots, \vec{v}_n$ in $\mathbb{R}^n$. Then the transformation $T\colon \mathbb{R}^n \to \mathbb{R}$ defined by

$$T(\vec{x}) = \det(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{i-1}, \vec{x}, \vec{v}_{i+1}, \dots, \vec{v}_n)$$

is *linear*.

Proof

First assume that $i = 1$, so

$$T(\vec{x}) = \det(\vec{x}, \vec{v}_2, \dots, \vec{v}_n)$$

We have to show that T satisfies the defining properties of a linear transformation:

- By the first defining property of determinants, we know scaling any row of a matrix by a number c scales the determinant by a factor of c. This implies that T satisfies the scalar multiple property, i.e., that

$$T(c\vec{x}) = \det(c\vec{x}, \vec{v}_2, \dots, \vec{v}_n) = c\det(\vec{x}, \vec{v}_2, \dots, \vec{v}_n) = cT(\vec{x})$$

- We claim that $T(\vec{v}+\vec{w}) = T(\vec{v}) + T(\vec{w})$.
  Case 1: If $\vec{w}$ is in $\operatorname{Span}\{\vec{v}, \vec{v}_2, \dots, \vec{v}_n\}$, then

$$\vec{w} = c\vec{v} + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$$

for some scalars $c, c_2, \dots, c_n$. Let A be the matrix with rows $\vec{v}+\vec{w}, \vec{v}_2, \dots, \vec{v}_n$, so $T(\vec{v}+\vec{w}) = \det(A)$. By performing the row operations

$$R_1 = R_1 - c_2R_2; \quad R_1 = R_1 - c_3R_3; \quad \dots \quad R_1 = R_1 - c_nR_n$$

the first row of the matrix A becomes

$$\vec{v} + \vec{w} - (c_2\vec{v}_2 + \cdots + c_n\vec{v}_n) = \vec{v} + c\vec{v} = (1+c)\vec{v}$$

Therefore,

$$\begin{aligned} T(\vec{v}+\vec{w}) = \det(A) &= \det((1+c)\vec{v}, \vec{v}_2, \dots, \vec{v}_n) \\ &= (1+c)\det(\vec{v}, \vec{v}_2, \dots, \vec{v}_n) \\ &= T(\vec{v}) + cT(\vec{v}) = T(\vec{v}) + T(c\vec{v}) \end{aligned}$$

Doing the opposite row operations

$$R_1 = R_1 + c_2R_2; \quad R_1 = R_1 + c_3R_3; \quad \dots \quad R_1 = R_1 + c_nR_n$$

to the matrix with rows $c\vec{v}, \vec{v}_2, \dots, \vec{v}_n$ shows that

$$\begin{aligned} T(c\vec{v}) &= \det(c\vec{v}, \vec{v}_2, \dots, \vec{v}_n) \\ &= \det(c\vec{v} + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n, \vec{v}_2, \dots, \vec{v}_n) \\ &= \det(\vec{w}, \vec{v}_2, \dots, \vec{v}_n) = T(\vec{w}) \end{aligned}$$

which finishes the proof of the first property in this case.

Case 2: If $\vec{w}$ is not in $\text{Span}\{\vec{v}, \vec{v}_2, \ldots, \vec{v}_n\}$. This implies that $\{\vec{v}, \vec{v}_2, \ldots, \vec{v}_n\}$ is linearly *dependent* (otherwise it would form a basis for $\mathbb{R}^n$), so $T(\vec{v}) = \vec{0}$. If $\vec{v}$ is not in $\text{Span}\{\vec{v}_2, \ldots, \vec{v}_n\}$, then $\{\vec{v}_2, \ldots, \vec{v}_n\}$ is linearly dependent by the increasing span criterion, so $T(\vec{x}) = \vec{0}$ for all $\vec{x}$, as the matrix with rows $\vec{x}, \vec{v}_2, \ldots, \vec{v}_n$ is not invertible. Clearly then, we have $T(\vec{v} + \vec{w}) = \vec{0} = T(\vec{v}) + T(\vec{w})$. which proves our first property in this case.

For $i \neq 1$, we note that

$$\begin{aligned} T(\vec{x}) &= \det(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{i-1}, \vec{x}, \vec{v}_{i+1}, \ldots, \vec{v}_n) \\ &= -\det(\vec{x}, \vec{v}_2, \ldots, \vec{v}_{i-1}, \vec{v}_1, \vec{v}_{i+1}, \ldots, \vec{v}_n) \end{aligned}$$

By the previously handled case, we know that $-T$ is a linear transformation, and therefore (since it is a scalar multiple of $-T$), T must be linear.

For example, we have

$$\begin{vmatrix} | & & | \\ \vec{v}_1 & a\vec{v} + b\vec{w} & \vec{v}_3 \\ | & & | \end{vmatrix} = a \begin{vmatrix} | & | & | \\ \vec{v}_1 & \vec{v} & \vec{v}_3 \\ | & | & | \end{vmatrix} + b \begin{vmatrix} | & | & | \\ \vec{v}_1 & \vec{w} & \vec{v}_3 \\ | & | & | \end{vmatrix}$$

Now is a good time to define the transpose of a matrix.

Definition: Transpose

The ***transpose*** of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose rows are the columns of A. In other words, the ij entry of A^T is a_{ji}.

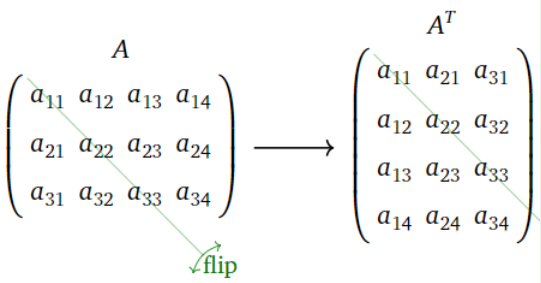


Figure **2.6.1.3**

Like inversion, transposition reverses the order of matrix multiplication.

Fact 2.6.1.1

Let A be an $m \times n$ matrix, and let B be an $n \times p$ matrix. Then

$$(AB)^T = B^T A^T$$

Proof

First suppose that A is a row vector and B is a column vector, i.e., $m = p = 1$. Then

$$AB = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

$$= \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = B^T A^T$$

Now we use the row-column rule for matrix multiplication. Let $\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_m$ be the rows of A, and let $\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_p$ be the columns of B, so

$$AB = \begin{bmatrix} \text{—}\vec{r}_1\text{—} \\ \text{—}\vec{r}_2\text{—} \\ \vdots \\ \text{—}\vec{r}_m\text{—} \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{c}_1 & \vec{c}_2 & \cdots & \vec{c}_p \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \vec{r}_1 \cdot \vec{c}_1 & \vec{r}_1 \cdot \vec{c}_2 & \cdots & \vec{r}_1 \cdot \vec{c}_p \\ \vec{r}_2 \cdot \vec{c}_1 & \vec{r}_2 \cdot \vec{c}_2 & \cdots & \vec{r}_2 \cdot \vec{c}_p \\ \vdots & \vdots & & \vdots \\ \vec{r}_m \cdot \vec{c}_1 & \vec{r}_m \cdot \vec{c}_2 & \cdots & \vec{r}_m \cdot \vec{c}_p \end{bmatrix}$$

By the case we handled above, we have $\vec{r}_i \cdot \vec{c}_j = \vec{c}_j^T \cdot \vec{r}_i^T$. Then

$$\begin{aligned} (AB)^T &= \begin{bmatrix} \vec{r}_1 \cdot \vec{c}_1 & \vec{r}_2 \cdot \vec{c}_1 & \cdots & \vec{r}_m \cdot \vec{c}_1 \\ \vec{r}_1 \cdot \vec{c}_2 & \vec{r}_2 \cdot \vec{c}_2 & \cdots & \vec{r}_m \cdot \vec{c}_2 \\ \vdots & \vdots & & \vdots \\ \vec{r}_1 \cdot \vec{c}_p & \vec{r}_2 \cdot \vec{c}_p & \cdots & \vec{r}_m \cdot \vec{c}_p \end{bmatrix} \\ &= \begin{bmatrix} \vec{c}_1^T \cdot \vec{r}_1^T & \vec{c}_1^T \cdot \vec{r}_2^T & \cdots & \vec{c}_1^T \cdot \vec{r}_m^T \\ \vec{c}_2^T \cdot \vec{r}_1^T & \vec{c}_2^T \cdot \vec{r}_2^T & \cdots & \vec{c}_2^T \cdot \vec{r}_m^T \\ \vdots & \vdots & & \vdots \\ \vec{c}_p^T \cdot \vec{r}_1^T & \vec{c}_p^T \cdot \vec{r}_2^T & \cdots & \vec{c}_p^T \cdot \vec{r}_m^T \end{bmatrix} \\ &= \begin{bmatrix} \text{—}\vec{c}_1^T\text{—} \\ \text{—}\vec{c}_2^T\text{—} \\ \vdots \\ \text{—}\vec{c}_p^T\text{—} \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{r}_1^T & \vec{r}_2^T & \cdots & \vec{r}_m^T \\ | & | & & | \end{bmatrix} = B^T A^T \end{aligned}$$

Proposition 2.6.1.2: Transpose Property

For any square matrix A, we have

$$\det(A) = \det(A^T)$$

Proof

Define the function $d(A) = \det(A^T)$. We will show that d satisfies the four defining properties of the determinant.

1. Let C' be the matrix obtained by doing a row replacement on the general matrix C, and let E be the elementary matrix for this row replacement, so $C' = EC$. The elementary matrix for a row replacement is either upper-triangular or lower-triangular, with ones on the diagonal:

$$R_1 = R_1 + 3R_3 : \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad R_3 = R_3 + 3R_1 : \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$$

   It follows that E^T is also either upper-triangular or lower-triangular, with ones on the diagonal, so $\det(E^T) = 1$. By the multiplicative property,

$$\begin{aligned} d(C') &= \det((C')^T) = \det((EC)^T) = \det(C^T E^T) \\ &= \det(C^T)\det(E^T) = \det(C^T) = d(C) \end{aligned}$$

2. Let C' be the matrix obtained by scaling a row of C by a factor of c, and let E be the elementary matrix for this row replacement, so $C' = EC$. Then E is a diagonal matrix:

$$R_2 = cR_2 : \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

   Thus $\det(E^T) = c$. By the multiplicative property,

$$d(C') = \det((C')^T) = \det((EC)^T) = \det(C^T E^T)$$
$$= \det(C^T)\det(E^T) = c\det(C^T) = c \cdot d(C)$$

3. Let C' be the matrix obtained by swapping two rows of C, and let E be the elementary matrix for this row replacement, so $C' = EC$. The E is equal to its own transpose:

$$R_1 \leftrightarrow R_2 : \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}^T$$

Since E (hence E^T) is obtained by performing one row swap on the identity matrix, we have $\det(E^T) = -1$. By the multiplicative property,

$$d(C') = \det((C')^T) = \det((EC)^T) = \det(C^T E^T)$$
$$= \det(C^T)\det(E^T) = -\det(C^T) = -d(C)$$

4. Since $I^T = I$, we have

$$d(I) = \det(I^T) = \det(I) = 1$$

Since d satisfies the four defining properties of the determinant, *it is defined exactly the same as, and must therefore be equal to, the determinant*. In other words, for all matrices A, we have

$$\det(A) = d(A) = \det(A^T)$$

The transpose property is very useful. For concreteness, we note that $\det(A) = \det(A^T)$ means, for instance, that

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{vmatrix}$$

This implies that the determinant has the curious feature that it also behaves well with respect to *column* operations. Indeed, a column operation on A is the same as a row operation on A^T, and $\det(A) = \det(A^T)$.

Corollary 2.6.1.1

The determinant satisfies the following properties with respect to column operations:

1. Doing a column replacement on A does not change $\det(A)$.
2. Scaling a column of A by a scalar c multiplies the determinant by c.
3. Swapping two columns of a matrix multiplies the determinant by -1.

The previous corollary makes it easier to compute the determinant: one is allowed to do row *and* column operations when simplifying the matrix. (Of course, one still has to keep track of how the row and column operations change the determinant.)

Example 2.6.1.1

Compute $\begin{vmatrix} 2 & 7 & 4 \\ 3 & 1 & 3 \\ 4 & 0 & 1 \end{vmatrix}$

Solution

It takes fewer column operations than row operations to make this matrix upper-triangular:

$$\begin{bmatrix} 2 & 7 & 4 \\ 3 & 1 & 3 \\ 4 & 0 & 1 \end{bmatrix} \xrightarrow{C_1 = C_1 - 4C_3} \begin{bmatrix} -14 & 7 & 4 \\ -9 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{C_1 = C_1 + 9C_2} \begin{bmatrix} 49 & 7 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

We performed two column replacements (no scaling or switching), which does not change the determinant; therefore,

$$\begin{vmatrix} 2 & 7 & 4 \\ 3 & 1 & 3 \\ 4 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 49 & 7 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{vmatrix} = \boxed{49}$$

By the transpose property, the determinant is also multilinear in the *columns* of a matrix:

$$\begin{vmatrix} | & | & | \\ \vec{v}_1 & a\vec{v}+b\vec{w} & \vec{v}_3 \\ | & | & | \end{vmatrix} = a \begin{vmatrix} | & | & | \\ \vec{v}_1 & \vec{v} & \vec{v}_3 \\ | & | & | \end{vmatrix} + b \begin{vmatrix} | & | & | \\ \vec{v}_1 & \vec{w} & \vec{v}_3 \\ | & | & | \end{vmatrix}$$

Illustration 2.6.1.1

We have

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

Therefore,

$$\begin{vmatrix} 2 & 7 & 4 \\ 3 & 1 & 3 \\ 4 & 0 & 1 \end{vmatrix} = 2\begin{vmatrix} 1 & 7 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{vmatrix} + 3\begin{vmatrix} 0 & 7 & 4 \\ 1 & 1 & 3 \\ 0 & 0 & 1 \end{vmatrix} + 4\begin{vmatrix} 0 & 7 & 4 \\ 0 & 1 & 3 \\ 1 & 0 & 1 \end{vmatrix}$$

This is the basic idea behind the upcoming formula using cofactor expansions.

A recursive formula must have a starting point. For cofactor expansions, the starting point is the case of 1×1 matrices. The definition of determinant directly implies that $I = [1]$ has determinant 1, so

$$\det([\ a\ ]) = a$$

To describe cofactor expansions, we need to introduce some notation.

Definition: Minor and Cofactor

Let A be an $n \times n$ matrix.

1. The (i,j) ***minor***, denoted A_{ij}, is the $(n-1) \times (n-1)$ matrix obtained from A by deleting the ith row and the jth column.
2. The (i,j) ***cofactor*** C_{ij} is defined in terms of the minor by

$$C_{ij} = (-1)^{i+j} \det(A_{ij})$$

Note that the signs of the cofactors follow a "checkerboard pattern." Namely, $(-1)^{i+j}$ is pictured in this matrix:

$$\begin{bmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{bmatrix}$$

Illustration 2.6.1.2

For

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

The minor A_{23} and cofactor C_{23} are

$$A_{23} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} \qquad C_{23} = (-1)^{2+3} \det \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} = (-1)(-6) = 6$$

Figure 2.6.1.1

The cofactors C_{ij} of an $n \times n$ matrix are determinants of $(n-1) \times (n-1)$ submatrices. Hence the following theorem is in fact a recursive procedure for computing the determinant.

Theorem 2.6.1.1: Cofactor Row Expansion

Let A be an $n \times n$ matrix with entries a_{ij}.

For any $i = 1, 2, \ldots, n$, we have

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij} = a_{i1} C_{i1} + a_{i2} C_{i2} + \cdots + a_{in} C_{in}$$

This is called ***cofactor expansion along the ith row***.

Proof

First we will prove that cofactor expansion along the i^{th} row computes the determinant for each i. Define a function $d_i : \{n \times n \text{ matrices}\} \to \mathbb{R}$ by

$$d_i(A) = \sum_{j=1}^{n} (-1)^{j+i} a_{ij} \det(A_{ij})$$

We want to show that $d_i(A) = \det(A)$.

This proof is done by induction. We start with a base case to show that the 2×2 matrix expanded along either row matches the determinant we have previously defined. This is fairly straightforward:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \implies |A| = ad - bc$$

If we expand along the first row: $d_1(A) = (-1)^{1+1} a \det([d]) + (-1)^{1+2} b \det([c]) = ad - bc$

If we expand along the second row: $d_2(A) = (-1)^{2+1} c \det([b]) + (-1)^{2+2} d \det([a]) = ad - bc$

Now, we can assume that we have proven that $d(A) = \det(A)$ for all $(n-1) \times (n-1)$ matrices, and show that this assumption implies the same is true for the $n \times n$ case.

This requires four smaller proofs, one for each of the four defining properties of determinant.

1. Doing a row replacement on A does not change $d(A)$

Case 1: r times the k^{th} row is added to the ℓ^{th} row to make the matrix B for $\ell \neq i$

Then each minor B_{ij} is the minor A_{ij} with one row replacement operation done to it. Since determinants are not affected by row replacement, this means the corresponding cofactors and matrix entries they multiply by are exactly equal:

$$d_i(B) = \sum_{j=1}^{n} (-1)^{j+i} a_{ij} \det(A_{ij}) = d_i(A)$$

Case 2: r times the k^{th} row is added to the i^{th} row to make the matrix B.

Then each minor B_{ij} is exactly equal to the matching minor A_{ij}, which means the corresponding cofactors are exactly equal:

$$d_i(B) = \sum_{j=1}^{n} (-1)^{j+i}(a_{ij} + ra_{kj}) \det(A_{ij}) = d_i(A) + rd_i(A')$$

Here we use A' to denote the matrix A with row i replaced entirely by row k. Once we show that $d_i(A') = 0$ we will have finished the proof of this case. Assume rows k and i are matching in A'

$$d_i(A') = \sum_{j=1}^{n} (-1)^{j+i} a_{kj} \det(A_{ij})$$

Now we expand the $n-1 \times n-1$ matrix $\det(A_{ij})$ by using $d_k(A_{ij})$. Consider the minor of A_{ij} formed when row k and column ℓ from the original matrix A are also removed. This matrix will appear exactly twice in our double cofactor expansion. Once with coefficient $(-1)^{k+i+j+\ell} a_{kj} a_{k\ell}$ and again with coefficient $(-1)^{k+i+j+\ell-1} a_{kj} a_{k\ell}$.

The reason the powers of (-1) do not match is because one occurs when $\ell < j$, so when column number j is removed from A then column ℓ in A_{ij} is exactly column number ℓ in the original matrix A. But when $\ell > j$, removing column number j from A changes the number of the ℓ^{th} column in A to the $\ell - 1^{th}$ column in A_{ij}. Thus by symmetry, the same cofactors will each occur twice with opposite coefficients and will always add to zero. The end result clearly becomes $d_k(A_{ij}) = 0$.

This proves that $d_i(A') = 0$ for our $n \times n$ matrix, and finishes the proof of case 2 for this property.

2. Scaling a row of A by a scalar c multiplies the determinant by c

Case 1: the i^{th} row is scaled by c to make the matrix B. This directly leads to:

$$d_i(B) = \sum_{j=1}^{n} (-1)^{j+i} c a_{ij} \det(A_{ij}) = c d_i(A)$$

Case 2: the k^{th} row is scaled by c to make the matrix B for $k \neq i$

In this case, we know that each minor B_{ij} is the minor A_{ij} with a row scaled by c. This means each cofactor of B is exactly c times the corresponding cofactor of A by the properties of determinants. This gives us

$$d_i(B) = \sum_{j=1}^{n} (-1)^{j+i} a_{ij} \cdot c \det(A_{ij}) = c d_i(A)$$

3. Swapping two rows of a matrix multiplies the determinant by -1

Assume we swap two rows in A, and call the resulting matrix B.

Note that the row operation $R_i \leftrightarrow R_k$ can be replicated by using four different row operations in order:

$$R_i = R_i - R_k; R_k = R_k + R_i; R_i = R_i - R_k, R_i = -1 \cdot R_i$$

Since we have proven properties 1 and 2 above, we know this means that $d(B) = -d(A)$ must hold as desired.

4. The determinant of the identity matrix I is equal to 1.

This is the easiest to check. The i^{th} row of the $n \times n$ matrix I is $\vec{e}_i$, which means:

$$d_i(I) = 0 \det(I_{i1}) + \cdots + 1 \det(I_{ii}) + \cdots + 0 \det(I_{in}) = \det(I_{ii}) = \det(I') = 1$$

Here we denote I' as the $(n-1) \times (n-1)$ identity matrix, which we know has determinant 1.

Finally, we have proven that $d(A) = \det(A)$ for all $n \times n$ matrices.

Corollary 2.6.1.2: Cofactor Column Expansion

Let A be an $n \times n$ matrix with entries a_{ij}.

For any $j = 1, 2, \ldots, n$, we have

$$\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

This is called ***cofactor expansion along the jth column***.

Proof

This is simply a combination of two facts: the columns of A are just the rows of A^T, and determinants are unaffected by transposes.

Note that these cofactor expansions actually give $2n$ different formulas for the determinant: one for each row and one for each column. For instance, the formula for cofactor expansion along the first column is

$$\begin{aligned}\det(A) = \sum_{i=1}^{n} a_{i1} C_{i1} &= a_{11}C_{11} + a_{21}C_{21} + \cdots + a_{n1}C_{n1} \\ &= a_{11}\det(A_{11}) - a_{21}\det(A_{21}) + a_{31}\det(A_{31}) - \cdots \pm a_{n1}\det(A_{n1})\end{aligned}$$

Remember, the determinant of a matrix is just a number, defined by the four defining properties, so to be clear:

Note 2.6.1.1

You obtain the same number by expanding cofactors along ***any*** row or column.

Example 2.6.1.2

Find the determinant of

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

Solution

We make the somewhat arbitrary choice to expand along the first row. The minors and cofactors are

$$A_{11} = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix} \qquad C_{11} = +\det\begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix} = 4$$

$$A_{12} = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} \qquad C_{12} = -\det\begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} = 1$$

$$A_{13} = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -2 & 2 \end{pmatrix} \qquad C_{13} = +\det\begin{pmatrix} -1 & 2 \\ -2 & 2 \end{pmatrix} = 2.$$

Figure 2.6.1.2

Thus,

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = (2)(4) + (1)(1) + (3)(2) = \boxed{15}$$

Illustration 2.6.1.3: The Determinant of a 2×2 Matrix

Let us compute (again) the determinant of a general 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The minors are

$$A_{11} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (d) \qquad A_{12} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (c)$$
$$A_{21} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (b) \qquad A_{22} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = (a).$$

Figure 2.6.1.3

The minors are all 1×1 matrices. As we have seen that the determinant of a 1×1 matrix is just the number inside of it, the cofactors are therefore

$$C_{11} = +\det(A_{11}) = d \qquad C_{12} = -\det(A_{12}) = -c$$
$$C_{21} = -\det(A_{21}) = -b \qquad C_{22} = +\det(A_{22}) = a$$

Expanding cofactors along the first column, we find that

$$\det(A) = aC_{11} + cC_{21} = ad - bc$$

which agrees with our previous findings.

Example 2.6.1.3: The Determinant of a 3×3 Matrix

We can also use cofactor expansions to find a formula for the determinant of a 3×3 matrix. Let us compute the determinant of

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

by expanding along the first row. The minors and cofactors are:

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \qquad C_{11} = +\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$$
$$A_{12} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \qquad C_{12} = -\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}$$
$$A_{13} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \qquad C_{13} = +\det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Figure 2.6.1.4

The determinant is:

$$\begin{aligned} \det(A) &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\ &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= \boxed{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}} \end{aligned}$$

The formula for the determinant of a 3×3 matrix looks too complicated to memorize outright. Fortunately, there is the following mnemonic device.

Recipe: Computing the Determinant of a 3×3 Matrix

To compute the determinant of a 3×3 matrix, first draw a larger matrix with the first two columns repeated on the right. Then add the products of the downward diagonals together, and subtract the products of the upward diagonals:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ \quad -a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \end{array}$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix} \quad - \quad \begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix}$$

Figure **2.6.1.5**

Alternatively, it is not necessary to repeat the first two columns if you allow your diagonals to "wrap around" the sides of a matrix, like in Pac-Man or Asteroids. (Forgive the ancient gaming reference.)

Example 2.6.1.4

Find the determinant of $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 0 & -1 \\ 4 & -3 & 1 \end{bmatrix}$

Solution

We repeat the first two columns on the right, then add the products of the downward diagonals and subtract the products of the upward diagonals:

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

Figure **2.6.1.6**

$$\begin{vmatrix} 1 & 3 & 5 \\ 2 & 0 & -1 \\ 4 & -3 & 1 \end{vmatrix} = \begin{array}{l} (1)(0)(1) + (3)(-1)(4) + (5)(2)(-3) \\ \quad -(5)(0)(4) - (1)(-1)(-3) - (3)(2)(1) \end{array} = -51$$

This clearly matches our cofactor formula, expanding along the second row:

$$\begin{aligned} \det(A) &= (-1)^{2+1}2 \cdot \begin{vmatrix} 3 & 5 \\ -3 & 1 \end{vmatrix} + (-1)^{2+2}0 \cdot \begin{vmatrix} 1 & 5 \\ 4 & 1 \end{vmatrix} + (-1)^{2+3}(-1) \cdot \begin{vmatrix} 1 & 3 \\ 4 & -3 \end{vmatrix} \\ &= -2 \cdot 18 + 0 \cdot (-19) + 1 \cdot (-15) = \boxed{-51} \end{aligned}$$

Cofactor expansions are most useful when computing the determinant of a matrix that has a row or column with several zero entries. Indeed, if the (i, j) entry of A is zero, then there is no reason to compute the (i, j) cofactor. In the following example we compute the determinant of a matrix with two zeros in the fourth column by expanding cofactors along the fourth column.

Example 2.6.1.5

Find the determinant of

$$A = \begin{bmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 0 \\ -1 & 6 & 4 & 0 \end{bmatrix}$$

Solution

The fourth column has two zero entries. We expand along the fourth column to find

$$\det(A) = 2\begin{vmatrix} -2 & -3 & 2 \\ 1 & 3 & -2 \\ -1 & 6 & 4 \end{vmatrix} - 5\begin{vmatrix} 2 & 5 & -3 \\ 1 & 3 & -2 \\ -1 & 6 & 4 \end{vmatrix} - 0\left|\text{don't care}\right| + 0\left|\text{don't care}\right|$$

We only have to compute two cofactors. We can find these determinants using any method we wish; for the sake of illustration, we will expand cofactors on one and use the formula for the 3×3 determinant on the other.

Expanding along the first column, we compute

$$\begin{vmatrix} -2 & -3 & 2 \\ 1 & 3 & -2 \\ -1 & 6 & 4 \end{vmatrix} = -2\begin{vmatrix} 3 & -2 \\ 6 & 4 \end{vmatrix} - \begin{vmatrix} -3 & 2 \\ 6 & 4 \end{vmatrix} - \begin{vmatrix} -3 & 2 \\ 3 & -2 \end{vmatrix}$$
$$= -2(24) - (-24) - 0 = -48 + 24 + 0 = -24$$

Using the formula for the 3×3 determinant, we have

$$\begin{vmatrix} 2 & 5 & -3 \\ 1 & 3 & -2 \\ -1 & 6 & 4 \end{vmatrix} = \begin{matrix} (2)(3)(4) + (5)(-2)(-1) + (-3)(1)(6) \\ -(2)(-2)(6) - (5)(1)(4) - (-3)(3)(-1) \end{matrix} = 11$$

Thus, we find that

$$\det(A) = 2(-24) - 5(11) = \boxed{-103}$$

Cofactor expansions are also very useful when computing the determinant of a matrix with unknown entries. Indeed, it is inconvenient to row reduce in this case, because one cannot be sure whether an entry containing an unknown is a pivot or not.

Example 2.6.1.6

Compute the determinant of this matrix containing the unknown λ:

$$A = \begin{bmatrix} -\lambda & 2 & 7 & 12 \\ 3 & 1-\lambda & 2 & -4 \\ 0 & 1 & -\lambda & 7 \\ 0 & 0 & 0 & 2-\lambda \end{bmatrix}$$

Solution

First we expand cofactors along the fourth row:

$$\det(A) = 0\det(\cdots) + 0\det(\cdots) + 0\det(\cdots) + (2-\lambda)\begin{vmatrix} -\lambda & 2 & 7 \\ 3 & 1-\lambda & 2 \\ 0 & 1 & -\lambda \end{vmatrix}$$

We only have to compute one cofactor. To do so, first we clear the $(3,3)$-entry by performing the column replacement $C_3 = C_3 + \lambda C_2$, which does not change the determinant:

$$\begin{vmatrix} -\lambda & 2 & 7 \\ 3 & 1-\lambda & 2 \\ 0 & 1 & -\lambda \end{vmatrix} = \begin{vmatrix} -\lambda & 2 & 7+2\lambda \\ 3 & 1-\lambda & 2+\lambda(1-\lambda) \\ 0 & 1 & 0 \end{vmatrix}$$

Now we expand cofactors along the third row to find

$$\begin{aligned} \begin{vmatrix} -\lambda & 2 & 7+2\lambda \\ 3 & 1-\lambda & 2+\lambda(1-\lambda) \\ 0 & 1 & 0 \end{vmatrix} &= (-1)^{2+3} \begin{vmatrix} -\lambda & 7+2\lambda \\ 3 & 2+\lambda(1-\lambda) \end{vmatrix} \\ &= -\Big(-\lambda\big(2+\lambda(1-\lambda)\big) - 3(7+2\lambda)\Big) \\ &= -\lambda^3 + \lambda^2 + 8\lambda + 21 \end{aligned}$$

Therefore, we have

$$\det(A) = (2-\lambda)(-\lambda^3+\lambda^2+8\lambda+21) = \boxed{\lambda^4 - 3\lambda^3 - 6\lambda^2 - 5\lambda + 42}$$

It is often most efficient to use a combination of several techniques when computing the determinant of a matrix. Indeed, when expanding cofactors on a matrix, one can compute the determinants of the cofactors in whatever way is most convenient. Or, one can perform row and column operations to clear some entries of a matrix before expanding cofactors, as in the previous example.

Note 2.6.1.2: Summary: Methods for Computing Determinants

We have several ways of computing determinants:

1. *Special formulas for* 2×2 *and* 3×3 *matrices.*
   This is usually the best way to compute the determinant of a small matrix, except for a 3×3 matrix with several zero entries.
2. *Cofactor expansion.*
   This is usually most efficient when there is a row or column with several zero entries, or if the matrix has unknown entries.
3. *Row and column operations.*
   This is generally the fastest when presented with a large matrix which does not have a row or column with a lot of zeros in it.
4. *Any combination of the above.*
   Cofactor expansion is recursive, but one can compute the determinants of the minors using whatever method is most convenient. Or, you can perform row and column operations to clear some entries of a matrix before expanding cofactors.

Remember, *all methods for computing the determinant yield the same number.*

2.6.1: Cofactor Expansion is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.1: Determinants- Definition** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **4.2: Cofactor Expansions** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.6.2: Cramer's Rule and Adjugates

If a student calculates enough inverses of 2×2 matrices, they start to see a pattern:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Longrightarrow \quad A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

It is also clear that the cofactors of this 2×2 matrix A are $C_{11} = d$, $C_{12} = -c$, $C_{21} = -b$, $C_{22} = a$, so we can rewrite the above formula as

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{bmatrix}$$

It turns out that this formula generalizes to $n \times n$ matrices.

Theorem 2.6.2.2

Let A be an invertible $n \times n$ matrix, with cofactors C_{ij}. Then

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n-1,1} & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n-1,2} & C_{n2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{1,n-1} & C_{2,n-1} & \cdots & C_{n-1,n-1} & C_{n,n-1} \\ C_{1n} & C_{2n} & \cdots & C_{n-1,n} & C_{nn} \end{bmatrix}$$

The matrix of cofactors is sometimes called the ***adjugate matrix*** of A, and is denoted $\mathbf{adj}(A)$:

$$\mathbf{adj}(A) = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n-1,1} & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n-1,2} & C_{n2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{1,n-1} & C_{2,n-1} & \cdots & C_{n-1,n-1} & C_{n,n-1} \\ C_{1n} & C_{2n} & \cdots & C_{n-1,n} & C_{nn} \end{bmatrix}$$

Note that the (i, j) cofactor C_{ij} goes in the (j, i) entry of the adjugate matrix, not the (i, j) entry: the adjugate matrix is the *transpose* of the cofactor matrix.

Remark

In fact, one always has $A \cdot \mathbf{adj}(A) = \mathbf{adj}(A) \cdot A = \det(A) I$, whether or not A is invertible.

Example 2.6.2.7

Use the adjugate to compute A^{-1}, where

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Solution

The minors are:

$$A_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad A_{12} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad A_{13} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

$$A_{21} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad A_{22} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad A_{23} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$A_{31} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad A_{32} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad A_{33} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

The cofactors are:

$$\begin{array}{lll} C_{11}=-1 & C_{12}=1 & C_{13}=-1 \\ C_{21}=1 & C_{22}=-1 & C_{23}=-1 \\ C_{31}=-1 & C_{32}=-1 & C_{33}=1 \end{array}$$

Expanding along the first row, we compute the determinant to be

$$\det(A)=1\cdot C_{11}+0\cdot C_{12}+1\cdot C_{13}=-2$$

Therefore, the inverse is

$$A^{-1}=\frac{1}{\det(A)}\begin{bmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{bmatrix}=\boxed{-\frac{1}{2}\begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix}}$$

It is clear from the previous example that the adjugate is a very inefficient way of computing the inverse of a matrix, compared to augmenting by the identity matrix and row reducing. However, it has its uses.

- If a matrix has unknown entries, then it is difficult to compute its inverse using row reduction, for the same reason it is difficult to compute the determinant that way: one cannot be sure whether an entry containing an unknown is a pivot or not.
- This formula is useful for theoretical purposes. Notice that the only denominators in the inverse using adjugates occur when dividing by the determinant: computing cofactors only involves multiplication and addition, never division. This means, for instance, that if the determinant is very small, then any measurement error in the entries of the matrix is greatly magnified when computing the inverse. In this way, it is useful in error analysis.

The proof of the adjugate inverse formula uses an interesting trick called *Cramer's Rule*, which gives a formula for the entries of the solution of an invertible matrix equation.

Theorem 2.6.2.3: Cramer's Rule

Let $\vec{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ be the solution of $A\vec{x}=\vec{b}$, where A is an invertible $n\times n$ matrix and $\vec{b}$ is a vector in $\mathbb{R}^n$. Let A_i be the matrix obtained from A by replacing the ith column by $\vec{b}$. Then

$$x_i=\frac{\det(A_i)}{\det(A)}$$

Proof

First suppose that A is the identity matrix, so that $\vec{x}=\vec{b}$. Then the matrix A_i looks like some variant of this:

$$\begin{bmatrix} 1 & 0 & b_1 & 0 \\ 0 & 1 & b_2 & 0 \\ 0 & 0 & b_3 & 0 \\ 0 & 0 & b_4 & 1 \end{bmatrix}$$

Expanding cofactors along the ith row, we see that $\det(A_i)=b_i$, so in this case,

$$x_i=b_i=\det(A_i)=\frac{\det(A_i)}{\det(A)}$$

Now let A be a general $n\times n$ matrix. One way to solve $A\vec{x}=\vec{b}$ is to row reduce the augmented matrix $[\,A\mid\vec{b}\,]$; the result is $[\,I\mid\vec{x}\,]$. By the case we handled above, it is enough to check that the quantity $\frac{\det(A_i)}{\det(A)}$ does not change when we do a row operation to $[\,A\mid\vec{b}\,]$, since $\frac{\det(A_i)}{\det(A)}=x_i$ when $A=I$.

1. Doing a row replacement on $[\,A\mid\vec{b}\,]$ does the same row replacement on A and on A_i:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right] \xrightarrow{R_2=R_2-2R_3} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21}-2a_{31} & a_{22}-2a_{32} & a_{23}-2a_{33} & b_2-2b_3 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right]$$

$$\left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right] \xrightarrow{R_2=R_2-2R_3} \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21}-2a_{31} & a_{22}-2a_{32} & a_{23}-2a_{33} \\ a_{31} & a_{32} & a_{33} \end{array}\right]$$

$$\left[\begin{array}{ccc} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{array}\right] \xrightarrow{R_2=R_2-2R_3} \left[\begin{array}{ccc} a_{11} & b_1 & a_{13} \\ a_{21}-2a_{31} & b_2-2b_3 & a_{23}-2a_{33} \\ a_{31} & b_3 & a_{33} \end{array}\right]$$

In particular, $\det(A)$ and $\det(A_i)$ are unchanged, so $\dfrac{\det(A_i)}{\det(A)}$ is unchanged.

2. Scaling a row of $[A \mid \vec{b}]$ by a factor of c scales the same row of A and of A_i by the same factor:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right] \xrightarrow{R_2=cR_2} \left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ ca_{21} & ca_{22} & ca_{23} & cb_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right]$$

$$\left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right] \xrightarrow{R_2=cR_2} \left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ ca_{21} & ca_{22} & ca_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right]$$

$$\left[\begin{array}{ccc} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{array}\right] \xrightarrow{R_2=cR_2} \left[\begin{array}{ccc} a_{11} & b_1 & a_{13} \\ ca_{21} & cb_2 & ca_{23} \\ a_{31} & b_3 & a_{33} \end{array}\right]$$

In particular, $\det(A)$ and $\det(A_i)$ are both scaled by a factor of c, so $\dfrac{\det(A_i)}{\det(A)}$ is unchanged.

3. Swapping two rows of $[A \mid \vec{b}]$ swaps the same rows of A and of A_i:

$$\left[\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right] \xrightarrow{R_1\leftrightarrow R_2} \left[\begin{array}{ccc|c} a_{21} & a_{22} & a_{23} & b_2 \\ a_{11} & a_{12} & a_{13} & b_1 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right]$$

$$\left[\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right] \xrightarrow{R_1\leftrightarrow R_2} \left[\begin{array}{ccc} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{array}\right]$$

$$\left[\begin{array}{ccc} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{array}\right] \xrightarrow{R_1\leftrightarrow R_2} \left[\begin{array}{ccc} a_{21} & b_2 & a_{23} \\ a_{11} & b_1 & a_{13} \\ a_{31} & b_3 & a_{33} \end{array}\right]$$

In particular, $\det(A)$ and $\det(A_i)$ are both negated, so $\dfrac{\det(A_i)}{\det(A)}$ is unchanged.

Example 2.6.2.8

Compute the solution of $A\vec{x} = \vec{b}$ using Cramer's rule, where

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Here the coefficients of A are unknown, but A may be assumed invertible.

Solution

First we compute the determinants of the matrices obtained by replacing the columns of A with b:

$$A_1 = \begin{bmatrix} 1 & b \\ 2 & d \end{bmatrix} \qquad \det(A_1) = d - 2b$$
$$A_2 = \begin{bmatrix} a & 1 \\ c & 2 \end{bmatrix} \qquad \det(A_2) = 2a - c$$

Now we compute

$$\frac{\det(A_1)}{\det(A)} = \frac{d-2b}{ad-bc} \qquad \frac{\det(A_2)}{\det(A)} = \frac{2a-c}{ad-bc}$$

It follows that

$$\boxed{\vec{x} = \frac{1}{ad-bc}\begin{bmatrix} d-2b \\ 2a-c \end{bmatrix}}$$

Now we use Cramer's rule to prove the adjugate formula of the inverse.

Proof

The jth column of A^{-1} is $\vec{x}_j = A^{-1}\vec{e}_j$. This vector is the solution of the matrix equation

$$A\vec{x} = A\left(A^{-1}\vec{e}_j\right) = I\vec{e}_j = \vec{e}_j$$

By Cramer's rule, the ith entry of $\vec{x}_j$ is $\dfrac{\det(A_i)}{\det(A)}$, where A_i is the matrix obtained from A by replacing the ith column of A by $\vec{e}_j$:

$$A_i = \begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} \\ a_{21} & a_{22} & 1 & a_{24} \\ a_{31} & a_{32} & 0 & a_{34} \\ a_{41} & a_{42} & 0 & a_{44} \end{bmatrix} \quad (i = 3,\ j = 2)$$

Expanding cofactors along the ith column, we see the determinant of A_i is exactly the (j,i)-cofactor C_{ji} of A. Therefore, the jth column of A^{-1} is

$$\vec{x}_j = \frac{1}{\det(A)}\begin{bmatrix} C_{j1} \\ C_{j2} \\ \vdots \\ C_{jn} \end{bmatrix}$$

and thus

$$A^{-1} = \begin{bmatrix} | & | & & | \\ \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \\ | & | & & | \end{bmatrix} = \frac{1}{\det(A)}\begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n-1,1} & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n-1,2} & C_{n2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{1,n-1} & C_{2,n-1} & \cdots & C_{n-1,n-1} & C_{n,n-1} \\ C_{1n} & C_{2n} & \cdots & C_{n-1,n} & C_{nn} \end{bmatrix}$$

2.6.2: Cramer's Rule and Adjugates is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **4.2: Cofactor Expansions** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.6.3: Practice Set

Exercise 2.6.3.1a

Find the determinant of $A = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 0 & 4 \\ 6 & 0 & -1 \end{bmatrix}$

Answer

Expanding on the second column gives:

$$\det(A) = (-1)^{1+2} \cdot 2 \cdot \begin{vmatrix} 3 & 4 \\ 6 & -1 \end{vmatrix} = -2\,(3 \cdot (-1) - 4 \cdot 6) = 54$$

Exercise 2.6.3.1b

Find the determinant of $B = \begin{bmatrix} 10 & 11 & 7 \\ 0 & 0 & -1 \\ 4 & 5 & 16 \end{bmatrix}$

Answer

Expanding on the second row gives:

$$\det(B) = (-1)^{2+3} \cdot (-1) \cdot \begin{vmatrix} 10 & 11 \\ 4 & 5 \end{vmatrix} = 1\,(10 \cdot 5 - 11 \cdot 4) = 6$$

Exercise 2.6.3.1c

Find the determinant of $C = \begin{bmatrix} 5 & 6 & 0 \\ 1 & 2 & 8 \\ 3 & 4 & 4 \end{bmatrix}$

Answer

Expanding on the first row gives:

$$\det(C) = (-1)^{1+1} \cdot 5 \cdot \begin{vmatrix} 2 & 8 \\ 4 & 4 \end{vmatrix} + (-1)^{1+2} \cdot 6 \cdot \begin{vmatrix} 1 & 8 \\ 3 & 4 \end{vmatrix} = 5\,(2 \cdot 4 - 8 \cdot 4) - 6\,(1 \cdot 4 - 8 \cdot 3) = 0$$

Exercise 2.6.3.1d

Find the determinant of $D = \begin{bmatrix} 4 & 2 & 15 & -1 \\ 0 & 0 & 4 & 0 \\ 6 & 0 & -11 & 5 \\ 0 & -4 & 21 & 8 \end{bmatrix}$

Answer

Expanding on the second row gives:

$$\det(D) = (-1)^{2+3} \cdot 4 \cdot \begin{vmatrix} 4 & 2 & -1 \\ 6 & 0 & 5 \\ 0 & -4 & 8 \end{vmatrix}$$

Expanding again on the bottom row gives:

$$\det(D) = -4\left((-1)^{3+2} \cdot (-4) \begin{vmatrix} 4 & -1 \\ 6 & 5 \end{vmatrix} + (-1)^{3+3} \cdot 8 \cdot \begin{vmatrix} 4 & 2 \\ 6 & 0 \end{vmatrix}\right) = -4\,(4\,(4 \cdot 5 - 6 \cdot (-1)) + 8\,(4 \cdot 0 - 6 \cdot 2)) = -32$$

Exercise 2.6.3.1e

Find the determinant of $E = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 6 & 7 & 8 & 9 \\ 5 & 0 & 0 & -8 \\ 8 & 3 & 3 & -5 \end{bmatrix}$

Answer

Best case here is expanding on the third row, but that still makes two 3×3 determinants to solve.
Instead we begin by doing a few row operations that do not affect the determinant:

$$E \xrightarrow{R_2=R_2-6R_1, R_3=R_3-5R_1, R_4=R_4-8R_1} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -5 & -10 & -15 \\ 0 & -10 & -15 & -28 \\ 0 & -13 & -21 & -37 \end{bmatrix} = E'$$

Now we expand on the first column:

$$\det(E) = \det(E') = (-1)^{1+1} \cdot 1 \cdot \begin{vmatrix} -5 & -10 & -15 \\ -10 & -15 & -28 \\ -13 & -21 & -37 \end{vmatrix}$$

This time we do column operations that do not affect determinants:

$$\begin{bmatrix} -5 & -10 & -15 \\ -10 & -15 & -28 \\ -13 & -21 & -37 \end{bmatrix} \xrightarrow{C_2=C_2-2C_1, C_3=C_3-3C_1} \begin{bmatrix} -5 & 0 & 0 \\ -10 & 5 & 2 \\ -13 & 5 & 2 \end{bmatrix}$$

And we can expand on our new first row:

$$\det(E) = 1 \cdot \left((-1)^{1+1} \cdot (-5) \cdot \begin{vmatrix} 5 & 2 \\ 5 & 2 \end{vmatrix} \right) = -5\,(5 \cdot 2 - 5 \cdot 2) = 0$$

Exercise 2.6.3.1f

Find the determinant of $F = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{8} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{2}{3} \\ 10 & 10 & 0 & 10 \\ 1 & 0 & 0 & 4 \end{bmatrix}$

Answer

Best case here is expanding on the fourth row or third column, but that still makes two 3×3 determinants to solve.
Instead we begin by doing a column operation that does not affect the determinant:

$$F \xrightarrow{C_1=C_1-C_2} \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{8} & \frac{1}{2} \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{2}{3} \\ 0 & 10 & 0 & 10 \\ 1 & 0 & 0 & 4 \end{bmatrix} = F'$$

We can expand this along the first column:

$$\det(F) = \det(F') = (-1)^{4+1} \cdot 1 \cdot \begin{vmatrix} \frac{1}{4} & \frac{1}{8} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{2}{3} \\ 10 & 0 & 10 \end{vmatrix}$$

Now we expand on the bottom row:

$$\det(F) = -1 \left((-1)^{3+1} \cdot 10 \cdot \begin{vmatrix} \frac{1}{8} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{vmatrix} + (-1)^{3+3} \cdot 10 \cdot \begin{vmatrix} \frac{1}{4} & \frac{1}{8} \\ \frac{1}{3} & \frac{1}{3} \end{vmatrix} \right)$$

$$= -10 \left(\frac{1}{8} \cdot \frac{2}{3} - \frac{1}{3} \cdot \frac{1}{2} \right) - 10 \left(\frac{1}{4} \cdot \frac{1}{3} - \frac{1}{3} \cdot \frac{1}{8} \right) = \frac{5}{12}$$

? Exercise 2.6.3.1*g*

Find the determinant of $G = \begin{bmatrix} 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 2 & -7 \\ 0 & 0 & 3 & 10 & 8 \\ 0 & -1 & 0 & 4 & 8 \\ 5 & 5 & 5 & 4 & -5 \end{bmatrix}$

Answer

Expanding on the first row *repeatedly* gives:

$$\det(G) = (-1)^{1+5} \cdot 4 \cdot \begin{vmatrix} 0 & 0 & 0 & 2 \\ 0 & 0 & 3 & 10 \\ 0 & -1 & 0 & 4 \\ 5 & 5 & 5 & 4 \end{vmatrix} = 4 \left((-1)^{1+4} \cdot 2 \cdot \begin{vmatrix} 0 & 0 & 3 \\ 0 & -1 & 0 \\ 5 & 5 & 5 \end{vmatrix} \right) = 4 \cdot (-2) \left((-1)^{1+3} \cdot 3 \cdot \begin{vmatrix} 0 & -1 \\ 5 & 5 \end{vmatrix} \right)$$

$$= 4 \cdot (-2) \cdot 3 \, (0 \cdot 5 - 5 \cdot (-1)) = -120$$

? Exercise 2.6.3.2*a*

Find the value(s) of h which make $A = \begin{bmatrix} 4-h & 2 \\ 3 & 5-h \end{bmatrix}$ invertible

Answer

We calculate $\det(A) = (4-h) \cdot (5-h) - 2 \cdot 3 = h^2 - 9h + 14 = (h-2)(h-7)$

A is invertible when $\det(A) \neq 0$, so our solution is $h \neq 2, 7$

? Exercise 2.6.3.2*b*

Find the value(s) of h which make $B = \begin{bmatrix} h & -3 \\ 2 & 5-h \end{bmatrix}$ invertible

Answer

We calculate $\det(B) = h \cdot (5-h) - 2 \cdot (-3) = -h^2 + 5h + 6 = -(h-6)(h+1)$

B is invertible when $\det(B) \neq 0$, so our solution is $h \neq -1, 6$

? Exercise 2.6.3.2*c*

Find the value(s) of h which make $C = \begin{bmatrix} h & 0 & -7 \\ 0 & 2 & 6 \\ 1 & 4 & 5 \end{bmatrix}$ invertible

Answer

We expand along the top row to find the determinant of C:

$$\det(C) = (-1)^{1+1} \cdot h \cdot \begin{vmatrix} 2 & 6 \\ 4 & 5 \end{vmatrix} + (-1)^{1+3} \cdot (-7) \cdot \begin{vmatrix} 0 & 2 \\ 1 & 4 \end{vmatrix} = h \, (2 \cdot 5 - 4 \cdot 6) - 7 \, (0 \cdot 4 - 1 \cdot 2) = -14h + 14$$

C is invertible when $\det(C) \neq 0$, so our solution is $h \neq 1$

? Exercise 2.6.3.2d

Find the value(s) of h which make $D = \begin{bmatrix} h & 0 & h \\ 5 & 7 & -2 \\ -1 & 4 & 3 \end{bmatrix}$ invertible

Answer

We expand along the top row to find the determinant of D:

$$\det(D) = (-1)^{1+1} \cdot h \cdot \begin{vmatrix} 7 & -2 \\ 4 & 3 \end{vmatrix} + (-1)^{1+3} \cdot h \cdot \begin{vmatrix} 5 & 7 \\ -1 & 4 \end{vmatrix} = h\,(7 \cdot 3 - 4 \cdot (-2)) + h\,(5 \cdot 4 - (-1) \cdot 7) = 56h$$

D is invertible when $\det(D) \neq 0$, so our solution is $h \neq 0$

? Exercise 2.6.3.2e

Find the value(s) of h which make $E = \begin{bmatrix} h & h & h \\ 1 & 4 & 3 \\ 3 & 0 & 1 \end{bmatrix}$ invertible

Answer

Before we do any expanding, we will do a couple of column operations that do not change our determinant:

$$E \xrightarrow{C_1 = C_1 - C_2, C_3 = C_3 - C_2} \begin{bmatrix} 0 & h & 0 \\ -3 & 4 & -1 \\ 3 & 0 & 1 \end{bmatrix} = E'$$

We expand along the top row to find the determinant of E:

$$\det(E) = \det(E') = (-1)^{1+2} \cdot h \cdot \begin{vmatrix} -3 & -1 \\ 3 & 1 \end{vmatrix} = -h\,((-3) \cdot 1 - 3 \cdot (-1)) = 0$$

E is invertible when $\det(E) \neq 0$, which is impossible regardless of what h is.

This matrix is **never invertible.**

? Exercise 2.6.3.2f

Find the value(s) of h which make $F = \begin{bmatrix} 5 & 0 & h & 1 \\ 2 & 0 & 2 & 1 \\ -11 & 4 & h & 1 \\ 1 & 0 & 2 & -1 \end{bmatrix}$ invertible

Answer

We expand along the second column to get:

$$\det(F) = (-1)^{3+2} \cdot 4 \cdot \begin{vmatrix} 5 & h & 1 \\ 2 & 2 & 1 \\ 1 & 2 & -1 \end{vmatrix}$$

Now we do a couple row operations that do not change our determinant:

$$\begin{bmatrix} 5 & h & 1 \\ 2 & 2 & 1 \\ 1 & 2 & -1 \end{bmatrix} \xrightarrow{R_1 = R_1 + R_3, R_2 = R_2 + R_3} \begin{bmatrix} 6 & h+2 & 0 \\ 3 & 4 & 0 \\ 1 & 2 & -1 \end{bmatrix}$$

We expand along the third column:

$$\det(F) = -4 \left((-1)^{3+3} \cdot (-1) \cdot \begin{vmatrix} 6 & h+2 \\ 3 & 4 \end{vmatrix} \right) = 4\,(6 \cdot 4 - 3 \cdot (h+2)) = 72 - 12h = 12(6-h)$$

F is invertible when $\det(F) \neq 0$, so our solution is $h \neq 6$

Exercise 2.6.3.3a

Use Cramer's Rule to solve the system $\begin{cases} 7x_1 + 6x_2 = 11 \\ 13x_1 + 5x_2 = 21 \end{cases}$

Answer

This is represented by $A\vec{x} = \vec{b}$ where $A = \begin{bmatrix} 7 & 6 \\ 13 & 5 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 11 \\ 21 \end{bmatrix}$

We find $\det(A) = 7 \cdot 5 - 13 \cdot 6 = -43$

Now we replace the first column of A with $\vec{b}$ to make $A_1 = \begin{bmatrix} 11 & 6 \\ 21 & 5 \end{bmatrix}$

This gives $\det(A_1) = 11 \cdot 5 - 21 \cdot 6 = -71$

Now we replace the second column of A with $\vec{b}$ to make $A_2 = \begin{bmatrix} 7 & 11 \\ 13 & 21 \end{bmatrix}$

This gives $\det(A_2) = 7 \cdot 21 - 13 \cdot 11 = 4$

Therefore by Cramer's Rule, we know $x_1 = \dfrac{\det(A_1)}{\det(A)} = \dfrac{71}{43}$ and $x_2 = \dfrac{\det(A_2)}{\det(A)} = -\dfrac{4}{43}$

Exercise 2.6.3.3b

Use Cramer's Rule to solve the system $\begin{cases} 4x_1 + 2x_2 - x_3 = 4 \\ x_2 + 3x_3 = 11 \\ 4x_1 - 3x_2 = 7 \end{cases}$

Answer

This is represented by $A\vec{x} = \vec{b}$ where $A = \begin{bmatrix} 4 & 2 & -1 \\ 0 & 1 & 3 \\ 4 & -3 & 0 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 4 \\ 11 \\ 7 \end{bmatrix}$

By expanding on the first column we get:

$\det(A) = (-1)^{1+1} \cdot 4 \cdot \begin{vmatrix} 1 & 3 \\ -3 & 0 \end{vmatrix} + (-1)^{3+1} \cdot 4 \cdot \begin{vmatrix} 2 & -1 \\ 1 & 3 \end{vmatrix} = 4(1 \cdot 0 - 3 \cdot (-3)) + 4(2 \cdot 3 - 1 \cdot (-1)) = 64$

Now we replace the first column of A with $\vec{b}$ to make $A_1 = \begin{bmatrix} 4 & 2 & -1 \\ 11 & 1 & 3 \\ 7 & -3 & 0 \end{bmatrix}$

Expanding on row three gives $\det(A_1) = (-1)^{3+1} \cdot 7 \cdot \begin{vmatrix} 2 & -1 \\ 1 & 3 \end{vmatrix} + (-1)^{3+2} \cdot (-3) \cdot \begin{vmatrix} 4 & -1 \\ 11 & 3 \end{vmatrix} = 7(2 \cdot 3 - (-1) \cdot 1) + 3(4 \cdot 3 - 11 \cdot (-1)) = 118$

Now we replace the second column of A with $\vec{b}$ to make $A_2 = \begin{bmatrix} 4 & 4 & -1 \\ 0 & 11 & 3 \\ 4 & 7 & 0 \end{bmatrix}$

Expanding on column one gives $\det(A_2) = (-1)^{1+1} \cdot 4 \cdot \begin{vmatrix} 11 & 3 \\ 7 & 0 \end{vmatrix} + (-1)^{3+1} \cdot 4 \cdot \begin{vmatrix} 4 & -1 \\ 11 & 3 \end{vmatrix} = 4(11 \cdot 0 - 3 \cdot 7) + 4(4 \cdot 3 - 11 \cdot (-1)) = 8$

Now we replace the third column of A with $\vec{b}$ to make $A_3 = \begin{bmatrix} 4 & 2 & 4 \\ 0 & 1 & 11 \\ 4 & -3 & 7 \end{bmatrix}$

Expanding on column one gives $\det(A_3) = (-1)^{1+1} \cdot 4 \cdot \begin{vmatrix} 1 & 11 \\ -3 & 7 \end{vmatrix} + (-1)^{3+1} \cdot 4 \cdot \begin{vmatrix} 2 & 4 \\ 1 & 11 \end{vmatrix} = 4(1 \cdot 7 - 11 \cdot (-3)) + 4(2 \cdot 11 - 1 \cdot 4) = 232$

Therefore by Cramer's Rule, we know $x_1 = \frac{\det(A_1)}{\det(A)} = \frac{59}{32}$, $x_2 = \frac{\det(A_2)}{\det(A)} = \frac{1}{8}$ and $x_3 = \frac{\det(A_3)}{\det(A)} = \frac{29}{8}$

? Exercise 2.6.3.3c

Use Cramer's Rule to solve the system $\begin{cases} \frac{2}{3}x_1 + \frac{1}{5}x_2 = \alpha \\ \frac{5}{2}x_1 - \frac{3}{4}x_2 = \beta \end{cases}$

Answer

This is represented by $A\vec{x} = \vec{b}$ where $A = \begin{bmatrix} \frac{2}{3} & \frac{1}{5} \\ \frac{5}{2} & -\frac{3}{4} \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$

We find $\det(A) = \frac{2}{3} \cdot \left(-\frac{3}{4}\right) - \frac{5}{2} \cdot \frac{1}{5} = -1$

Now we replace the first column of A with $\vec{b}$ to make $A_1 = \begin{bmatrix} \alpha & \frac{1}{5} \\ \beta & -\frac{3}{4} \end{bmatrix}$

This gives $\det(A_1) = \alpha \cdot \left(-\frac{3}{4}\right) - \beta \cdot \frac{1}{5} = \frac{-15\alpha - 4\beta}{20}$

Now we replace the second column of A with $\vec{b}$ to make $A_2 = \begin{bmatrix} \frac{2}{3} & \alpha \\ \frac{5}{2} & \beta \end{bmatrix}$

We find $\det(A_2) = \frac{2}{3} \cdot \beta - \frac{5}{2} \cdot \alpha = \frac{-15\alpha + 4\beta}{6}$

Therefore by Cramer's Rule, we know $x_1 = \frac{\det(A_1)}{\det(A)} = \frac{15\alpha + 4\beta}{20}$ and $x_2 = \frac{\det(A_2)}{\det(A)} = \frac{15\alpha - 4\beta}{6}$

2.6.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

2.7: Determinants and Volumes

Objectives

1. Understand the relationship between the determinant of a matrix and the volume of a parallelepiped.
2. Learn to use determinants to compute volumes of parallelograms and triangles.
3. Learn to use determinants to compute the volume of some curvy shapes like ellipses.
4. *Pictures:* parallelepiped, the image of a curvy shape under a linear transformation.
5. *Theorem:* determinants and volumes.
6. *Vocabulary word:* ***parallelepiped***.

In this section we give a geometric interpretation of determinants, in terms of *areas and volumes.* This will shed light on the reason behind three of the four defining properties of the determinant. It is also a crucial ingredient in the change-of-variables formula in multivariable calculus.

This page titled 2.7: Determinants and Volumes is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **4.3: Determinants and Volumes** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.7: Determinants for Area and Volume

The determinant computes the volume of the following kind of geometric object.

Definition: Parallelepiped

The ***parallelepiped*** determined by n vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ in $\mathbb{R}^n$ is the subset

$$P = \left\{ a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n \ : \ 0 \leq a_1, a_2, \ldots, a_n \leq 1 \right\}$$

In other words, a parallelepiped is the set of all linear combinations of n vectors with coefficients in $[0, 1]$. We can draw parallelepipeds using the parallelogram law for vector addition.

Illustration 2.7.1: The Unit Cube

The parallelepiped determined by the standard coordinate vectors $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ is the unit n-dimensional cube.

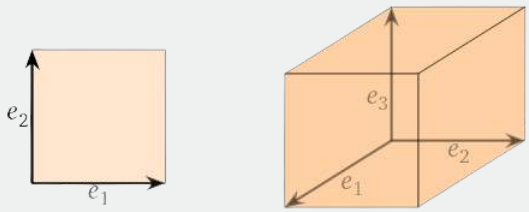


Figure 2.7.1

Illustration 2.7.2: Parallelograms

When $n = 2$, a parallelepiped is just a parallelogram in $\mathbb{R}^2$. Note that the edges come in parallel pairs.

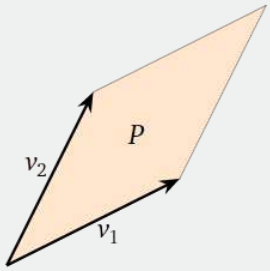


Figure 2.7.2

Illustration 2.7.3

When $n = 3$, a parallelepiped is a kind of a skewed cube. Note that the faces come in parallel pairs.

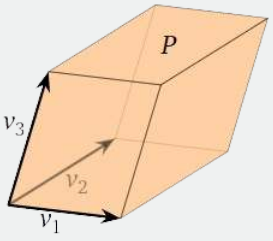


Figure 2.7.3

When does a parallelepiped have zero volume? This can happen only if the parallelepiped is flat, i.e., it is squashed into a lower dimension.

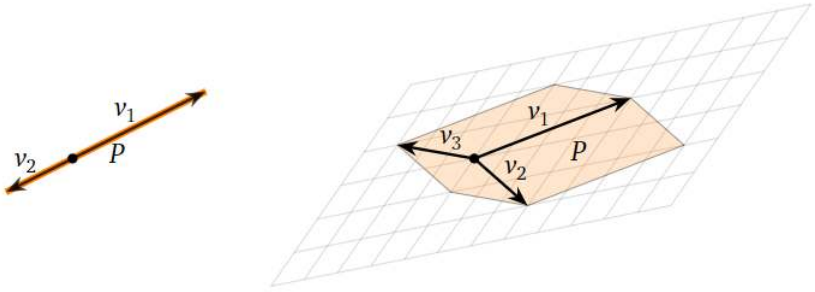


Figure 2.7.4

This means exactly that $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ is *linearly dependent*, which is equivalent to saying the matrix with rows $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ has determinant equal to zero. To summarize:

Key Observation 2.7.1

The parallelepiped defined by $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ has zero volume if and only if the matrix with rows $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ has zero determinant.

The key observation above is only the beginning of the story: the volume of a parallelepiped is *always* a determinant.

Theorem 2.7.1: Determinants and Volumes

Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ be vectors in $\mathbb{R}^n$, let P be the parallelepiped determined by these vectors, and let A be the matrix with rows $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$. Then the absolute value of the determinant of A is the volume of P:

$$|\det(A)| = \text{volume of } P$$

Proof

Since the four defining properties characterize the determinant, they also characterize the absolute value of the determinant. Explicitly, $|\det|$ is a function on square matrices which satisfies these properties:

1. Doing a row replacement on A does not change $|\det(A)|$.
2. Scaling a row of A by a scalar c multiplies $|\det(A)|$ by $|c|$.
3. Swapping two rows of a matrix does not change $|\det(A)|$.
4. The determinant of the identity matrix I is equal to 1.

The absolute value of the determinant is the *only* such function: indeed, if you do some number of row operations on A to obtain a matrix B in row echelon form, then

$$|\det(A)| = \left| \frac{(\text{product of the diagonal entries of } B)}{(\text{product of scaling factors used})} \right|$$

For a square matrix A, we abuse notation and let $\text{vol}(A)$ denote the volume of the parallelepiped determined by the rows of A. Then we can regard vol as a function from the set of square matrices to the real numbers. We will show that vol also satisfies the above four properties.

1. For simplicity, we consider a row replacement of the form $R_n = R_n + cR_i$. The volume of a parallelepiped is the volume of its base, times its height: here the "base" is the parallelepiped determined by $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{n-1}$, and the "height" is the perpendicular distance of $\vec{v}_n$ from the base.

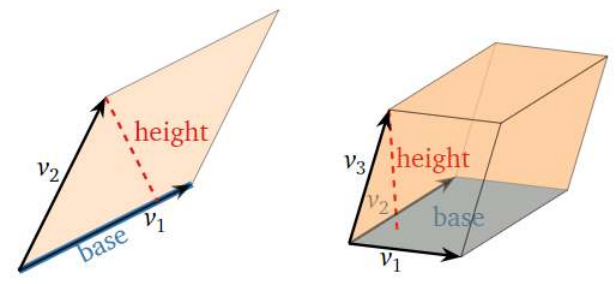


Figure 2.7.5

Translating $\vec{v}_n$ by a multiple of $\vec{v}_i$ moves $\vec{v}_n$ in a direction parallel to the base. This changes neither the base nor the height! Thus, $\text{vol}(A)$ is unchanged by row replacements.

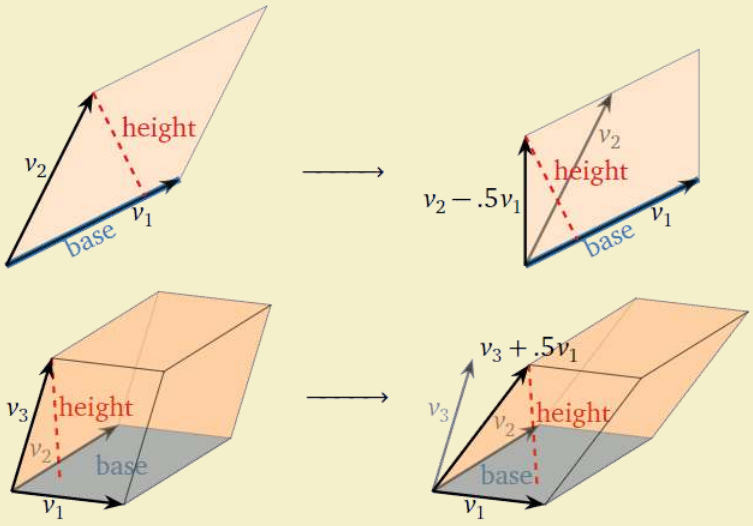
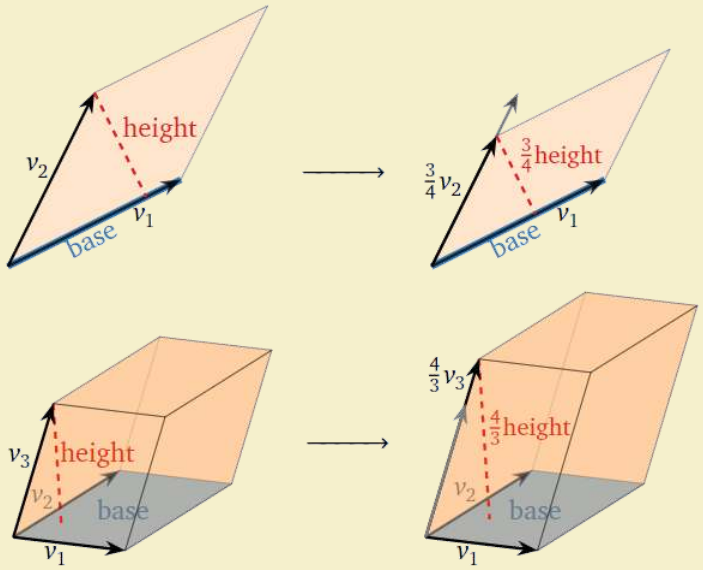


Figure 2.7.6

2. For simplicity, we consider a row scale of the form $R_n = cR_n$. This scales the length of $\vec{v}_n$ by a factor of $|c|$, which also scales the perpendicular distance of $\vec{v}_n$ from the base by a factor of $|c|$. Thus, $\text{vol}(A)$ is scaled by $|c|$.



https://math.libretexts.org/@go/page/186426

Figure 2.7.7

3. Swapping two rows of A just reorders the vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, hence has no effect on the parallelepiped determined by those vectors. Thus, $\text{vol}(A)$ is unchanged by row swaps.

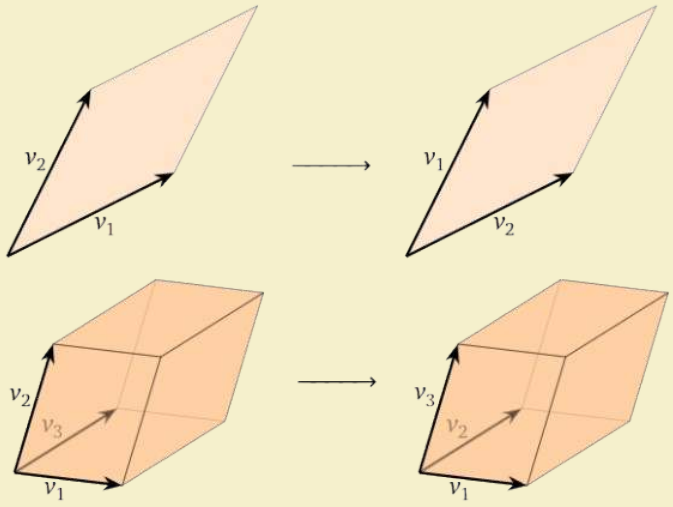


Figure 2.7.8

4. The rows of the identity matrix I are the standard coordinate vectors $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. The associated parallelepiped is the unit cube, which has volume 1. Thus, $\text{vol}(I) = 1$.

Since $|\det|$ is the only function satisfying these properties, we have

$$\text{volume of } P = \text{vol}(A) = |\det(A)|$$

This completes the proof.

Since $\det(A) = \det(A^T)$ by the transpose property, the absolute value of $\det(A)$ is also equal to the volume of the parallelepiped determined by the *columns* of A as well.

Illustration 2.7.4: Length

A 1×1 matrix A is just a number $\begin{pmatrix} a \end{pmatrix}$. In this case, the parallelepiped P determined by its one row is just the interval $[0, a]$ (or $[a, 0]$ if $a < 0$). The "volume" of a region in $\mathbb{R}^1 = \mathbb{R}$ is just its length, so it is clear in this case that $\text{volume of } P = |a|$.

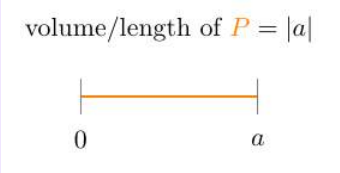
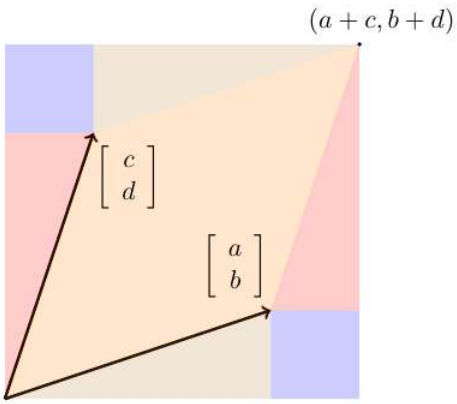


Figure 2.7.9

Illustration 2.7.5: Area

When A is a 2×2 matrix, its rows determine a parallelogram in $\mathbb{R}^2$. The "volume" of a region in $\mathbb{R}^2$ is its area, so we obtain a formula for the area of a parallelogram: it is the determinant of the matrix whose rows are the vectors forming two adjacent sides of the parallelogram. It is perhaps surprising that it is possible to compute the area of a parallelogram without trigonometry. Here is a visual proof of the formula:



$$\begin{aligned} &\text{volume/area of } P \\ &= (a+c)(b+d) - 2\cdot bc - 2\cdot \tfrac{1}{2}cd - 2\cdot \tfrac{1}{2}ab \\ &= \left|\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)\right| = |ad-bc| \end{aligned}$$

Figure 2.7.10

Example 2.7.1

Find the area of the parallelogram with sides $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ -3 \end{bmatrix}$

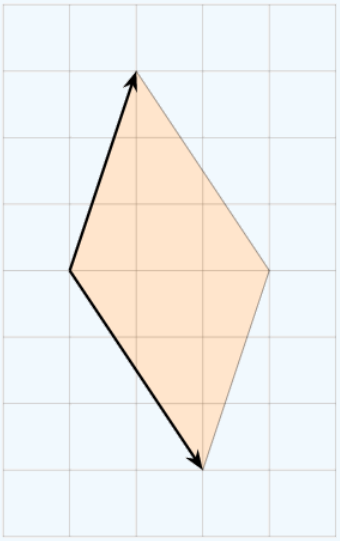


Figure 2.7.11

Solution

The area is

$$\left|\det\left(\begin{bmatrix} 1 & 3 \\ 2 & -3 \end{bmatrix}\right)\right| = |-3-6| = \boxed{9}$$

Example 2.7.2

Find the area of the parallelogram in the picture.

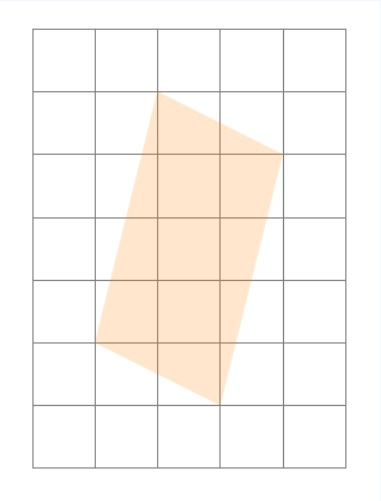


Figure 2.7.12

Solution

We choose two adjacent sides to be the rows of a matrix. We choose the top two:

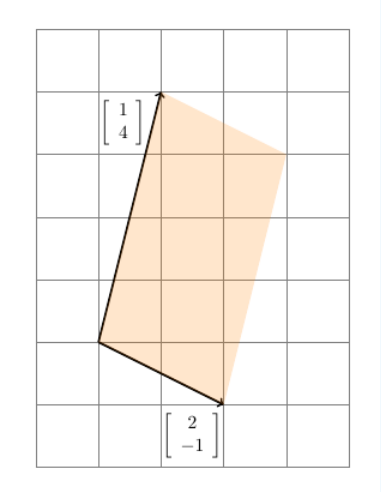


Figure 2.7.13

Note that we do not need to know where the origin is in the picture: vectors are determined by their length and direction, not where they start. The area is

$$\left|\det\left(\begin{bmatrix} 1 & 4 \\ 2 & -1 \end{bmatrix}\right)\right| = |-1-8| = \boxed{9}$$

Example 2.7.3: Area of a triangle

Find the area of the triangle with vertices $(1,2)$, $(3,6)$, $(4,5)$

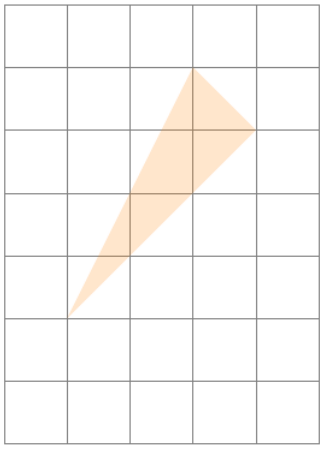


Figure 2.7.14

Solution

Doubling a triangle makes a parallelogram. We choose two of its sides to be the rows of a matrix.

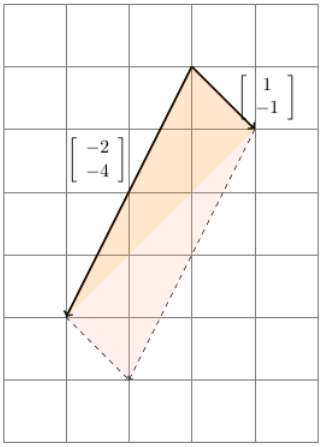


Figure 2.7.15

The area of the parallelogram is

$$\left|\det\left(\begin{bmatrix} -2 & -4 \\ 1 & -1 \end{bmatrix}\right)\right| = |2+4| = 6$$

so the area of the triangle is $\boxed{3}$

You might be wondering: if the absolute value of the determinant is a volume, what is the geometric meaning of the determinant without the absolute value? The next remark explains that we can think of the determinant as a *signed* volume. If you have taken an integral calculus course, you probably computed negative areas under curves; the idea here is similar.

Remark: Signed volumes

Theorem 2.7.1 on determinants and volumes tells us that the *absolute value* of the determinant is the volume of a parallelepiped. This raises the question of whether the sign of the determinant has any geometric meaning.

A 1×1 matrix A is just a number $\begin{pmatrix} a \end{pmatrix}$. In this case, the parallelepiped P determined by its one row is just the interval $[0, a]$ if $a \geq 0$, and it is $[a, 0]$ if $a < 0$. In this case, the sign of the determinant determines whether the interval is to the left or the right of

the origin.

For a 2×2 matrix with rows v_1, v_2, the sign of the determinant determines whether v_2 is counterclockwise or clockwise from v_1. That is, if the counterclockwise angle from v_1 to v_2 is less than 180°, then the determinant is positive; otherwise it is negative (or zero).

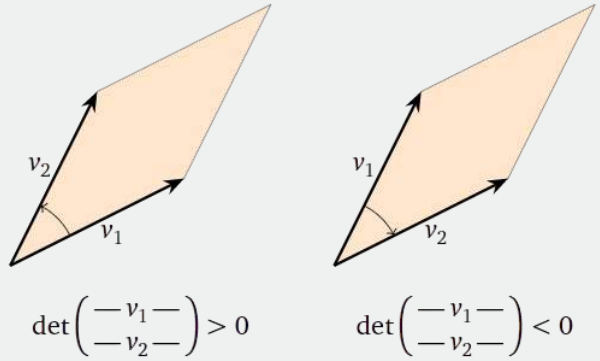


Figure 2.7.16

For example, if $v_1 = \begin{bmatrix} a \\ b \end{bmatrix}$, then the counterclockwise rotation of v_1 by 90° is $v_2 = \begin{bmatrix} -b \\ a \end{bmatrix}$, and

$$\det\left(\begin{bmatrix} a & b \\ -b & a \end{bmatrix}\right) = a^2 + b^2 > 0$$

On the other hand, the *clockwise* rotation of v_1 by 90° is $\begin{bmatrix} b \\ -a \end{bmatrix}$, and

$$\det\left(\begin{bmatrix} a & b \\ b & -a \end{bmatrix}\right) = -a^2 - b^2 < 0$$

For a 3×3 matrix with rows v_1, v_2, v_3, the *right-hand rule* determines the sign of the determinant. If you point the index finger of your right hand in the direction of v_1 and your middle finger in the direction of v_2, then the determinant is positive if your thumb points roughly in the direction of v_3, and it is negative otherwise.

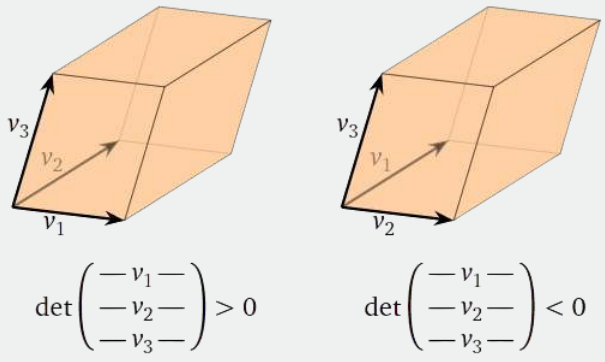


Figure 2.7.17

In higher dimensions, the notion of signed volume is still important, but it is usually *defined* in terms of the sign of a determinant.

Let A be an $n \times n$ matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, and let $T\colon \mathbb{R}^n \to \mathbb{R}^n$ be the associated matrix transformation, $T(\vec{x}) = A\vec{x}$. Then $T(\vec{e}_1) = \vec{v}_1$ and $T(\vec{e}_2) = \vec{v}_2$, so T takes the unit cube C to the parallelepiped P determined by $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$:

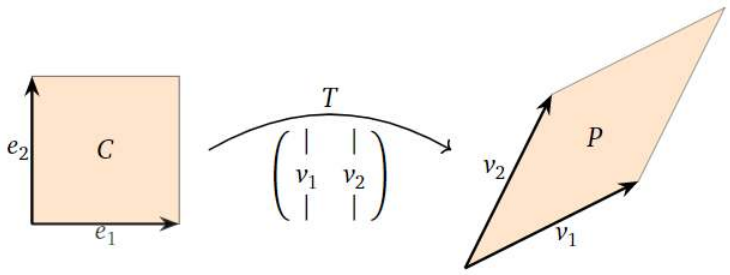


Figure 2.7.18

Since the unit cube has volume 1 and its image has volume $|\det(A)|$, the transformation T scaled the volume of the cube by a factor of $|\det(A)|$. To rephrase:

Note 2.7.1

If A is an $n \times n$ matrix with corresponding matrix transformation $T\colon \mathbb{R}^n \to \mathbb{R}^n$, and if C is the unit cube in $\mathbb{R}^n$, then the volume of $T(C)$ is $|\det(A)|$.

The notation $T(S)$ means the image of the region S under the transformation T. In set builder notation, this is the subset

$$T(S) = \{T(\vec{x}) \,:\, \vec{x} \text{ in } S\}$$

In fact, T scales the volume of *any* region in $\mathbb{R}^n$ by the same factor, even for curvy regions.

Theorem 2.7.2

Let A be an $n \times n$ matrix, and let $T\colon \mathbb{R}^n \to \mathbb{R}^n$ be the associated matrix transformation $T(\vec{x}) = A\vec{x}$. If S is any region in $\mathbb{R}^n$, then

$$\text{volume/area of } T(S) = |\det(A)| \cdot \text{volume/area of } S$$

Proof

Let C be the unit cube, let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be the columns of A, and let P be the parallelepiped determined by these vectors, so $T(C) = P$ and volume of $P = |\det(A)|$. For $\epsilon > 0$ we let ϵC be the cube with side lengths ϵ, i.e., the parallelepiped determined by the vectors $\epsilon\vec{e}_1, \epsilon\vec{e}_2, \ldots, \epsilon\vec{e}_n$, and we define ϵP similarly. By the second defining property, T takes ϵC to ϵP. The volume of ϵC is ϵ^n (we scaled each of the n standard vectors by a factor of ϵ) and the volume of ϵP is $\epsilon^n|\det(A)|$ (for the same reason), so we have shown that T scales the volume of ϵC by $|\det(A)|$.

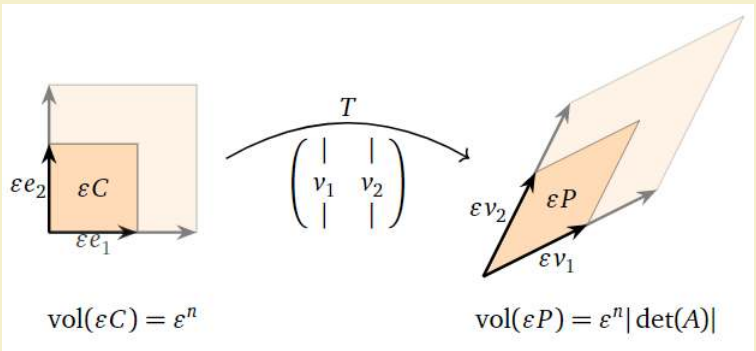


Figure 2.7.19

We can see that the image of a translate of ϵC is a translate of ϵP:

$$T(\vec{x} + \epsilon C) = T(\vec{x}) + \epsilon T(C) = T(\vec{x}) + \epsilon P$$

Since a translation does not change volumes, this proves that T scales the volume of a translate of ϵC by $|\det(A)|$.

At this point, we need to use techniques from multivariable calculus, so we only give an idea of the rest of the proof. Any region S can be approximated by a collection of very small cubes of the form $\vec{x} + \epsilon C$. The image $T(S)$ is then approximated by the image of this collection of cubes, which is a collection of very small parallelepipeds of the form $T(\vec{x}) + \epsilon P$.

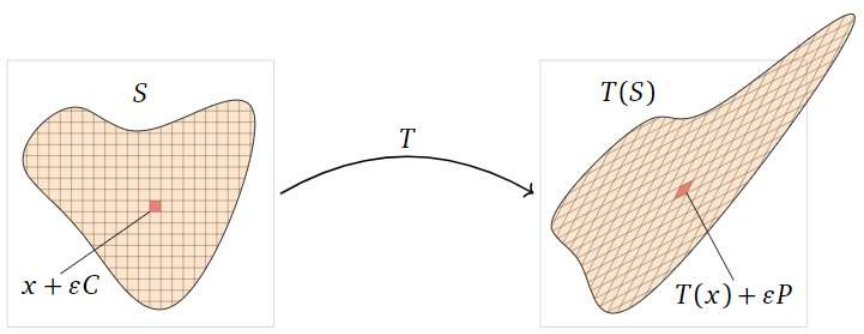


Figure 2.7.20

The volume of S is closely approximated by the sum of the volumes of the cubes; in fact, as ϵ goes to zero, the limit of this sum is precisely $\text{volume of } S$. Likewise, the volume of $T(S)$ is equal to the sum of the volumes of the parallelepipeds, taken in the limit as $\epsilon \to 0$. The key point is that *the volume of each cube is scaled by* $|\det(A)|$. Therefore, the sum of the volumes of the parallelepipeds is $|\det(A)|$ times the sum of the volumes of the cubes.

This proves that $\text{volume/area of } T(S) = |\det(A)| \cdot \text{volume/area of } S$

Example 2.7.4

Let S be a half circle of radius 1, let

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$

and define $T: \mathbb{R}^2 \to \mathbb{R}^2$ by $T(\vec{x}) = A\vec{x}$. What is the area of $T(S)$?

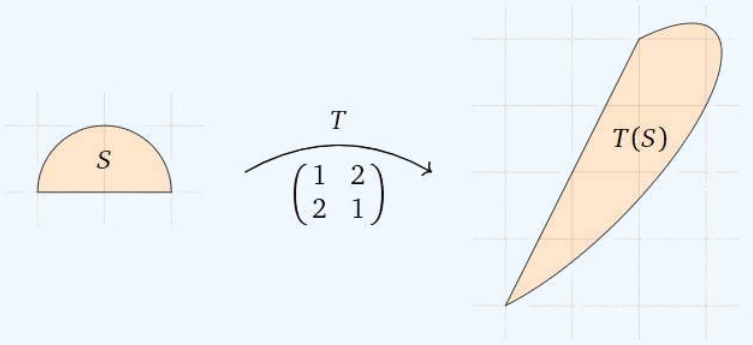


Figure 2.7.21

Solution

The area of the unit circle is π, so the area of S is $\frac{\pi}{2}$. The transformation T scales areas by a factor of $|\det(A)| = |1 - 4| = 3$, so

$$\text{area of } T(S) = 3 \cdot \text{area of } S = \boxed{\frac{3\pi}{2}}$$

Example 2.7.5: Area of an ellipse

Find the area of the interior of the ellipse defined by the equation

$$\left(\frac{2x-y}{2}\right)^2+\left(\frac{y+3x}{3}\right)^2=1$$

Solution

This ellipse is obtained from the unit circle $X^2+Y^2=1$ by the linear change of coordinates

$$X=\frac{2x-y}{2}$$
$$Y=\frac{y+3x}{3}$$

In other words, if we define a linear transformation $T\colon\mathbb{R}^2\to\mathbb{R}^2$ by

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}\frac{2x-y}{2}\\ \frac{y+3x}{3}\end{bmatrix}$$

then $T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)$ lies on the unit circle C whenever $\begin{bmatrix}x\\y\end{bmatrix}$ lies on E.

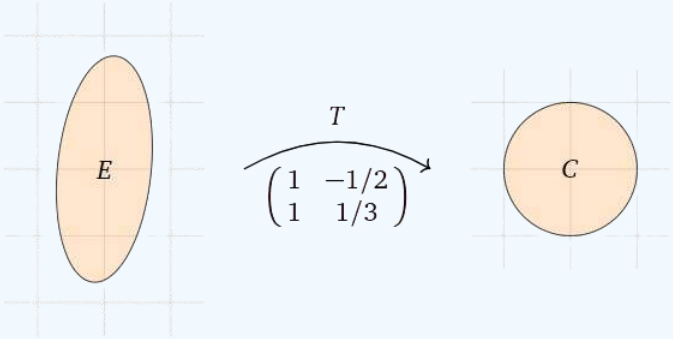


Figure 2.7.22

We compute the standard matrix A for T by evaluating on the standard coordinate vectors:

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)=\begin{bmatrix}1\\1\end{bmatrix}\qquad T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)=\begin{bmatrix}-\frac{1}{2}\\ \frac{1}{3}\end{bmatrix}\qquad\Longrightarrow\qquad A=\begin{bmatrix}1&-\frac{1}{2}\\1&\frac{1}{3}\end{bmatrix}$$

Therefore, T scales areas by a factor of $|\det(A)|=\left|\frac{1}{3}+\frac{1}{2}\right|=\frac{5}{6}$. The area of the unit circle is π, so

$$\pi=\text{area of } C=\text{area of } T(E)=|\det(A)|\cdot\text{area of } E=\frac{5}{6}\cdot\text{area of } E$$

and thus the area of E is $\boxed{\frac{6\pi}{5}}$

Remark: Multiplicativity of $|\det|$

The above Theorem 2.7.2 also gives a geometric reason for multiplicativity of the (absolute value of the) determinant. Indeed, let A and B be $n\times n$ matrices, and let $T,U\colon\mathbb{R}^n\to\mathbb{R}^n$ be the corresponding matrix transformations. If C is the unit cube, then

$$\begin{aligned}\text{volume of } (T\circ U)(C)=\text{volume of }\big(T(U(C))\big)&=|\det(A)|\cdot\text{volume of } U(C)\\&=|\det(A)|\cdot|\det(B)|\cdot\text{volume of } C\\&=|\det(A)|\cdot|\det(B)|\cdot 1\end{aligned}$$

On the other hand, the matrix for the composition $T\circ U$ is the product AB, so

$$\text{volume of } (T\circ U)(C)=|\det(AB)|\cdot\text{volume of } C=|\det(AB)|$$

Thus $|\det(AB)|=|\det(A)|\cdot|\det(B)|$.

2.7: Determinants for Area and Volume is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

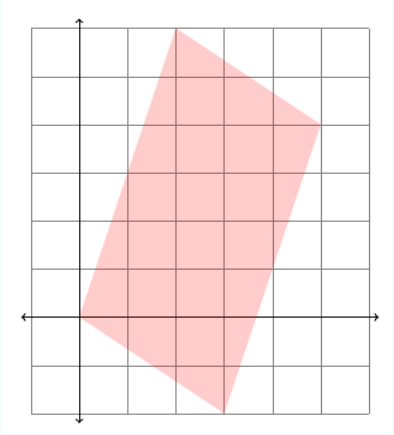
- **4.3: Determinants and Volumes** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

2.7: Practice Set

Exercise 2.7.1a

Find the area of the parallelogram with adjacent sides determined by the vectors $\vec{a} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$

Answer



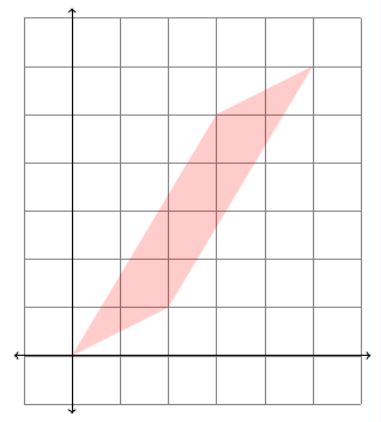
We take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns: $\begin{vmatrix} 2 & 3 \\ 6 & -2 \end{vmatrix} = 2 \cdot (-2) - 3 \cdot 6 = -22$

So the area is $|-22| = 22$

Exercise 2.7.1b

Find the area of the parallelogram with vertices $(0, 0), (2, 1), (3, 5)$, and $(5, 6)$.

Answer



First we need the vectors connecting these vertices:

$\vec{a} = \begin{bmatrix} 2-0 \\ 1-0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 3-0 \\ 5-0 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$

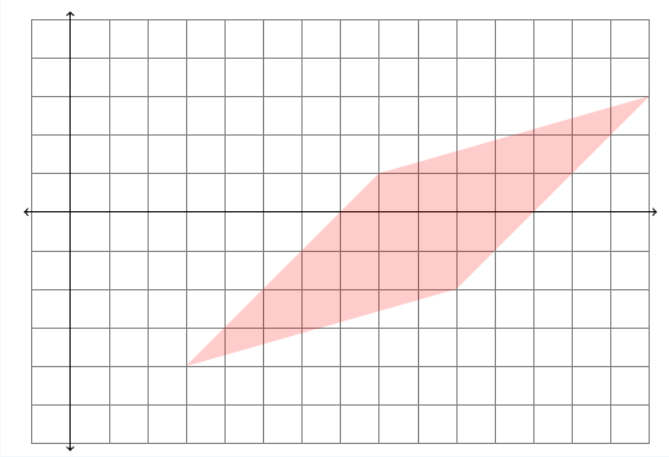
Now take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns: $\begin{vmatrix} 2 & 3 \\ 1 & 5 \end{vmatrix} = 2 \cdot 5 - 3 \cdot 1 = 7$

So the area is $|7| = 7$

Exercise 2.7.1c

Find the area of the parallelogram with vertices $(3,-4),(8,1),(10,-2)$, and $(15,3)$.

Answer



First we need the vectors connecting these vertices:

$\vec{a} = \begin{bmatrix} 8-3 \\ 1-(-4) \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 10-3 \\ -2-(-4) \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$

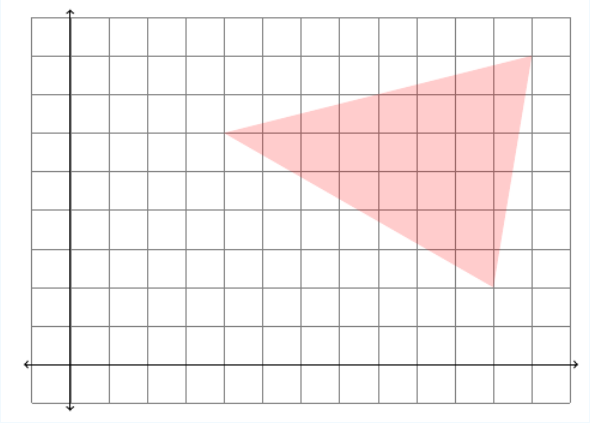
Now take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns: $\begin{vmatrix} 5 & 7 \\ 5 & 2 \end{vmatrix} = 5 \cdot 2 - 5 \cdot 7 = -25$

So the area is $|-25| = 25$

Exercise 2.7.1d

Find the area of the triangle with vertices $(4,6),(11,2)$, and $(12,8)$.

Answer



First we need the vectors connecting these vertices:

$\vec{a} = \begin{bmatrix} 11-4 \\ 2-6 \end{bmatrix} = \begin{bmatrix} 7 \\ -4 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 12-4 \\ 8-6 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$

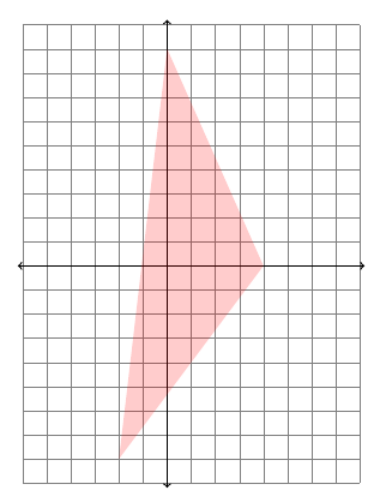
Now take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns: $\begin{vmatrix} 7 & 8 \\ -4 & 2 \end{vmatrix} = 7 \cdot 2 - (-4) \cdot 8 = 46$

So the area is $\frac{1}{2}|46| = 23$

Exercise 2.7.1e

Find the area of the triangle with vertices $(-2,-8),(0,9)$, and $(4,0)$.

Answer



First we need the vectors connecting these vertices:

$\vec{a} = \begin{bmatrix} 0-(-2) \\ 9-(-8) \end{bmatrix} = \begin{bmatrix} 2 \\ 17 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 4-(-2) \\ 0-(-8) \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$

Now take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns: $\begin{vmatrix} 2 & 6 \\ 17 & 8 \end{vmatrix} = 2 \cdot 8 - 17 \cdot 6 = -86$

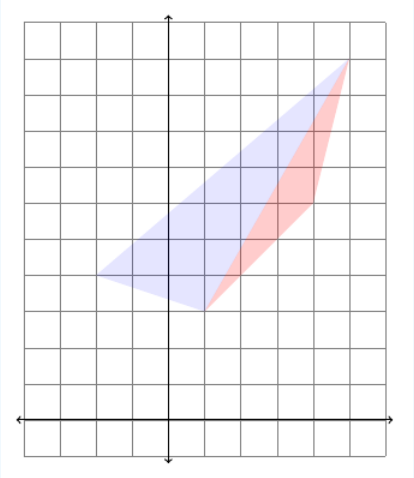
So the area is $\frac{1}{2}|-86| = 43$

Exercise 2.7.1f

Find the area of the quadrilateral with vertices $(1,3),(-2,4),(4,6),(5,10)$.

Answer

Note that this is **not a parallelogram**. Instead think of it as two triangles added together!



Triangle 1 with vertices $(1,3),(-2,4),(5,10)$ in blue:

$\vec{a} = \begin{bmatrix} -2-1 \\ 4-3 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 5-1 \\ 10-3 \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$

Now take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns: $\begin{vmatrix} -3 & 4 \\ 1 & 7 \end{vmatrix} = -3 \cdot 7 - 4 \cdot 1 = -25$

So the area of this triangle is $\frac{1}{2}|-25| = \frac{25}{2}$

Triangle 2 with vertices $(1,3), (4,6), (5,10)$ in red:

$\vec{c} = \begin{bmatrix} 4-1 \\ 6-3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ and $\vec{d} = \begin{bmatrix} 5-1 \\ 10-3 \end{bmatrix} = \begin{bmatrix} 4 \\ 7 \end{bmatrix}$

Now take the determinant of the matrix with $\vec{c}$ and $\vec{d}$ as columns: $\begin{vmatrix} 3 & 4 \\ 3 & 7 \end{vmatrix} = 3 \cdot 7 - 3 \cdot 4 = 9$

So the area of this triangle is $\frac{1}{2}|9| = \frac{9}{2}$

Adding these together gives us a total area of $\frac{9}{2} + \frac{25}{2} = 17$

Exercise 2.7.1g

Find the volume of the parallelepiped with adjacent sides $\vec{a} = \begin{bmatrix} 2 \\ 6 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 3 \\ -2 \\ 3 \end{bmatrix}$ and $\vec{c} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix}$

Answer

We take the determinant of the matrix with $\vec{a}, \vec{b}$, and $\vec{c}$ as columns by expanding along the first row:

$$\begin{vmatrix} 2 & 3 & 1 \\ 6 & -2 & 1 \\ 1 & 3 & 5 \end{vmatrix} = 2\begin{vmatrix} -2 & 1 \\ 3 & 5 \end{vmatrix} - 3\begin{vmatrix} 6 & 1 \\ 1 & 5 \end{vmatrix} + 1\begin{vmatrix} 6 & -2 \\ 1 & 3 \end{vmatrix} = -93$$

So the volume is $|-93| = 93$

Exercise 2.7.1h

Find the volume of the parallelepiped with adjacent sides $\vec{a} = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$ and $\vec{c} = \begin{bmatrix} 3 \\ 6 \\ 2 \end{bmatrix}$

Answer

We take the determinant of the matrix with $\vec{a}, \vec{b}$, and $\vec{c}$ as columns by expanding along the first row:

$$\begin{vmatrix} -1 & 2 & 3 \\ 2 & -1 & 6 \\ 2 & -2 & 2 \end{vmatrix} = -1\begin{vmatrix} -1 & 6 \\ -2 & 2 \end{vmatrix} - 2\begin{vmatrix} 2 & 6 \\ 2 & 2 \end{vmatrix} + 3\begin{vmatrix} 2 & -1 \\ 2 & -2 \end{vmatrix} = 0$$

So the volume is $|0| = 0$ since these vectors are *coplanar*.

Exercise 2.7.2a

Find the value(s) of h which make the parallelogram with edges $\vec{a} = \begin{bmatrix} 1 \\ h \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ have an area of exactly 10.

Answer

We take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns:

$$\begin{vmatrix} 1 & 4 \\ h & 2 \end{vmatrix} = 1 \cdot 2 - h \cdot 4 = 2 - 4h$$

Now set the area of this equal to 10 and solve:

$$|2-4h| = 10 \longrightarrow \begin{cases} 2-4h = 10 \\ 2-4h = -10 \end{cases} \longrightarrow \begin{cases} h = -2 \\ h = 3 \end{cases}$$

So our answers are $h = -2, 3$

Exercise 2.7.2b

Find the value(s) of h which make the parallelogram with edges $\vec{a} = \begin{bmatrix} h \\ h+1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 2 \\ h-1 \end{bmatrix}$ have an area of exactly 2.

Answer

We take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns:

$$\begin{vmatrix} h & 2 \\ h+1 & h-1 \end{vmatrix} = h \cdot (h-1) - (h+1) \cdot 2 = h^2 - 3h - 2$$

Now set the area of this equal to 2 and solve:

$$|h^2 - 3h - 2| = 2 \longrightarrow \begin{cases} h^2 - 3h - 2 = 2 \\ h^2 - 3h - 2 = -2 \end{cases} \longrightarrow \begin{cases} h = -1, 4 \\ h = 0, 3 \end{cases}$$

So our answers are $h = -1, 0, 3, 4$

Exercise 2.7.2c

Find the value(s) of h which make the triangle with vertices $(0,0), (4,5)$, and $(2,h)$ have an area of exactly 3.

Answer

First we need the vectors connecting these vertices:

$$\vec{a} = \begin{bmatrix} 4-0 \\ 5-0 \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 2-0 \\ h-0 \end{bmatrix} = \begin{bmatrix} 2 \\ h \end{bmatrix}$$

Now take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns:

$$\begin{vmatrix} 4 & 2 \\ 5 & h \end{vmatrix} = 4 \cdot h - 2 \cdot 5 = 4h - 10$$

Now set the area of this equal to 3 and solve:

$$\tfrac{1}{2}|4h - 10| = 3 \longrightarrow \begin{cases} 2h - 5 = 3 \\ 2h - 5 = -3 \end{cases} \longrightarrow \begin{cases} h = 4 \\ h = 1 \end{cases}$$

So our answers are $h = 1, 4$

Exercise 2.7.2d

Find the value(s) of h which make the triangle with vertices $(-3,2), (h,h)$, and $(5,-1)$ have an area of exactly 6.

Answer

First we need the vectors connecting these vertices:

$$\vec{a} = \begin{bmatrix} h-(-3) \\ h-2 \end{bmatrix} = \begin{bmatrix} h+3 \\ h-2 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 5-(-3) \\ -1-2 \end{bmatrix} = \begin{bmatrix} 8 \\ -3 \end{bmatrix}$$

Now take the determinant of the matrix with $\vec{a}$ and $\vec{b}$ as columns: $\begin{vmatrix} h+3 & 8 \\ h-2 & -3 \end{vmatrix} = (h+3) \cdot (-3) - (h-2) \cdot 8 = 7 - 11h$

Now set the area of this equal to 6 and solve:

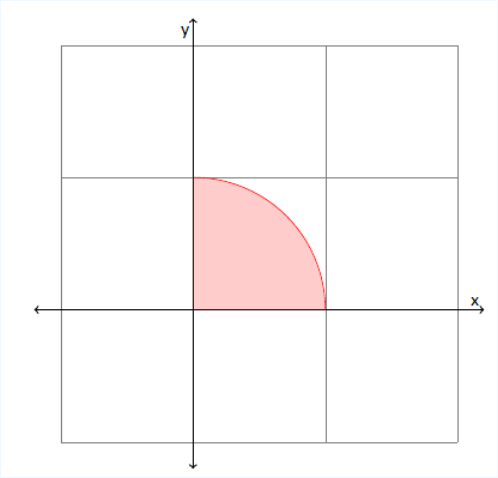
$$\tfrac{1}{2}|7 - 11h| = 6 \longrightarrow \begin{cases} 7 - 11h = 12 \\ 7 - 11h = -12 \end{cases} \longrightarrow \begin{cases} h = -\frac{5}{11} \\ h = \frac{19}{11} \end{cases}$$

So our answers are $h = -\frac{5}{11}, \frac{19}{11}$

Exercise 2.7.3a

Find the area of the image of the region $S = \{x^2 + y^2 \leq 1, x \geq 0, y \geq 0\}$ under the transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x + y \\ 2x - 3y \end{bmatrix}$

Answer



The set S is clearly a quarter circle, so $\text{area of } S = \frac{1}{4}\pi \cdot 1^2 = \frac{\pi}{4}$

The linear transformation T has matrix representation $A = \begin{bmatrix} 2 & 1 \\ 2 & -3 \end{bmatrix}$

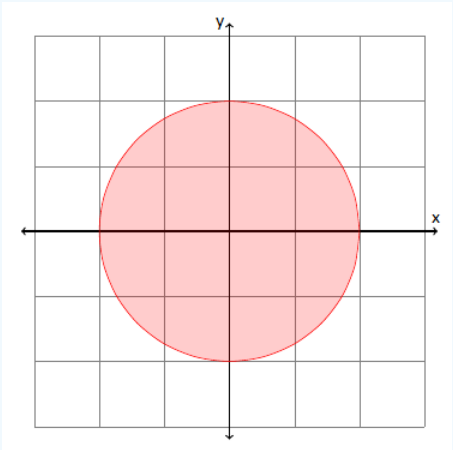
We calculate $\det(A) = 2 \cdot (-3) - 2 \cdot 1 = -8$

So our answer is $\text{area of } T(S) = |\det(A)| \cdot \text{area of } S = |-8| \cdot \frac{\pi}{4} = 2\pi$

Exercise 2.7.3b

Find the area of the image of the region $S = \{x^2 + y^2 \leq 4\}$ under the transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 4x + 9y \\ 2x + 7y \end{bmatrix}$

Answer



The set S is a full circle, so $\text{area of } S = \pi \cdot 2^2 = 4\pi$

https://math.libretexts.org/@go/page/186427

The linear transformation T has matrix representation $A = \begin{bmatrix} 4 & 9 \\ 2 & 7 \end{bmatrix}$

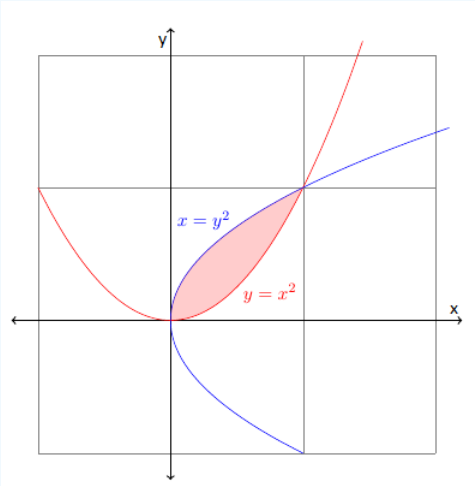
We calculate $\det(A) = 4 \cdot 7 - 2 \cdot 9 = 10$

So our answer is area of $T(S) = |\det(A)| \cdot$ area of $S = |10| \cdot 4\pi = 40\pi$

Exercise 2.7.3c

Find the area of the image of the region $S = \{y \geq x^2, x \geq y^2\}$ under the transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x + y \\ 2x + 4y \end{bmatrix}$

Answer



The set S is between two parabolas, so we use an integral to find its area:

area of $S = \int_0^1 \sqrt{x} - x^2 dx = \frac{1}{3}$

The linear transformation T has matrix representation $A = \begin{bmatrix} 2 & 1 \\ 2 & 4 \end{bmatrix}$

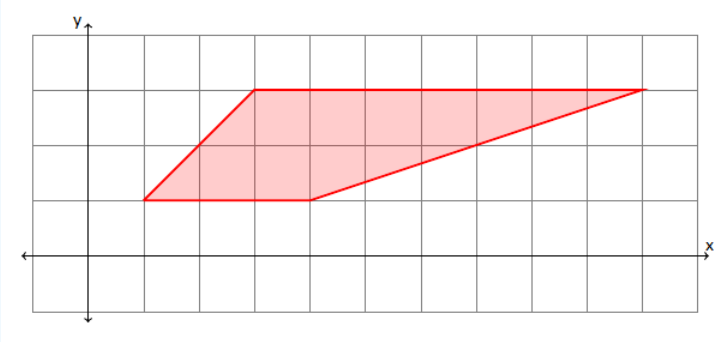
We calculate $\det(A) = 2 \cdot 4 - 1 \cdot 2 = 6$

So our answer is area of $T(S) = |\det(A)| \cdot$ area of $S = |6| \cdot \frac{1}{3} = 2$

Exercise 2.7.3d

Find the area of the image of the region $S = \{1 \leq y \leq 3, y \leq x \leq 3y + 1\}$ under the transformation $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -x + 2y \\ 5x + 8y \end{bmatrix}$

Answer



The set S is a trapezoid with vertices at $(1,1),(3,3),(10,3),(4,1)$, which has parallel bases of length $10-3=7$ and $4-1=3$ and vertical height of $3-1=2$ so we have area of $S=\frac{7+3}{2}\cdot 2=10$

The linear transformation T has matrix representation $A=\begin{bmatrix}-1 & 2\\ 5 & 8\end{bmatrix}$

We calculate $\det(A)=-1\cdot 8-5\cdot 2=-18$

So our answer is area of $T(S)=|\det(A)|\cdot$ area of $S=|-18|\cdot 10=180$

Exercise 2.7.4a

Find the area of the region $S=\{(x+2y)^2+(3x+y)^2\leq 100\}$

Answer

This ellipse looks nicer with the substitution: $u=x+2y$ and $v=3x+y$

Which maps it to the set $S'=\{u^2+v^2\leq 100\}$, clearly the area of $S'=\pi\cdot 10^2=100\pi$

The transformation that this substitution represents is $T\left(\begin{bmatrix}x\\ y\end{bmatrix}\right)=\begin{bmatrix}x+2y\\ 3x+y\end{bmatrix}=\begin{bmatrix}1 & 2\\ 3 & 1\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$

We find the determinant of this matrix: $\det(A)=1\cdot 1-2\cdot 3=-5$

We know $100\pi=$ area of $S'=|\det(A)|\cdot$ area of $S=5\cdot$ area of S

Therefore we must have area of $S=20\pi$

Exercise 2.7.4b

Find the area of the region $S=\{(\frac{1}{3}x-\frac{1}{2}y)^2+(4x+3y)^2\leq 12, y\geq\frac{2}{3}x\}$

Answer

This ellipse looks nicer with the substitution: $u=\frac{1}{3}x-\frac{1}{2}y$ and $v=4x+3y$

Which maps it to the set $S'=\{u^2+v^2\leq 12, u\leq 0\}$, clearly the area of $S'=\frac{1}{2}\pi\cdot(\sqrt{12})^2=6\pi$

The transformation that this substitution represents is $T\left(\begin{bmatrix}x\\ y\end{bmatrix}\right)=\begin{bmatrix}\frac{1}{3}x-\frac{1}{2}y\\ 4x+3y\end{bmatrix}=\begin{bmatrix}\frac{1}{3} & -\frac{1}{2}\\ 4 & 3\end{bmatrix}\begin{bmatrix}x\\ y\end{bmatrix}$

We find the determinant of this matrix: $\det(A)=\frac{1}{3}\cdot 3-4\cdot(-\frac{1}{2})=3$

We know $6\pi=$ area of $S'=|\det(A)|\cdot$ area of $S=3\cdot$ area of S

Therefore we must have area of $S=2\pi$

Exercise 2.7.4c

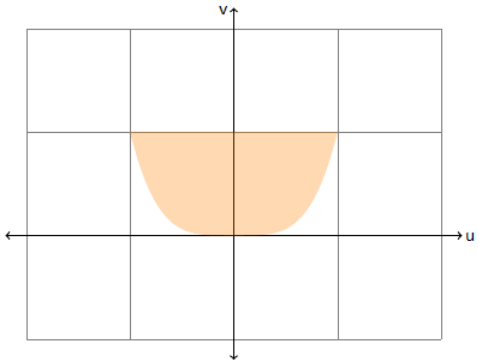
Find the area of the region S between the curves $x+y=(2x-3y)^4$ and $y=1-x$

Answer

These curves look nicer with the substitution: $u=2x-3y$ and $v=x+y$

This makes them into $v=u^4$ and $v=1$

Which maps the set $S'=\{-1\leq u\leq 1, u^4\leq v\leq 1\}$, pictured below.



We can use integrals to find the $\text{area of } S' = \int_{-1}^{1} 1 - u^4 du = \frac{8}{5}$

The transformation that this substitution represents is $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 2x - 3y \\ x + y \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

We find the determinant of this matrix: $\det(A) = 2 \cdot 1 - 1 \cdot (-3) = 5$

We know $\frac{8}{5} = \text{area of } S' = |\det(A)| \cdot \text{ area of } S = 5 \cdot \text{ area of } S$

Therefore we must have $\text{area of } S = \frac{8}{25}$

2.7: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

3: Vector Geometry

3.1: Change of Basis
- 3.1: Coordinatization
- 3.1: Similar Matrices
- 3.1: Practice Set

3.2: Eigenvalues and Eigenvectors
- 3.2: Eigenvalues and Eigenvectors Definitions
- 3.2: Eigenspaces
- 3.2: Characteristic Polynomial
- 3.2: Practice Set

3.3: Diagonalization and Geometry of Eigenvalues
- 3.3: Diagonalization
- 3.3: Geometry of Diagonalizable Matrices
- 3.3: Practice Set

3.4: Complex Eigenvalues
- 3.4: Complex Eigenvalues and Eigenvectors
- 3.4: Geometry of Matrices with Complex Eigenvalues
- 3.4: Practice Set

3.5: Dot Products and Orthogonal Complements
- 3.5: Dot Products and Orthogonality
- 3.5: Orthogonal Complements
- 3.5: Practice Set

3.6: Orthogonal Projection and Least Squares
- 3.6: Orthogonal Decomposition
- 3.6: Least Squares Solutions
- 3.6: Practice Set

3.7: Orthogonal Sets
- 3.7: Orthogonal Sets and Projection
- 3.7: Gram-Schmidt Process
- 3.7: Practice Set

This page titled 3: Vector Geometry is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

3.1: Change of Basis

Learning Objectives

1. Learn to view a basis as a coordinate system on a subspace.
2. Find a matrix similar to a given matrix A with respect to a change in basis.
3. *Recipes:* compute the $\mathcal{B}$-coordinates of a vector, compute the usual coordinates of a vector from its $\mathcal{B}$-coordinates
4. *Picture:* the $\mathcal{B}$-coordinates of a vector using its location on a nonstandard coordinate grid.
5. *Vocabulary words:* $\mathcal{B}$-***coordinates, similar matrices***

In this section, we define ways to describe vectors in $\mathbb{R}^n$ in terms of their linear combinations with respect to a given basis. This allows us to better understand transformations based on how they act on a basis.

3.1: Change of Basis is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **2.8: Bases as Coordinate Systems** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.1: Coordinatization

In this section, we interpret a basis of a subspace V as a *coordinate system* on V, and we learn how to write a vector in V in that coordinate system.

Fact 3.1.1

If $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m\}$ is a basis for a subspace V, then any vector $\vec{x}$ in V can be written as a linear combination

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$$

in exactly one way.

Proof

Recall that to say $\mathcal{B}$ is a *basis* for V means that $\mathcal{B}$ spans V and $\mathcal{B}$ is linearly independent. Since $\mathcal{B}$ spans V, we can write any $\vec{x}$ in V as a linear combination of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$. For uniqueness, suppose that we had two such expressions:

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$$
$$\vec{x} = c_1'\vec{v}_1 + c_2'\vec{v}_2 + \cdots + c_m'\vec{v}_m$$

Subtracting the first equation from the second yields

$$\vec{0} = \vec{x} - \vec{x} = (c_1' - c_1)\vec{v}_1 + (c_2' - c_2)\vec{v}_2 + \cdots + (c_m' - c_m)\vec{v}_m$$

Since $\mathcal{B}$ is linearly independent, the only solution to the above equation is the trivial solution: all the coefficients must be zero. It follows that $c_i - c_i' = 0$ for all i, which proves that $c_1 = c_1'$, $c_2 = c_2'$, $\ldots$, $c_m = c_m'$.

Illustration 3.1.1

Consider the standard basis of $\mathbb{R}^3$

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

According to the above fact, every vector in $\mathbb{R}^3$ can be written as a linear combination of $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$, with unique coefficients. For example,

$$\vec{v} = \begin{bmatrix} 3 \\ 5 \\ -2 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 5\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - 2\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 3\vec{e}_1 + 5\vec{e}_2 - 2\vec{e}_3$$

In this case, *the coordinates of $\vec{v}$ are exactly the coefficients of $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$.*

What exactly are coordinates, anyway? One way to think of coordinates is that they give directions for how to get to a certain point from the origin. In the above example, the linear combination $3\vec{e}_1 + 5\vec{e}_2 - 2\vec{e}_3$ can be thought of as the following list of instructions: start at the origin, travel 3 units east, then travel 5 units north, then 2 units down into the ground.

Definition: $\mathcal{B}$-Coordinate Vector

Let $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m\}$ be a basis of a subspace V, and let

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$$

be a vector in V. The coefficients $c_1, c_2, \cdots, c_m$ are the ***coordinates of $\vec{x}$ with respect to $\mathcal{B}$***. The ***$\mathcal{B}$-coordinate vector of $\vec{x}$*** is the vector:

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \quad \text{in } \mathbb{R}^m$$

If we change the basis, then we can still give instructions for how to get to the point $(3,\ 5,\ -2)$, but the instructions will be different. Say for example we take the basis

$$\vec{v}_1 = \vec{e}_1 + \vec{e}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

We can write $(3,\ 5,\ -2)$ in this basis as $3\vec{v}_1 + 2\vec{v}_2 - 2\vec{v}_3$. In other words: start at the origin, travel northeast 3 times as far as $\vec{v}_1$, then 2 units north, then 2 units down. In this situation, we can say that "3 is the $\vec{v}_1$-coordinate of $(3,\ 5,\ -2)$, 2 is the $\vec{v}_2$-coordinate of $(3,\ 5,\ -2)$, and -2 is the $\vec{v}_3$-coordinate of $(3,\ 5,\ -2)$."

Note 3.1.1

The above definition gives a way of using $\mathbb{R}_m$ to *label* the points of a subspace of dimension m: a point is simply labeled by its $\mathcal{B}$-coordinate vector. For instance, if we choose a basis for a plane, we can label the points of that plane with the points of $\mathbb{R}^2$.

Example 3.1.1: A Nonstandard Coordinate System of $\mathbb{R}^2$

Define

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathcal{B} = \{\vec{v}_1,\ \vec{v}_2\}$$

(a) Verify that $\mathcal{B}$ is a basis for $\mathbb{R}^2$.

(b) If $[\vec{w}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, then what is $\vec{w}$?

(c) Find the $\mathcal{B}$-coordinates of $\vec{v} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

Solution

(a) We know that any two linearly independent vectors form a basis for $\mathbb{R}^2$. Clearly $\vec{v}_1, \vec{v}_2$ are not multiples of each other, so they are linearly independent.

(b) To say $[\vec{w}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ means that 1 is the $\vec{v}_1$-coordinate of w, and that 2 is the $\vec{v}_2$-coordinate:

$$\vec{w} = \vec{v}_1 + 2\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \boxed{\begin{bmatrix} 3 \\ -1 \end{bmatrix}}$$

(c) We have to solve the vector equation $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2$ in the unknowns $c_1,\ c_2$. We form an augmented matrix and row reduce:

$$\left[\begin{array}{cc|c} 1 & 1 & 5 \\ 1 & -1 & 3 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & 4 \\ 0 & 1 & 1 \end{array}\right]$$

We have $c_1 = 4$ and $c_2 = 1$, so $\vec{v} = 4\vec{v}_1 + \vec{v}_2$ and $\boxed{[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}}$

In the following picture, we indicate the coordinate system defined by $\mathcal{B}$ by drawing lines parallel to the "$\vec{v}_1$-axis" and "$\vec{v}_2$-axis". Using this grid it is easy to see that the $\mathcal{B}$-coordinates of $\vec{v}$ are $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$, and that the $\mathcal{B}$-coordinates of $\vec{w}$ are $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

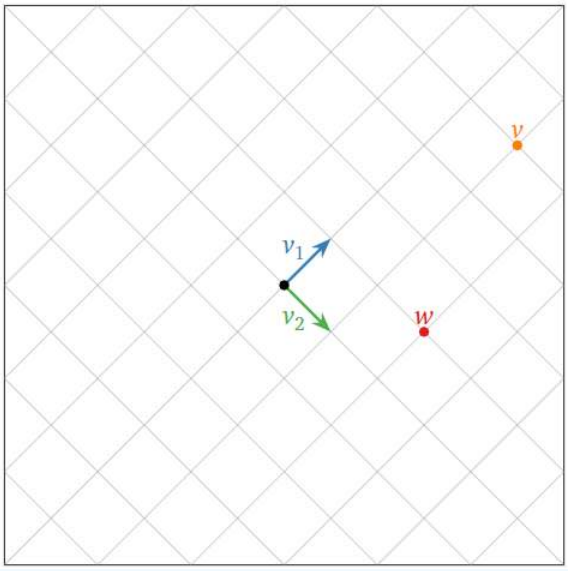


Figure 3.1.1

This picture could be the grid of streets in Palo Alto, California. Residents of Palo Alto refer to northwest as “north” and to northeast as “east”. There is a reason for this: the old road to San Francisco is called El Camino Real, and that road runs from the southeast to the northwest in Palo Alto. So when a Palo Alto resident says “go south two blocks and one block east”, they are giving directions from the origin to the Whole Foods at w.

Illustration 3.1.2

Let

$$\vec{v}_1 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

These form a basis $\mathcal{B}$ for a plane $V = \text{Span}\{\vec{v}_1,\ \vec{v}_2\}$ in $\mathbb{R}^3$. We indicate the coordinate system defined by $\mathcal{B}$ by drawing lines parallel to the “$\vec{v}_1$-axis” and “$\vec{v}_2$-axis”:

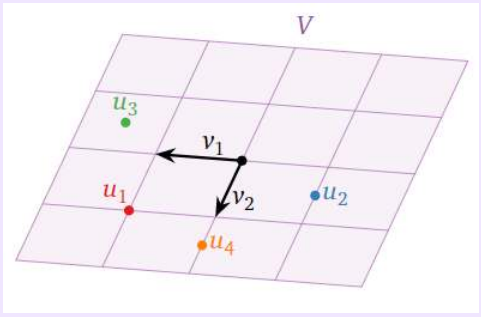


Figure 3.1.2

We can see from the picture that the $\vec{v}_1$-coordinate of $\vec{u}_1$ is equal to 1, as is the $\vec{v}_2$-coordinate, so $[\vec{u}_1]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Similarly, we have

$$[\vec{u}_2]_{\mathcal{B}} = \begin{bmatrix} -1 \\ \frac{1}{2} \end{bmatrix} \quad [\vec{u}_3]_{\mathcal{B}} = \begin{bmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{bmatrix} \quad [\vec{u}_4]_{\mathcal{B}} = \begin{bmatrix} 0 \\ \frac{3}{2} \end{bmatrix}$$

Recipes: $\mathcal{B}$-Coordinates

If $\mathcal{B} = \{\vec{v}_1,\ \vec{v}_2, \cdots, \vec{v}_m\}$ is a basis for a subspace V and $\vec{x}$ is in V, then

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \quad \text{means} \quad \vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$$

Finding the $\mathcal{B}$-coordinates of $\vec{x}$ means solving the vector equation

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$$

in the unknowns $c_1,\ c_2, \cdots, c_m$. This generally means row reducing the augmented matrix

$$\left[\begin{array}{cccc|c} | & | & & | & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m & \vec{x} \\ | & | & & | & | \end{array}\right]$$

If $\mathcal{B}$ is a basis for $\mathbb{R}^m$, then the matrix $Q_{\mathcal{B}} = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix}$ is square and invertible. In this case, the change-of-basis formulas become:

$$Q_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} = \vec{x} \quad \text{and} \quad Q_{\mathcal{B}}^{-1}\vec{x} = [\vec{x}]_{\mathcal{B}}$$

Here the matrix $Q_{\mathcal{B}}$ is called a **change-of-basis matrix** from $\mathcal{B}$ to the standard basis.

Its inverse $Q_{\mathcal{B}}^{-1}$ is also a change-of-basis matrix, but from the standard basis to $\mathcal{B}$.

This means that the map $T_{\mathcal{B}} : \mathbb{R}^m \to \mathbb{R}^m$ which sends each vector $\vec{x}$ to its $\mathcal{B}$-coordinate vector $[\vec{x}]_{\mathcal{B}}$ is a matrix transformation, and therefore must be linear.

Illustration 3.1.3

Consider again the problem in Example 3.1.1 with

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathcal{B} = \{\vec{v}_1,\ \vec{v}_2\}$$

Part (b) asked if $[\vec{w}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, then what is $\vec{w}$?

The change-of-basis matrix from $\mathcal{B}$ to the standard basis is

$$Q_{\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Therefore, we can find $\vec{w}$ by calculating:

$$\vec{w} = Q_{\mathcal{B}}[\vec{w}]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$$

We can see that this matches our answer from earlier.

Part (c) asked to find the $\mathcal{B}$-coordinates of $\vec{v} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$

(c) The change-of-basis matrix from the standard basis to $\mathcal{B}$ is

$$Q_{\mathcal{B}}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

Therefore, we can find $[\vec{v}]_{\mathcal{B}}$ by calculating:

$$[\vec{v}]_{\mathcal{B}} = Q_{\mathcal{B}}^{-1}\vec{v} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 5 \\ 3 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

Again, this matches our answer from earlier.

Example 3.1.2: Using the Change-of-Basis Matrix

Let $\mathcal{B} = \left\{\vec{b}_1, \vec{b}_2\right\} = \left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix}\right\}$

(a) Find a vector $\vec{x}$ such that $[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$

(b) Find $[\vec{y}]_{\mathcal{B}}$ for the vector $\vec{y} = \begin{bmatrix} 7 \\ 4 \end{bmatrix}$

Solution

(a) The change-of-basis matrix from $\mathcal{B}$ to the standard basis is

$$Q_{\mathcal{B}} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$$

Therefore, we can find $\vec{x}$ by calculating:

$$\vec{x} = Q_{\mathcal{B}}[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 4 \\ -1 \end{bmatrix} = \begin{bmatrix} 8-3 \\ 4-2 \end{bmatrix} = \boxed{\begin{bmatrix} 5 \\ 2 \end{bmatrix}}$$

We can verify: $4\vec{b}_1 + (-1)\vec{b}_2 = 4\begin{bmatrix} 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 8-3 \\ 4-2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$

(b) The change-of-basis matrix from the standard basis to $\mathcal{B}$ is

$$Q_{\mathcal{B}}^{-1} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$$

Therefore, we can find $[\vec{y}]_{\mathcal{B}}$ by calculating:

$$[\vec{y}]_{\mathcal{B}} = Q_{\mathcal{B}}^{-1}\vec{y} = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 7 \\ 4 \end{bmatrix} = \begin{bmatrix} 14-12 \\ -7+8 \end{bmatrix} = \boxed{\begin{bmatrix} 2 \\ 1 \end{bmatrix}}$$

Again we verify: $2\vec{b}_1 + 1\vec{b}_2 = 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 4+3 \\ 2+2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \end{bmatrix}$

3.1: Coordinatization is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **2.8: Bases as Coordinate Systems** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.1: Similar Matrices

Consider the linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} b \\ a \end{bmatrix}$

This transformation has standard matrix representation

$$[T] = \begin{bmatrix} | & | \\ T(\vec{e}_1) & T(\vec{e}_2) \\ | & | \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Now, let us see how T treats the vectors in the basis $\mathcal{B} = \left\{\vec{b}_1, \vec{b}_1\right\}$ where $\vec{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{b}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$:

$$T(\vec{b}_1) = T\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad T(\vec{b}_2) = T\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Now, let's change our perspective on this transformation by coordinatizing each input and output above with respect to $\mathcal{B}$:

$$\left[\vec{b}_1\right]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \left[\vec{b}_2\right]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\left[T\left(\vec{b}_1\right)\right]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \left[T\left(\vec{b}_2\right)\right]_{\mathcal{B}} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

We know there is a different linear transformation U which mimics the results of this "coordinatized version" of T:

$$U\left(\left[\vec{b}_1\right]_{\mathcal{B}}\right) = U\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \left[T\left(\vec{b}_1\right)\right]_{\mathcal{B}}$$

$$U\left(\left[\vec{b}_2\right]_{\mathcal{B}}\right) = U\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \left[T\left(\vec{b}_2\right)\right]_{\mathcal{B}}$$

It is straightforward to see that the matrix representation of U is

$$[U] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

We know that T "switches" the x and y component (like you would do for finding the inverse of a function), which geometrically means it reflects a vector across the line $y = x$.

Figure 3.1.3: The transformation T reflects each vector across the line $y = x$

On the other hand, the transformation U is defined by what T does to vectors, but with the "addresses" of each point renamed with respect to a new basis.

Clearly there is some relationship between the matrices $[T]$ and $[U]$, and you may have (correctly) guessed that it involves the change of basis matrix

$$Q_{\mathcal{B}} = \begin{bmatrix} | & | \\ \vec{b}_1 & \vec{b}_2 \\ | & | \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

The previous example worked nicely because the basis vectors were chosen cleverly. But the same method works for any basis and any target matrix. Here's another example with a different basis and a different diagonal matrix.

Illustration 3.1.4

Let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix} \right\}$ be a basis of $\mathbb{R}^2$. Can we find a matrix A which satisfies the equation below for all possible vectors $\vec{v} \in \mathbb{R}^2$?

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} [\vec{v}]_{\mathcal{B}} = [A\vec{v}]_{\mathcal{B}}$$

We define the matrix $B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

To find the columns of A we just need to calculate $A\vec{e}_1$ and $A\vec{e}_2$.

First we plug in $\vec{v} = \vec{e}_1$ to our equation, which turns into

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} [\vec{e}_1]_{\mathcal{B}} = [A\vec{e}_1]_{\mathcal{B}}$$

Clearly we need to find $[\vec{e}_1]_{\mathcal{B}}$, which we can do by using $Q_{\mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$, or more specifically its inverse $Q_{\mathcal{B}}^{-1} = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix}$

$$[\vec{e}_1]_{\mathcal{B}} = Q_{\mathcal{B}}^{-1}\vec{e}_1 = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -5 \\ 3 \end{bmatrix}$$

Now we can directly compute $B[\vec{e}_1]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -5 \\ 3 \end{bmatrix} = \begin{bmatrix} -10 \\ 9 \end{bmatrix}$

Plugging this back into our equation above, we see that: $\begin{bmatrix} -10 \\ 9 \end{bmatrix} = [A\vec{e}_1]_{\mathcal{B}}$

We can use $Q_{\mathcal{B}}$ here to find $A\vec{e}_1$:

$$A\vec{e}_1 = Q_{\mathcal{B}}[A\vec{e}_1]_{\mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} -10 \\ 9 \end{bmatrix} = \begin{bmatrix} 8 \\ 15 \end{bmatrix}$$

Now we repeat the process for $\vec{v} = \vec{e}_2$:

$$[\vec{e}_2]_{\mathcal{B}} = Q_{\mathcal{B}}^{-1}\vec{e}_2 = \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$B[\vec{e}_2]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -3 \end{bmatrix} = [A\vec{e}_2]_{\mathcal{B}}$$

$$A\vec{e}_2 = Q_{\mathcal{B}}[A\vec{e}_2]_{\mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 4 \\ -3 \end{bmatrix} = \begin{bmatrix} -2 \\ -3 \end{bmatrix}$$

Putting this together, we have found our matrix A must be $A = \begin{bmatrix} 8 & -2 \\ 15 & -3 \end{bmatrix}$

To demonstrate it is the correct answer, note that for our basis $\mathcal{B} = \{\vec{b}_1, \vec{b}_2\}$ we have:

$$A\vec{b}_1 = \begin{bmatrix} 8 & -2 \\ 15 & -3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix} = 2\vec{b}_1 + 0\vec{b}_2$$

$$A\vec{b}_2 = \begin{bmatrix} 8 & -2 \\ 15 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 6 \\ 15 \end{bmatrix} = 0\vec{b}_1 + 3\vec{b}_2$$

Since A is linear, for any vector $\vec{v} = \alpha\vec{b}_1 + \beta\vec{b}_2$, (meaning $[\vec{v}]_{\mathcal{B}} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$), we will have

$$\begin{aligned} A\vec{v} &= A(\alpha\vec{b}_1 + \beta\vec{b}_2) = \alpha A\vec{b}_1 + \beta A\vec{b}_2 \\ &= \alpha(2\vec{b}_1 + 0\vec{b}_2) + \beta(0\vec{b}_1 + 3\vec{b}_2) \\ &= (2\alpha)\vec{b}_1 + (3\beta)\vec{b}_2 \end{aligned}$$

Thus, $[A\vec{v}]_{\mathcal{B}} = \begin{bmatrix} 2\alpha \\ 3\beta \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} [\vec{v}]_{\mathcal{B}}$ as desired.

The Illustration above shows that the matrix $A = \begin{bmatrix} 8 & -2 \\ 15 & -3 \end{bmatrix}$ and the matrix $B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ are related in a particular way.

In fact it is not hard to see that

$$A = \begin{bmatrix} 8 & -2 \\ 15 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -5 & 2 \\ 3 & -1 \end{bmatrix} = Q_{\mathcal{B}} B Q_{\mathcal{B}}^{-1}$$

That is to say that B does to $[\vec{v}]_{\mathcal{B}}$ the same action that A does to $\vec{v}$.

In fact, for any invertible matrix $P = \begin{bmatrix} | & | & & | \\ \vec{b}_1 & \vec{b}_2 & \cdots & \vec{b}_m \\ | & | & & | \end{bmatrix}$, we can consider the set $\mathcal{B} = \{\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_m\}$. We know that this set must be a basis for $\mathbb{R}^m$, and we can consider $P = Q_{\mathcal{B}}$ to be a change of basis matrix.

So any time we have $A = PBP^{-1}$, we know that A and B are related in this same way:

$$[A\vec{v}]_{\mathcal{B}} = B[\vec{v}]_{\mathcal{B}}$$

The work above inspires a new definition.

Definition: Similar Matrices

Two $n \times n$ matrices A and B are ***similar*** if there exists an invertible $n \times n$ matrix P such that $A = PBP^{-1}$.

Illustration 3.1.5

The matrices

$$\begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

are similar because

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

as the reader can verify.

Note 3.1.2

Similarity is unrelated to row equivalence. Any invertible matrix is row equivalent to I, but I is the only matrix similar to I. For instance,

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

are row equivalent but not similar.

As suggested by its name, similarity is what is called an *equivalence relation*. This means that it satisfies the following properties.

Proposition 3.1.1

Let A, B, and C be $n \times n$ matrices.

1. **Reflexivity:** A is similar to itself.
2. **Symmetry:** if A is similar to B, then B is similar to A.
3. **Transitivity:** if A is similar to B and B is similar to C, then A is similar to C.

Proof

1. Taking $P = I = I^{-1}$, we have $A = IAI^{-1}$.
2. Suppose that $A = PBP^{-1}$. Multiplying both sides on the left by P^{-1} and on the right by P gives

$$P^{-1}AP = P^{-1}(PBP^{-1})P = B$$

Since $(P^{-1})^{-1} = P$, we have $B = P^{-1}A(P^{-1})^{-1}$, so that B is similar to A.

3. Suppose that $A = PBP^{-1}$ and $B = QCQ^{-1}$. Substituting for B and remembering that $(PQ)^{-1} = Q^{-1}P^{-1}$, we have

$$A = P(QCQ^{-1})P^{-1} = (PQ)C(PQ)^{-1},$$

which shows that A is similar to C.

Illustration 3.1.6

We saw in Illustration 3.1.5 that the matrices

$$\begin{bmatrix} 8 & -2 \\ 15 & -3 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

are similar. We also saw Illustration 3.1.6 that the matrices

$$\begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

are similar.

It follows that

$$\begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 8 & -2 \\ 15 & -3 \end{bmatrix}$$

are similar to each other.

This can be verified directly by calculating

$$\begin{bmatrix} 4 & 2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 8 & -2 \\ 15 & -3 \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}^{-1}$$

Example 3.1.3

Let $P = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$

Compute $A = PBP^{-1}$ and verify that A and B are similar.

Solution

First compute P^{-1}:

$$P^{-1} = \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix}$$

Next compute BP^{-1}:

$$BP^{-1} = \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 3 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 9-1 & -6+1 \\ 0-2 & 0+2 \end{bmatrix} = \begin{bmatrix} 8 & -5 \\ -2 & 2 \end{bmatrix}$$

Finally compute $A = P(BP^{-1})$:

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

Thus

$$\boxed{A = \begin{bmatrix} 4 & -1 \\ 2 & 1 \end{bmatrix} = PBP^{-1}}$$

By definition we see that A and B are similar.

Similar matrices may look different, but they represent the same linear transformation in different coordinate systems. In the next few sections, we will see how choosing the right basis can simplify a matrix dramatically. This can save us a lot of effort when doing complicated calculations, like computing large powers (or roots) of a matrix.

3.1: Similar Matrices is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **5.3: Similarity** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.1: Practice Set

Exercise 3.1.1

Consider the vectors $\vec{b}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\vec{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ which form the basis $\mathcal{B} = \left\{\vec{b}_1, \vec{b}_2\right\}$

(a) Find the change of basis matrix P which satisfies $P[\vec{x}]_{\mathcal{B}} = \vec{x}$ for all $\vec{x} \in \mathbb{R}^2$

(b) Find the change of basis matrix Q which satisfies $Q\vec{x} = [\vec{x}]_{\mathcal{B}}$ for all $\vec{x} \in \mathbb{R}^2$

(c) Find $\vec{u}$ if $[\vec{u}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$

(d) Find $[\vec{v}]_{\mathcal{B}}$ if $\vec{v} = \begin{bmatrix} 3 \\ -5 \end{bmatrix}$

(e) Is there a non-zero vector $\vec{y}$ which has $[\vec{y}]_{\mathcal{B}} = \vec{y}$? If so, find one.

(f) If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ find the matrix A' satisfying $A'[\vec{x}]_{\mathcal{B}} = [A\vec{x}]_{\mathcal{B}}$

(g) If $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ find the matrix B' satisfying $B[\vec{x}]_{\mathcal{B}} = [B'\vec{x}]_{\mathcal{B}}$

Answer

(a) This matrix must have columns $\vec{b}_1, \vec{b}_2$ so $P = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$

(b) This matrix is the inverse of P in part (a), so $Q = P^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}$

(c) $\vec{u} = P[\vec{u}]_{\mathcal{B}} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ -5 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$

(d) $[\vec{v}]_{\mathcal{B}} = Q\vec{v} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ -5 \end{bmatrix} = \begin{bmatrix} 4 \\ -5 \end{bmatrix}$

(e) This $\vec{y} = \begin{bmatrix} a \\ b \end{bmatrix}$ would satisfy of $P\vec{y} = \vec{y}$:

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2a+b \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \implies \begin{cases} 2a+b = a \\ b = b \end{cases} \implies \begin{cases} a = -b \\ b \text{ free} \end{cases}$$

So we have infinitely many such$\vec{y}$, but they are all in $\text{Span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$

Any non-zero vector $\vec{y}$ in the span above is an acceptable answer.

(f) $A'[\vec{x}]_{\mathcal{B}} = A'Q\vec{x}$ and $[A\vec{x}]_{\mathcal{B}} = QA\vec{x}$. Therefore in order for $A'Q = QA$, we must have

$$A' = QAQ^{-1} = QAP = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a-c & \frac{a+b-c-d}{2} \\ 2c & c+d \end{bmatrix}$$

(g) $B[\vec{x}]_{\mathcal{B}} = BQ\vec{x}$ and $[B'\vec{x}]_{\mathcal{B}} = QB'\vec{x}$. Therefore in order for $BQ = QB'$, we must have

$$B' = Q^{-1}BQ = PBQ = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a+\frac{c}{2} & -a+2b-\frac{c}{2}+d \\ \frac{c}{2} & -\frac{c}{2}+d \end{bmatrix}$$

Exercise 3.1.2

Consider the vectors $\vec{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{b}_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ which form the basis $\mathcal{B} = \left\{\vec{b}_1, \vec{b}_2\right\}$

(a) Find the change of basis matrix P which satisfies $P[\vec{x}]_{\mathcal{B}} = \vec{x}$ for all $\vec{x} \in \mathbb{R}^2$

(b) Find the change of basis matrix Q which satisfies $Q\vec{x} = [\vec{x}]_{\mathcal{B}}$ for all $\vec{x} \in \mathbb{R}^2$

(c) Find $\vec{u}$ if $[\vec{u}]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

(d) Find $[\vec{v}]_{\mathcal{B}}$ if $\vec{v} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

(e) Is there a non-zero vector $\vec{y}$ which has $[\vec{y}]_{\mathcal{B}} = \vec{y}$? If so, find one.

(f) If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ find the matrix A' satisfying $A'[\vec{x}]_{\mathcal{B}} = [A\vec{x}]_{\mathcal{B}}$

(g) If $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ find the matrix B' satisfying $B[\vec{x}]_{\mathcal{B}} = [B'\vec{x}]_{\mathcal{B}}$

Answer

(a) This matrix must have columns $\vec{b}_1, \vec{b}_2$ so $P = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$

(b) This matrix is the inverse of P in part (a), so $Q = P^{-1} = \begin{bmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{bmatrix}$

(c) $\vec{u} = P[\vec{u}]_{\mathcal{B}} = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \end{bmatrix}$

(d) $[\vec{v}]_{\mathcal{B}} = Q\vec{v} = \begin{bmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

(e) This $\vec{y} = \begin{bmatrix} a \\ b \end{bmatrix}$ would satisfy of $P\vec{y} = \vec{y}$:

$$\begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+3b \\ 2a-b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \implies \begin{cases} a+3b = a \\ 2a-b = b \end{cases} \implies \begin{cases} a = 0 \\ b = 0 \end{cases}$$

There are no non-zero vectors $\vec{y}$ that satisfy the requirements.

(f) $A'[\vec{x}]_{\mathcal{B}} = A'Q\vec{x}$ and $[A\vec{x}]_{\mathcal{B}} = QA\vec{x}$. Therefore in order for $A'Q = QA$, we must have

$$A' = QAQ^{-1} = QAP = \begin{bmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} \frac{a+2b+3c+6d}{7} & \frac{3a-b+9c-3d}{7} \\ \frac{2a+4b-c-2d}{7} & \frac{6a-2b-3c+d}{7} \end{bmatrix}$$

(g) $B[\vec{x}]_{\mathcal{B}} = BQ\vec{x}$ and $[B'\vec{x}]_{\mathcal{B}} = QB'\vec{x}$. Therefore in order for $BQ = QB'$, we must have

$$B' = Q^{-1}BQ = PBQ = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{bmatrix} = \begin{bmatrix} \frac{a+2b+3c+6d}{7} & \frac{3a-b+9c-3d}{7} \\ \frac{2a+4b-c-2d}{7} & \frac{6a-2b-3c+d}{7} \end{bmatrix}$$

? Exercise 3.1.3

Consider the vectors $\vec{b}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ which form the basis $\mathcal{B} = \left\{\vec{b}_1, \vec{b}_2\right\}$

(a) Sketch the grid lines corresponding to $\mathcal{B}$ in the plane (lines where the $\vec{b}_1$-coordinate is constant, and lines where the $\vec{b}_2$-coordinate is constant).

(b) Estimate the $\mathcal{B}$-coordinates of the vector $\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ from your sketch.

(c) Compute $[\vec{v}]_{\mathcal{B}}$ exactly for $\vec{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and verify your estimate.

Answer

(a) The purple line below points in the direction of the new "positive x-axis" (which is the same as the old positive x-axis. The blue line points in the direction of the new "positive y-axis" (which is the old line $y = x$, traversed up and rightward)



(b) From the sketch, $\vec{v}$ appears to be at approximately $\vec{b}_1$-coordinate 1 and $\vec{b}_2$-coordinate 2, so $[\vec{v}]_{\mathcal{B}} \approx \begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

(c) $P = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $Q = P^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$. Then $[\vec{v}]_{\mathcal{B}} = Q\vec{v} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. This matches the estimate.

Exercise 3.1.4

Let $A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

(a) Compute $C = P^{-1}AP$.

(b) Is C similar to A? Why or why not?

(c) Find a basis $\mathcal{B}$ such that C is the matrix representation of the same linear transformation as A in $\mathcal{B}$-coordinates.

Answer

(a) $P^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$. Then $C = P^{-1}AP = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix}$

(b) Yes, C is similar to A because $C = P^{-1}AP$ with P invertible.

(c) The basis $\mathcal{B}$ is the set of columns of P, so $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}$

Exercise 3.1.5

Consider the vectors $\vec{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\vec{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, and $\vec{b}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ which form the basis $\mathcal{B} = \left\{ \vec{b}_1, \vec{b}_2, \vec{b}_3 \right\}$ of $\mathbb{R}^3$

(a) Find the change-of-basis matrix P which satisfies $P[\vec{x}]_{\mathcal{B}} = \vec{x}$ for all $\vec{x} \in \mathbb{R}^3$

(b) Find the change-of-basis matrix Q which satisfies $Q\vec{x} = [\vec{x}]_{\mathcal{B}}$ for all $\vec{x} \in \mathbb{R}^3$

(c) Find $[\vec{v}]_{\mathcal{B}}$ if $\vec{v} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

(d) Find $\vec{w}$ if $[\vec{w}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

Answer

(a) $P = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

(b) $Q = P^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

(c) $[\vec{v}]_{\mathcal{B}} = Q\vec{v} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

(d) $\vec{w} = P[\vec{w}]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 1 \end{bmatrix}$

Exercise 3.1.6

Let $\mathcal{B} = \left\{ \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix} \right\}$ and $C = \begin{bmatrix} -5a^3 + 6b^3 & 15(a^3 - b^3) \\ 2(b^3 - a^3) & 6a^3 - 5b^3 \end{bmatrix}$ where a and b are real numbers.

(a) Find a matrix D satisfing $[C\vec{x}]_{\mathcal{B}} = D[\vec{x}]_{\mathcal{B}}$ for all $\vec{x} \in \mathbb{R}^2$

(b) Find a matrix A such that $A^3 = C$.
Hint: If $A = PXP^{-1}$, then $A^3 = PX^3P^{-1}$

Answer

(a) Let $P = \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix}$ be the change-of-basis matrix from $\mathcal{B}$ to the standard basis, so that $P[\vec{x}]_{\mathcal{B}} = \vec{x}$ and $[\vec{x}]_{\mathcal{B}} = P^{-1}\vec{x}$.

We have $[C\vec{x}]_{\mathcal{B}} = P^{-1}C\vec{x}$. We want this to equal $D[\vec{x}]_{\mathcal{B}} = DP^{-1}\vec{x}$ for all $\vec{x}$. Thus $P^{-1}C = DP^{-1}$, so $D = P^{-1}CP$

First compute $P^{-1} = \begin{bmatrix} -1 & 3 \\ 2 & -5 \end{bmatrix}$. Now we compute $D = P^{-1}CP$:

$$D = \begin{bmatrix} -1 & 3 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} -5a^3 + 6b^3 & 15(a^3 - b^3) \\ 2(b^3 - a^3) & 6a^3 - 5b^3 \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 3 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} 5(-5a^3 + 6b^3) + 2(15a^3 - 15b^3) & 3(-5a^3 + 6b^3) + 1(15a^3 - 15b^3) \\ 5(2b^3 - 2a^3) + 2(6a^3 - 5b^3) & 3(2b^3 - 2a^3) + 1(6a^3 - 5b^3) \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 3 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} 5a^3 & 3b^3 \\ 2a^3 & b^3 \end{bmatrix} = \begin{bmatrix} -5a^3 + 6a^3 & -3b^3 + 3b^3 \\ 10a^3 - 10a^3 & 6b^3 - 5b^3 \end{bmatrix} = \begin{bmatrix} a^3 & 0 \\ 0 & b^3 \end{bmatrix}$$

(b) From part (a), we have $C = PDP^{-1}$ with $P = \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} a^3 & 0 \\ 0 & b^3 \end{bmatrix}$.

Using the hint: if $A = PXP^{-1}$, then $A^3 = PX^3P^{-1}$. We want $A^3 = C = PDP^{-1}$. So we need $X^3 = D$

Choose $X = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, since then $X^3 = \begin{bmatrix} a^3 & 0 \\ 0 & b^3 \end{bmatrix} = D$

Therefore,

$$A = PXP^{-1} = \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 2 & -5 \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} -a & 3a \\ 2b & -5b \end{bmatrix}$$

$$= \begin{bmatrix} -5a + 6b & 15a - 15b \\ -2a + 2b & 6a - 5b \end{bmatrix}$$

One can verify that $A^3 = C$ by direct computation (if you have a lot of time)

Exercise 3.1.7

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If $\mathcal{B}$ is a basis of $\mathbb{R}^n$, there is at least one non-zero vector $\vec{v}$ satisfying $[\vec{v}]_{\mathcal{B}} = \vec{v}$.

(b) If A, B, C are all $n \times n$ matrices, and $AC = CB$, then A and B are similar matrices.

(c) If A and B are similar, then they have the same determinant.

(d) If A is similar to B and B is similar to C, then A is similar to C.

(e) A square matrix A is similar to itself.

(f) If A is similar to B and n is a positive integer, then A^n is similar to B^n.

(g) If A is similar to B and A is invertible, then B is invertible.

(h) If $A \neq I$, then A is similar to I.

(i) If P is invertible, then it is a change-of-basis matrix for some basis $\mathcal{B}$ to the standard basis.

(j) If P is invertible, then it is a change-of-basis matrix from the standard basis to some basis $\mathcal{B}$.

Answer

(a) **Sometimes True Sometimes False**

For the standard basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ this holds for all $\vec{v}$. In Exercise 2 above we saw this can be false.

(b) **Sometimes True Sometimes False**

When C is invertible, this is true. But it is definitely not true for an example like C is full of zeros.

(c) **Always True**

If $A = PBP^{-1}$, then $\det(A) = \det(PBP^{-1}) = \det(P)\det(B)\det(P^{-1}) = \det(P)\det(B) \cdot \dfrac{1}{\det(P)} = \det(B)$.

(d) **Always True**

If $A = PBP^{-1}$ and $B = QCQ^{-1}$, then $A = P(QCQ^{-1})P^{-1} = (PQ)C(PQ)^{-1}$, so A and C are similar.

(e) **Always True**

We know I is invertible and $A = IAI^{-1}$.

(f) **Always True**

If $A = PBP^{-1}$, then $A^n = (PBP^{-1})^n = PBP^{-1} \cdot PBP^{-1} \cdot PBP^{-1} = PB^nP^{-1}$.

(g) **Always True**

Since $A = PBP^{-1}$, then $B = P^{-1}AP$ which is a product of invertible matrices, so B is invertible.

(h) **Always False**

Since $PIP^{-1} = PP^{-1} = I$ for every invertible matrix P, the only matrix similar to I is itself.

(i) **Always True**

If $P = \begin{bmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & \cdots & | \end{bmatrix}$, then P is a change of basis matrix from $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ to the standard basis.

(j) **Always True**

If $P^{-1} = \begin{bmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & \cdots & | \end{bmatrix}$, then P is a change of basis matrix from the standard basis to $\mathcal{B} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$.

3.1: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.1: Eigenvalues and Eigenvectors

Learning Objectives

1. Learn the definitions of eigenvector and eigenvalue.
2. Learn to find eigenvectors and eigenvalues geometrically.
3. Learn to decide if a number is an eigenvalue of a matrix, and if so, how to find an associated eigenvector.
4. Find all eigenvalues of a matrix using the characteristic polynomial.
5. *Recipes:* find a basis for the λ-eigenspace, the characteristic polynomial of a 2×2 matrix.
6. *Pictures:* whether or not a vector is an eigenvector, eigenvectors of standard matrix transformations.
7. *Theorem:* the expanded invertible matrix theorem.
8. *Vocabulary word:* ***eigenspace***
9. *Essential vocabulary words:* ***eigenvector***, ***eigenvalue, characteristic polynomial***

In this section, we define eigenvalues and eigenvectors. These form the most important facet of the structure theory of square matrices. As such, eigenvalues and eigenvectors tend to play a key role in the real-life applications of linear algebra.

This page titled 3.1: Eigenvalues and Eigenvectors is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **5.1: Eigenvalues and Eigenvectors** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **5.2: The Characteristic Polynomial** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.2: Eigenvalues and Eigenvectors Definitions

Definition: Eigenvector and Eigenvalue

Let A be an $n \times n$ matrix.

1. An ***eigenvector*** of A is a *nonzero* vector $\vec{v}$ in $\mathbb{R}^n$ such that $A\vec{v} = \lambda\vec{v}$, for some scalar λ.
2. An ***eigenvalue*** of A is a scalar λ such that the equation $A\vec{v} = \lambda\vec{v}$ has a *nontrivial* solution.

If $A\vec{v} = \lambda\vec{v}$ for $\vec{v} \neq \vec{0}$, we say that λ is the ***eigenvalue for*** $\vec{v}$, and that $\vec{v}$ is an ***eigenvector for*** λ.

The German prefix "eigen" roughly translates to "self" or "own". An eigenvector of A is a vector that is taken to a multiple of itself by the matrix transformation $T(\vec{x}) = A\vec{x}$, which perhaps explains the terminology. On the other hand, "eigen" is often translated as "characteristic"; we may think of an eigenvector as describing an intrinsic, or characteristic, property of A.

Note 3.2.1

Eigenvalues and eigenvectors are only for square matrices.

Note 3.2.2

Eigenvectors are *by definition nonzero*. Eigenvalues may be equal to zero.

We do not consider the zero vector to be an eigenvector: since $A\vec{0} = \vec{0} = \lambda\vec{0}$ for *every* scalar λ, the associated eigenvalue would be undefined.

If someone hands you a matrix A and a vector $\vec{v}$, it is easy to check if $\vec{v}$ is an eigenvector of A: simply multiply $\vec{v}$ by A and see if $A\vec{v}$ is a scalar multiple of $\vec{v}$.

Illustration 3.2.1: Verifying eigenvectors

Consider the matrix

$$A = \begin{bmatrix} 2 & 2 \\ -4 & 8 \end{bmatrix} \quad \text{and vectors} \quad \vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \vec{w} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Which are eigenvectors? What are their eigenvalues?

We have

$$A\vec{v} = \begin{bmatrix} 2 & 2 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\vec{v}$$

Hence, $\vec{v}$ is an eigenvector of A, with eigenvalue $\lambda = 4$. On the other hand,

$$A\vec{w} = \begin{bmatrix} 2 & 2 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

which is not a scalar multiple of $\vec{w}$. Hence, $\vec{w}$ is not an eigenvector of A.

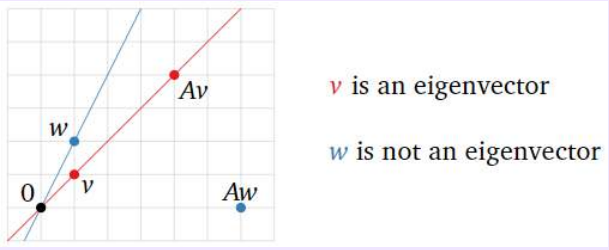


Figure 3.2.1

Illustration 3.2.2: Verifying eigenvectors

Consider the matrix

$$A = \begin{bmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \quad \text{and vectors} \quad \vec{v} = \begin{bmatrix} 16 \\ 4 \\ 1 \end{bmatrix} \quad \vec{w} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}.$$

Which are eigenvectors? What are their eigenvalues?

We have

$$A\vec{v} = \begin{bmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 16 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 32 \\ 8 \\ 2 \end{bmatrix} = 2\vec{v}$$

Hence, $\vec{v}$ is an eigenvector of A, with eigenvalue $\lambda = 2$. On the other hand,

$$A\vec{w} = \begin{bmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 28 \\ 1 \\ 1 \end{bmatrix}$$

which is not a scalar multiple of $\vec{w}$. Hence, $\vec{w}$ is not an eigenvector of A.

Illustration 3.2.3: An eigenvector with eigenvalue 0

Let

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \qquad \vec{v} = \begin{bmatrix} -3 \\ 1 \end{bmatrix}.$$

Is $\vec{v}$ an eigenvector of A? If so, what is its eigenvalue?

The product is

$$A\vec{v} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} -3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0\vec{v}$$

Hence, $\vec{v}$ is an eigenvector with eigenvalue *zero*.

As noted above, an eigen*value* is allowed to be zero, but an eigen*vector* is not.

To say that $A\vec{v} = \lambda\vec{v}$ means that $A\vec{v}$ and $\lambda\vec{v}$ are *collinear with the origin*. So, an eigenvector of A is a nonzero vector $\vec{v}$ such that $A\vec{v}$ and $\vec{v}$ lie on the same line through the origin. In this case, $A\vec{v}$ is a scalar multiple of $\vec{v}$; the eigenvalue is the scaling factor.

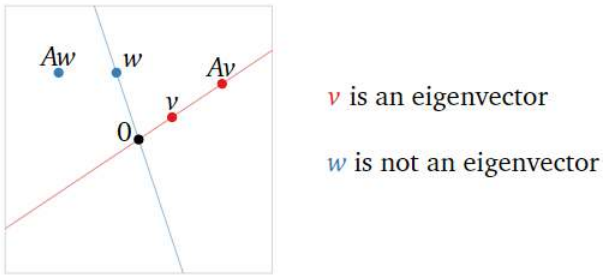


Figure 3.2.2

For matrices that arise as the standard matrix of a linear transformation, it is often best to draw a picture, then find the eigenvectors and eigenvalues geometrically by studying which vectors are not moved off of their line. For a transformation that is defined geometrically, it is not necessary even to compute its matrix to find the eigenvectors and eigenvalues.

Illustration 3.2.4: Reflection

Here is an example of this. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation that reflects over the line L defined by $y = -x$, and let A be the matrix for T. We will find the eigenvalues and eigenvectors of A without doing any computations.

This transformation is defined geometrically, so we draw a picture.

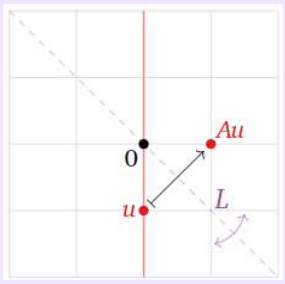


Figure 3.2.3

The vector $\vec{u}$ is not an eigenvector, because $A\vec{u}$ is not collinear with $\vec{u}$ and the origin.

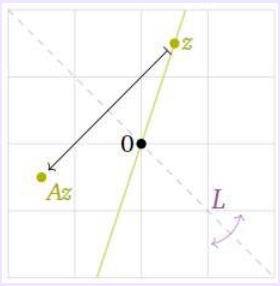


Figure 3.2.4

The vector $\vec{z}$ is not an eigenvector either.

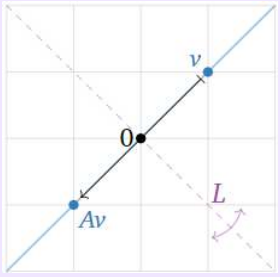


Figure 3.2.5

The vector $\vec{v}$ is an eigenvector because $A\vec{v}$ is collinear with $\vec{v}$ and the origin. The vector $A\vec{v}$ has the same length as $\vec{v}$, but the opposite direction, so the associated eigenvalue is -1.

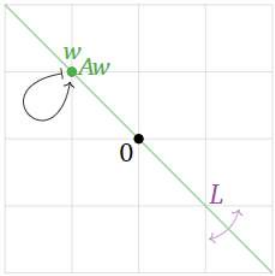


Figure 3.2.6

The vector $\vec{w}$ is an eigenvector because $A\vec{w}$ is collinear with $\vec{w}$ and the origin: indeed, $A\vec{w}$ is equal to $\vec{w}$! This means that $\vec{w}$ is an eigenvector with eigenvalue 1.

It appears that all eigenvectors lie either on L, or on the line perpendicular to L. The vectors on L have eigenvalue 1, and the vectors perpendicular to L have eigenvalue -1.

We will now give a few more examples of this nature

Illustration 3.2.5: Projection

Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation that projects a vector vertically onto the x-axis, and let A be the matrix for T. Find the eigenvalues and eigenvectors of A without doing any computations.

This transformation is defined geometrically, so we draw a picture.

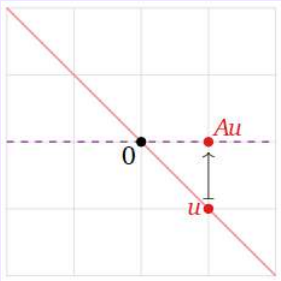


Figure 3.2.7

The vector $\vec{u}$ is not an eigenvector, because $A\vec{u}$ is not collinear with $\vec{u}$ and the origin.

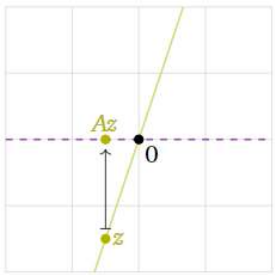


Figure 3.2.8

The vector $\vec{z}$ is not an eigenvector either.

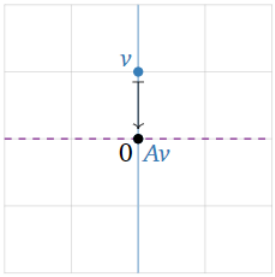


Figure 3.2.9

The vector $\vec{v}$ is an eigenvector. Indeed, $A\vec{v}$ is the zero vector, which is collinear with $\vec{v}$ and the origin; since $A\vec{v} = 0\vec{v}$, the associated eigenvalue is 0.

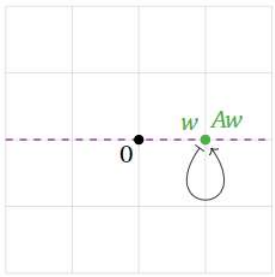


Figure 3.2.10

The vector $\vec{w}$ is an eigenvector because $A\vec{w}$ is collinear with $\vec{w}$ and the origin: indeed, $A\vec{w}$ is equal to $\vec{w}$! This means that $\vec{w}$ is an eigenvector with eigenvalue 1.

It appears that all eigenvectors lie on the x-axis or the y-axis. The vectors on the x-axis have eigenvalue 1, and the vectors on the y-axis have eigenvalue 0.

Example 3.2.1: Identity

Find all eigenvalues and eigenvectors of the identity matrix I.

Solution

The identity matrix has the property that $I\vec{v} = \vec{v}$ for all vectors $\vec{v}$ in $\mathbb{R}^n$. We can write this as $I\vec{v} = 1 \cdot \vec{v}$, so $\boxed{\text{every nonzero vector is an eigenvector with eigenvalue } 1}$

Example 3.2.2: Rotation

Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation that rotates counterclockwise by 90°, and let A be the matrix for T. Find the eigenvalues and eigenvectors of A without doing any computations.

Solution

If $\vec{v}$ is any nonzero vector, then $A\vec{v}$ is rotated by an angle of 90° from $\vec{v}$. Therefore, $A\vec{v}$ is not on the same line as $\vec{v}$, so $\vec{v}$ is not an eigenvector. And of course, the zero vector is never an eigenvector.

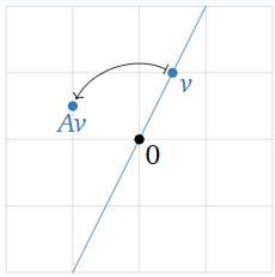


Figure 3.2.11

Therefore, this matrix $\boxed{\text{has no eigenvectors or eigenvalues}}$ *(at least no real number ones...)*

Here we mention one basic fact about eigenvectors.

Fact 3.2.1: Eigenvectors with Distinct Eigenvalues are Linearly Independent

Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ be eigenvectors of a matrix A, and suppose that the corresponding eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ are *distinct* (all different from each other). Then $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ is linearly independent.

Proof

Suppose that $\{\vec{v}_1\vec{v}_2, \dots, \vec{v}_k\}$ were linearly *dependent*. According to the increasing span criterion, this means that for some j, the vector $\vec{v}_j$ is in $\text{Span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{j-1}\}$. If we choose the *first* such j, then $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{j-1}\}$ is linearly independent. Note that $j > 1$ since $\vec{v}_1 \neq 0$.

Since $\vec{v}_j$ is in $\text{Span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{j-1}\}$, we can write

$$\vec{v}_j = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_{j-1}\vec{v}_{j-1}$$

for some scalars $c_1, c_2, \dots, c_{j-1}$. Multiplying both sides of the above equation by A gives

$$\begin{aligned} \lambda_j\vec{v}_j = A\vec{v}_j &= A\big(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_{j-1}\vec{v}_{j-1}\big) \\ &= c_1A\vec{v}_1 + c_2A\vec{v}_2 + \cdots + c_{j-1}A\vec{v}_{j-1} \\ &= c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2 + \cdots + c_{j-1}\lambda_{j-1}\vec{v}_{j-1}. \end{aligned}$$

Subtracting λ_j times the first equation from the second gives

$$\vec{0} = \lambda_j\vec{v}_j - \lambda_j\vec{v}_j = c_1(\lambda_1 - \lambda_j)\vec{v}_1 + c_2(\lambda_2 - \lambda_j)\vec{v}_2 + \cdots + c_{j-1}(\lambda_{j-1} - \lambda_j)\vec{v}_{j-1}.$$

Since $\lambda_i \neq \lambda_j$ for $i < j$, this is an equation of linear dependence among $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{j-1}$, which is impossible because those vectors are linearly independent. Therefore, $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ must have been linearly independent after all.

When $k = 2$, this says that if $\vec{v}_1, \vec{v}_2$ are eigenvectors with eigenvalues $\lambda_1 \neq \lambda_2$, then $\vec{v}_2$ is not a multiple of $\vec{v}_1$. In fact, any nonzero multiple $c\vec{v}_1$ of $\vec{v}_1$ is also an eigenvector with eigenvalue λ_1:

$$A(c\vec{v}_1) = cA\vec{v}_1 = c(\lambda_1\vec{v}_1) = \lambda_1(c\vec{v}_1).$$

As a consequence of the above Fact 3.2.1, we have the following.

Note 3.2.3

An $n \times n$ matrix A has at most n eigenvalues.

3.2: Eigenvalues and Eigenvectors Definitions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.1: Eigenvalues and Eigenvectors** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.2: Eigenspaces

Suppose that A is a square matrix. We already know how to check if a given vector is an eigenvector of A and in that case to find the eigenvalue. Our next goal is to check if a given real number is an eigenvalue of A and in that case to find all of the corresponding eigenvectors. Again this will be straightforward, but more involved.

Let A be an $n \times n$ matrix, and let λ be a scalar. The eigenvectors with eigenvalue λ, if any, are the nonzero solutions of the equation $A\vec{v} = \lambda\vec{v}$. We can rewrite this equation as follows:

$$\begin{aligned} & A\vec{v} = \lambda\vec{v} \\ \iff \quad & A\vec{v} - \lambda\vec{v} = \vec{0} \\ \iff \quad & A\vec{v} - \lambda I\vec{v} = \vec{0} \\ \iff \quad & (A - \lambda I)\vec{v} = \vec{0} \end{aligned}$$

Therefore, the eigenvectors of A with eigenvalue λ, if any, are the nontrivial solutions of the matrix equation $(A - \lambda I)\vec{v} = \vec{0}$, i.e., the nonzero vectors in $\text{Null}(A - \lambda I)$. If this equation has no nontrivial solutions, then λ is not an eigenvalue of A.

The above observation is important because it says that *finding the eigenvectors for a given eigenvalue means solving a homogeneous system of equations.* For instance, if

$$A = \begin{bmatrix} 7 & 1 & 3 \\ -3 & 2 & -3 \\ -3 & -2 & -1 \end{bmatrix}$$

then an eigenvector with eigenvalue λ is a nontrivial solution of the matrix equation

$$\begin{bmatrix} 7 & 1 & 3 \\ -3 & 2 & -3 \\ -3 & -2 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

This translates to the system of equations

$$\begin{cases} 7x + y + 3z = \lambda x \\ -3x + 2y - 3z = \lambda y \\ -3x - 2y - z = \lambda z \end{cases} \longrightarrow \begin{cases} (7-\lambda)x + y + 3z = 0 \\ -3x + (2-\lambda)y - 3z = 0 \\ -3x - 2y + (-1-\lambda)z = 0 \end{cases}$$

This is the same as the homogeneous matrix equation

$$\begin{bmatrix} 7-\lambda & 1 & 3 \\ -3 & 2-\lambda & -3 \\ -3 & -2 & -1-\lambda \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \vec{0}$$

i.e., $(A - \lambda I)\vec{v} = \vec{0}$.

Definition: λ-eigenspace

Let A be an $n \times n$ matrix, and let λ be an eigenvalue of A. The ***λ-eigenspace***, E_λ, of A is the solution set of $(A - \lambda I)\vec{v} = \vec{0}$, i.e.,

$$E_\lambda = \text{Null}(A - \lambda I)$$

The eigenspace E_λ is a subspace because it is the null space of a matrix, namely, the matrix $A - \lambda I$. This subspace consists of the zero vector and all eigenvectors of A with eigenvalue λ.

Note 3.2.4

Since a nonzero subspace is infinite, *every eigenvalue has infinitely many eigenvectors*. (For example, multiplying an eigenvector by a nonzero scalar gives another eigenvector.) On the other hand, there can be at most n linearly independent eigenvectors of an $n \times n$ matrix, since $\mathbb{R}^n$ has dimension n.

Example 3.2.3: Computing eigenspaces

For each of the numbers $\lambda = -2, 1, 3$, decide if λ is an eigenvalue of the matrix

$$A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$$

and if so, compute a basis for the λ-eigenspace.

Solution

The number 3 is an eigenvalue of A if and only if $\text{Null}(A - 3I)$ is nonzero. Hence, we have to solve the matrix equation $(A - 3I)\vec{v} = \vec{0}$. We have

$$A - 3I = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} - 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -4 \\ -1 & -4 \end{bmatrix}$$

The reduced row echelon form of this matrix is

$$\begin{bmatrix} 1 & 4 \\ 0 & 0 \end{bmatrix} \xrightarrow{\text{parametric form}} \begin{cases} x &= -4y \\ y &= y \end{cases} \xrightarrow{\text{parametric vector form}} \begin{bmatrix} x \\ y \end{bmatrix} = y\begin{bmatrix} -4 \\ 1 \end{bmatrix}$$

Since y is a free variable, the null space of $A - 3I$ is nonzero, so 3 is an eigenvalue. $\boxed{\text{A basis for } E_3 \text{ is } \left\{\begin{bmatrix} -4 \\ 1 \end{bmatrix}\right\}}$

Concretely, we have shown that the eigenvectors of A with eigenvalue 3 are exactly the nonzero multiples of $\begin{bmatrix} -4 \\ 1 \end{bmatrix}$. In particular, $\begin{bmatrix} -4 \\ 1 \end{bmatrix}$ is an eigenvector, which we can verify:

$$\begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}\begin{bmatrix} -4 \\ 1 \end{bmatrix} = \begin{bmatrix} -12 \\ 3 \end{bmatrix} = 3\begin{bmatrix} -4 \\ 1 \end{bmatrix}$$

The number 1 is an eigenvalue of A if and only if $\text{Null}(A - I)$ is nonzero. Hence, we have to solve the matrix equation $(A - I)\vec{v} = \vec{0}$. We have

$$A - I = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -4 \\ -1 & -2 \end{bmatrix}$$

This matrix has determinant -6, so it is invertible. This means we have $\text{Null}(A - I) = \{\vec{0}\}$, so $\boxed{1 \text{ is not an eigenvalue}}$

The eigenvectors of A with eigenvalue -2, if any, are the nonzero solutions of the matrix equation $(A + 2I)\vec{v} = \vec{0}$. We have

$$A + 2I = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & -4 \\ -1 & 1 \end{bmatrix}$$

The reduced row echelon form of this matrix is

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \xrightarrow{\text{parametric form}} \begin{cases} x &= y \\ y &= y \end{cases} \xrightarrow{\text{parametric vector form}} \begin{bmatrix} x \\ y \end{bmatrix} = y\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Hence there exist eigenvectors with eigenvalue -2, namely, any nonzero multiple of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. $\boxed{\text{A basis for } E_{-2} \text{ is } \left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}}$

Recipe: Eigenspaces

Let A be an $n \times n$ matrix and let λ be a number.

1. λ is an eigenvalue of A if and only if $(A - \lambda I)\vec{v} = \vec{0}$ has a nontrivial solution, if and only if $\text{Null}(A - \lambda I) \neq \{\vec{0}\}$.

2. In this case, finding a basis for the λ-eigenspace of A means finding a basis for $\text{Null}(A-\lambda I)$, which can be done by finding the parametric vector form of the solutions of the homogeneous system of equations $(A-\lambda I)\vec{v}=\vec{0}$.
3. The dimension of E_λ of A is equal to the number of free variables in the system of equations $(A-\lambda I)\vec{v}=\vec{0}$, which is the number of columns of $A-\lambda I$ without pivots.
4. The eigenvectors with eigenvalue λ are the nonzero vectors in $\text{Null}(A-\lambda I)$, or equivalently, the nontrivial solutions of $(A-\lambda I)\vec{v}=\vec{0}$.

3.2: Eigenspaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.1: Eigenvalues and Eigenvectors** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.2: Characteristic Polynomial

In this section, we will give a method for computing all of the *eigenvalues* of a matrix. This does not reduce to solving a system of linear equations: indeed, it requires solving a *nonlinear* equation in one variable, namely, finding the roots of the characteristic polynomial.

Definition: Characteristic Polynomial

Let A be an $n \times n$ matrix. The ***characteristic polynomial*** of A is the function $f(\lambda)$ given by:

$$f(\lambda) = \det(A - \lambda I)$$

We will see below that the characteristic polynomial is in fact a polynomial. Finding the characteristic polynomial means computing the determinant of the matrix $A - \lambda I$, whose entries contain the unknown λ.

Example 3.2.4

Find the characteristic polynomial of the matrix

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}$$

Solution

We have

$$\begin{aligned} f(\lambda) = \det(A - \lambda I) &= \det\left(\begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) \\ &= \begin{vmatrix} 3-\lambda & 2 \\ 2 & 0-\lambda \end{vmatrix} \\ &= (3-\lambda)(-\lambda) - 2 \cdot 2 \\ &= \boxed{\lambda^2 - 3\lambda - 4} \end{aligned}$$

Example 3.2.5

Find the characteristic polynomial of the matrix

$$A = \begin{bmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$$

Solution

We compute the determinant by expanding cofactors along the third column:

$$\begin{aligned} f(\lambda) = \det(A - \lambda I) &= \begin{vmatrix} -\lambda & 6 & 8 \\ \frac{1}{2} & -\lambda & 0 \\ 0 & \frac{1}{2} & -\lambda \end{vmatrix} \\ &= 8\left(\frac{1}{4} - 0 \cdot (-\lambda)\right) - \lambda\left(\lambda^2 - 6 \cdot \frac{1}{2}\right) \\ &= \boxed{-\lambda^3 + 3\lambda + 2} \end{aligned}$$

The point of the characteristic polynomial is that we can use it to compute eigenvalues.

Theorem 3.2.1: Eigenvalues are Roots of the Characteristic Polynomial

Let A be an $n \times n$ matrix, and let $f(\lambda) = \det(A - \lambda I)$ be its characteristic polynomial. Then a number λ_0 is an eigenvalue of A if and only if $f(\lambda_0) = 0$.

Proof

The matrix equation $(A - \lambda_0 I)\vec{x} = \vec{0}$ has a nontrivial solution if and only if $\det(A - \lambda_0 I) = 0$. Therefore,

$$\begin{aligned} \lambda_0 \text{ is an eigenvalue of } A &\iff A\vec{x} = \lambda_0\vec{x} \text{ has a nontrivial solution} \\ &\iff (A - \lambda_0 I)\vec{x} = \vec{0} \text{ has a nontrivial solution} \\ &\iff A - \lambda_0 I \text{ is not invertible} \\ &\iff \det(A - \lambda_0 I) = 0 \\ &\iff f(\lambda_0) = 0 \end{aligned}$$

Remark

By the above, the characteristic polynomial of an $n \times n$ matrix is a polynomial of degree n. Since a polynomial of degree n has at most n roots, this gives another proof of the fact that an $n \times n$ matrix has at most n eigenvalues.

Example 3.2.6: Finding eigenvalues

Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}$$

Solution

In the above Example 3.2.9 we computed the characteristic polynomial of A to be $f(\lambda) = \lambda^2 - 3\lambda - 4$. We can factor this to solve for its roots:

$$(\lambda - 4)(\lambda + 1) = 0 \implies \lambda = -1, 4$$

Therefore, the eigenvalues are -1 and 4.

To compute the eigenvectors, we solve the homogeneous system of equations $(A - \lambda I)\vec{x} = \vec{0}$ for each eigenvalue λ.

When $\lambda = -1$, we have

$$A - (-1)I = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}$$

The parametric form of the general solution is $x = -\frac{1}{2}y$, so the eigenspace $\boxed{E_{-1} = \text{Span}\left\{\begin{bmatrix} -\frac{1}{2} \\ 1 \end{bmatrix}\right\}}$

When $\lambda = 4$, we have

$$A - 4I = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}$$

The parametric form of the general solution is $x = 2y$, so the eigenspace $\boxed{E_4 = \text{Span}\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}}$

Example 3.2.7: Finding eigenvalues

Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{bmatrix}$$

Solution

In the above Example 3.2.10 we computed the characteristic polynomial of A to be $f(\lambda) = -\lambda^3 + 3\lambda + 2$. We factor this completely:

$$f(\lambda) = -(\lambda - 2)(\lambda + 1)^2$$

so the only eigenvalues are $\lambda = 2, -1$.

We compute the 2-eigenspace, E_2 by solving the homogeneous system $(A - 2I)\vec{x} = \vec{0}$. We have

$$A - 2I = \begin{bmatrix} -2 & 6 & 8 \\ \frac{1}{2} & -2 & 0 \\ 0 & \frac{1}{2} & -2 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & -16 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{bmatrix}$$

The parametric form and parametric vector form of the solutions are:

$$\begin{cases} x &= 16z \\ y &= 4z \\ z &= z \end{cases} \longrightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix} = z \begin{bmatrix} 16 \\ 4 \\ 1 \end{bmatrix}$$

Therefore, the 2-eigenspace is the line

$$\boxed{E_1 = \text{Span}\left\{\begin{bmatrix} 16 \\ 4 \\ 1 \end{bmatrix}\right\}}$$

We compute the -1-eigenspace by solving the homogeneous system $(A + I)\vec{x} = \vec{0}$. We have

$$A + I = \begin{bmatrix} 1 & 6 & 8 \\ \frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

The parametric form and parametric vector form of the solutions are:

$$\begin{cases} z &= 4z \\ y &= -2z \\ z &= z \end{cases} \longrightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix} = z \begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}$$

Therefore, the -1-eigenspace is the line

$$\boxed{E_{-1} = \text{Span}\left\{\begin{bmatrix} 4 \\ -2 \\ 1 \end{bmatrix}\right\}}$$

3.2: Characteristic Polynomial is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.2: The Characteristic Polynomial** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.2: Practice Set

Exercise 3.2.1a

Find all eigenvalues and eigenspaces for the matrix $A = \begin{bmatrix} 0 & \pi \\ 0 & 0 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 0-\lambda & \pi \\ 0 & 0-\lambda \end{vmatrix} = (0-\lambda)(0-\lambda) - 0 \cdot \pi = \lambda^2$$

The roots of this are our eigenvalues, giving us: $\lambda = 0$

$$E_0 = \text{Null}\,(A - 0 \cdot I) = \text{Null}\left(\begin{bmatrix} 0 & \pi \\ 0 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$$

Exercise 3.2.1b

Find all eigenvalues and eigenspaces for the matrix $B = \begin{bmatrix} 5 & 12 \\ 3 & 5 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(B - \lambda I) = \begin{vmatrix} 5-\lambda & 12 \\ 3 & 5-\lambda \end{vmatrix} = (5-\lambda)(5-\lambda) - 3 \cdot 12 = \lambda^2 - 10\lambda - 11$$

The roots of this are our eigenvalues, giving us: $\lambda = -1, 11$

$$E_{-1} = \text{Null}\,(B - (-1) \cdot I) = \text{Null}\left(\begin{bmatrix} 6 & 12 \\ 3 & 6 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}$$

$$E_{11} = \text{Null}\,(B - 11 \cdot I) = \text{Null}\left(\begin{bmatrix} -6 & 12 \\ 3 & -6 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$$

Exercise 3.2.1c

Find all eigenvalues and eigenspaces for the matrix $C = \begin{bmatrix} -2 & 5 \\ 7 & 8 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(C - \lambda I) = \begin{vmatrix} -2-\lambda & 5 \\ 7 & 8-\lambda \end{vmatrix} = (-2-\lambda)(8-\lambda) - 7 \cdot 5 = \lambda^2 - 6\lambda - 51$$

The roots of this are our eigenvalues, giving us: $\lambda = 3 \pm 2\sqrt{15}$

$$E_{3+2\sqrt{15}} = \text{Null}\left(C - (3+2\sqrt{15}) \cdot I\right) = \text{Null}\left(\begin{bmatrix} -5-2\sqrt{15} & 5 \\ 7 & 5-2\sqrt{15} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & \frac{5-2\sqrt{15}}{7} \\ 0 & 0 \end{bmatrix}\right)$$
$$= \text{Span}\left\{\begin{bmatrix} -5+2\sqrt{15} \\ 7 \end{bmatrix}\right\}$$

$$E_{3-2\sqrt{15}} = \text{Null}\left(C - (3-2\sqrt{15}) \cdot I\right) = \text{Null}\left(\begin{bmatrix} -5+2\sqrt{15} & 5 \\ 7 & 5+2\sqrt{15} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & \frac{5+2\sqrt{15}}{7} \\ 0 & 0 \end{bmatrix}\right)$$
$$= \text{Span}\left\{\begin{bmatrix} -5-2\sqrt{15} \\ 7 \end{bmatrix}\right\}$$

Exercise 3.2.1d

Find all eigenvalues and eigenspaces for the matrix $D = \begin{bmatrix} -1 & -3 \\ 4 & 6 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(D - \lambda I) = \begin{vmatrix} -1-\lambda & -3 \\ 4 & 6-\lambda \end{vmatrix} = (-1-\lambda)(6-\lambda) - 4 \cdot (-3) = \lambda^2 - 5\lambda + 6$$

The roots of this are our eigenvalues, giving us: $\lambda = 2, 3$

$$E_2 = \text{Null}\,(D - 2 \cdot I) = \text{Null}\left(\begin{bmatrix} -3 & -3 \\ 4 & 4 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$$

$$E_3 = \text{Null}\,(D - 3 \cdot I) = \text{Null}\left(\begin{bmatrix} -4 & -3 \\ 4 & 3 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & \frac{3}{4} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -3 \\ 4 \end{bmatrix}\right\}$$

Exercise 3.2.1e

Find all eigenvalues and eigenspaces for the matrix $E = \begin{bmatrix} \frac{1}{2} & \frac{2}{3} \\ 1 & -\frac{1}{3} \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(E - \lambda I) = \begin{vmatrix} \frac{1}{2}-\lambda & \frac{2}{3} \\ 1 & -\frac{1}{3}-\lambda \end{vmatrix} = (\tfrac{1}{2}-\lambda)(-\tfrac{1}{3}-\lambda) - 1 \cdot \tfrac{2}{3} = \lambda^2 - \tfrac{1}{6}\lambda - \tfrac{5}{6} = \tfrac{1}{6}\left(6\lambda^2 - \lambda - 5\right)$$

The roots of this are our eigenvalues, giving us: $\lambda = 1, -\frac{5}{6}$

$$E_1 = \text{Null}\,(E - 1 \cdot I) = \text{Null}\left(\begin{bmatrix} -\frac{1}{2} & \frac{2}{3} \\ 1 & -\frac{4}{3} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -\frac{4}{3} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 4 \\ 3 \end{bmatrix}\right\}$$

$$E_{-\frac{5}{6}} = \text{Null}\left(E - \left(-\tfrac{5}{6}\right) \cdot I\right) = \text{Null}\left(\begin{bmatrix} \frac{4}{3} & \frac{2}{3} \\ 1 & \frac{1}{2} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right\}$$

Exercise 3.2.1f

Find all eigenvalues and eigenspaces for the matrix $F = \begin{bmatrix} \sqrt[4]{3} & 0 \\ 0 & \sqrt[3]{4} \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(F - \lambda I) = \begin{vmatrix} \sqrt[4]{3}-\lambda & 0 \\ 0 & \sqrt[3]{4}-\lambda \end{vmatrix} = (\sqrt[4]{3}-\lambda)(\sqrt[3]{4}-\lambda) - 0 \cdot 0 = (\sqrt[4]{3}-\lambda)(\sqrt[3]{4}-\lambda)$$

The roots of this are our eigenvalues, giving us: $\lambda = \sqrt[4]{3}, \sqrt[3]{4}$

$$E_{\sqrt[4]{3}} = \text{Null}\left(F - \sqrt[4]{3} \cdot I\right) = \text{Null}\left(\begin{bmatrix} 0 & 0 \\ 0 & \sqrt[3]{4} - \sqrt[4]{3} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$$

$$E_{\sqrt[3]{4}} = \text{Null}\left(F - \sqrt[3]{4} \cdot I\right) = \text{Null}\left(\begin{bmatrix} \sqrt[4]{3} - \sqrt[3]{4} & 0 \\ 0 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$$

Exercise 3.2.1g

Find all eigenvalues and eigenspaces for the matrix $G = \begin{bmatrix} 0 & a^2 \\ b^2 & 0 \end{bmatrix}$ where $a, b > 0$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(G - \lambda I) = \begin{vmatrix} 0-\lambda & a^2 \\ b^2 & 0-\lambda \end{vmatrix} = (0-\lambda)(0-\lambda) - a^2 \cdot b^2 = \lambda^2 - a^2 b^2$$

The roots of this are our eigenvalues, giving us: $\lambda = \pm ab$

$$E_{ab} = \text{Null}\,(G - ab \cdot I) = \text{Null}\left(\begin{bmatrix} -ab & a^2 \\ b^2 & -ab \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -\frac{a}{b} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} a \\ b \end{bmatrix}\right\}$$

$$E_{-ab} = \text{Null}\,(G - (-ab) \cdot I) = \text{Null}\left(\begin{bmatrix} ab & a^2 \\ b^2 & ab \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & \frac{a}{b} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -a \\ b \end{bmatrix}\right\}$$

Exercise 3.2.2a

Find all eigenvalues and eigenspaces for the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 0 & 0 & 5 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 2 & 2-\lambda & 2 \\ 0 & 0 & 5-\lambda \end{vmatrix} = (5-\lambda)\begin{vmatrix} 1-\lambda & 1 \\ 2 & 2-\lambda \end{vmatrix} = (\lambda - 5)(-\lambda^2 + 3\lambda)$$

The roots of this are our eigenvalues, giving us: $\lambda = 0, 3, 5$

$$E_0 = \text{Null}\,(A - 0 \cdot I) = \text{Null}\left(\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 0 & 0 & 5 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\}$$

$$E_3 = \text{Null}\,(A - 3 \cdot I) = \text{Null}\left(\begin{bmatrix} -2 & 1 & 1 \\ 2 & -1 & 2 \\ 0 & 0 & 2 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right\}$$

$$E_5 = \text{Null}\,(A - 5 \cdot I) = \text{Null}\left(\begin{bmatrix} -4 & 1 & 1 \\ 2 & -3 & 2 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\right\}$$

Exercise 3.2.2b

Find all eigenvalues and eigenspaces for the matrix $B = \begin{bmatrix} 1 & 5 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & -1 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(B - \lambda I) = \begin{vmatrix} 1-\lambda & 5 & 1 \\ 0 & 2-\lambda & -1 \\ 0 & 0 & -1-\lambda \end{vmatrix} = (-1-\lambda)\begin{vmatrix} 1-\lambda & 5 \\ 0 & 2-\lambda \end{vmatrix} = (-1-\lambda)(1-\lambda)(2-\lambda)$$

The roots of this are our eigenvalues, giving us: $\lambda = -1, 1, 2$

$$E_{-1} = \text{Null}\,(B-(-1)\cdot I) = \text{Null}\left(\begin{bmatrix} 2 & 5 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & \frac{4}{3} \\ 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix}\right\}$$

$$E_{1} = \text{Null}\,(B-1\cdot I) = \text{Null}\left(\begin{bmatrix} 0 & 5 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -2 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}$$

$$E_{2} = \text{Null}\,(B-2\cdot I) = \text{Null}\left(\begin{bmatrix} -1 & 5 & 1 \\ 0 & 0 & -1 \\ 0 & 0 & -3 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -5 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}\right\}$$

Exercise 3.2.2c

Find all eigenvalues and eigenspaces for the matrix $C = \begin{bmatrix} 1 & 1 & -2 \\ 5 & 0 & 5 \\ 7 & 2 & 1 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(C-\lambda I) = \begin{vmatrix} 1-\lambda & 1 & -2 \\ 5 & 0-\lambda & 5 \\ 7 & 2 & 1-\lambda \end{vmatrix} = -\lambda^3 + 2\lambda^2$$

The roots of this are our eigenvalues, giving us: $\lambda = 0, 2$

$$E_{0} = \text{Null}\,(C-0\cdot I) = \text{Null}\left(\begin{bmatrix} 1 & 1 & -2 \\ 5 & 0 & 5 \\ 7 & 2 & 1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}\right\}$$

$$E_{2} = \text{Null}\,(C-2\cdot I) = \text{Null}\left(\begin{bmatrix} -1 & 1 & -2 \\ 5 & -2 & 5 \\ 7 & 2 & -1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & -\frac{5}{3} \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 5 \\ 3 \end{bmatrix}\right\}$$

Exercise 3.2.2d

Find all eigenvalues and eigenspaces for the matrix $D = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 1 \\ 3 & 0 & 1 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(D-\lambda I) = \begin{vmatrix} 1-\lambda & 0 & 1 \\ 0 & 2-\lambda & 1 \\ 3 & 0 & 1-\lambda \end{vmatrix} = (2-\lambda)\begin{vmatrix} 1-\lambda & 1 \\ 3 & 1-\lambda \end{vmatrix} = (2-\lambda)(\lambda^2-2\lambda-2)$$

The roots of this are our eigenvalues, giving us: $\lambda = 2, 1\pm\sqrt{3}$

$$E_{2} = \text{Null}\,(D-2\cdot I) = \text{Null}\left(\begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 1 \\ 3 & 0 & -1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$$

$$E_{1+\sqrt{3}} = \text{Null}\left(D-(1+\sqrt{3})\cdot I\right) = \text{Null}\left(\begin{bmatrix} -\sqrt{3} & 0 & 1 \\ 0 & 1-\sqrt{3} & 1 \\ 3 & 0 & -\sqrt{3} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -\frac{1}{\sqrt{3}} \\ 0 & 1 & \frac{1}{1-\sqrt{3}} \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{1-\sqrt{3}} \\ 1 \end{bmatrix}\right\}$$

$$E_{1-\sqrt{3}} = \text{Null}\left(D-(1-\sqrt{3})\cdot I\right) = \text{Null}\left(\begin{bmatrix} \sqrt{3} & 0 & 1 \\ 0 & 1+\sqrt{3} & 1 \\ 3 & 0 & \sqrt{3} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & \frac{1}{\sqrt{3}} \\ 0 & 1 & \frac{1}{1+\sqrt{3}} \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -\frac{1}{\sqrt{3}} \\ -\frac{1}{1+\sqrt{3}} \\ 1 \end{bmatrix}\right\}$$

Exercise 3.2.2e

Find all eigenvalues and eigenspaces for the matrix $E = \begin{bmatrix} -2 & 5 & -5 \\ 2 & 1 & 3 \\ 6 & -6 & 10 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(E - \lambda I) = \begin{vmatrix} -2-\lambda & 5 & -5 \\ 2 & 1-\lambda & 3 \\ 6 & -6 & 10-\lambda \end{vmatrix} = -\lambda^3 + 9\lambda^2 - 26\lambda + 24$$

The roots of this are our eigenvalues, giving us: $\lambda = 2, 3, 4$

$$E_2 = \text{Null}\,(E - 2\cdot I) = \text{Null}\left(\begin{bmatrix} -4 & 5 & -5 \\ 2 & -1 & 3 \\ 6 & -6 & 8 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & \frac{5}{3} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -5 \\ -1 \\ 3 \end{bmatrix}\right\}$$

$$E_3 = \text{Null}\,(E - 3\cdot I) = \text{Null}\left(\begin{bmatrix} -5 & 5 & -5 \\ 2 & -2 & 3 \\ 6 & -6 & 7 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$$

$$E_4 = \text{Null}\,(E - 4\cdot I) = \text{Null}\left(\begin{bmatrix} -6 & 5 & -5 \\ 2 & -3 & 3 \\ 6 & -6 & 6 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}$$

Exercise 3.2.2f

Find all eigenvalues and eigenspaces for the matrix $F = \begin{bmatrix} 2 & 1 & -2 \\ 4 & 0 & 4 \\ 1 & 3 & -4 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(F - \lambda I) = \begin{vmatrix} 2-\lambda & 1 & -2 \\ 4 & 0-\lambda & 4 \\ 1 & 3 & -4-\lambda \end{vmatrix} = -\lambda^3 - 2\lambda^2 + 22\lambda - 28$$

The roots of this are our eigenvalues, giving us: $\lambda = 2, -2 \pm 3\sqrt{2}$

$$E_2 = \text{Null}\,(F - 2\cdot I) = \text{Null}\left(\begin{bmatrix} 0 & 1 & -2 \\ 4 & -2 & 4 \\ 1 & 3 & -6 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}\right\}$$

$$E_{-2+3\sqrt{2}} = \text{Null}\left(F - (-2+3\sqrt{2})\cdot I\right) = \text{Null}\left(\begin{bmatrix} 4-3\sqrt{2} & 1 & -2 \\ 4 & 2-3\sqrt{2} & 4 \\ 1 & 3 & -2-3\sqrt{2} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & \frac{-10+3\sqrt{2}}{41} \\ 0 & 1 & \frac{-24-42\sqrt{2}}{41} \\ 0 & 0 & 0 \end{bmatrix}\right)$$

$$= \text{Span}\left\{\begin{bmatrix} \frac{10-3\sqrt{2}}{41} \\ \frac{24+42\sqrt{2}}{41} \\ 1 \end{bmatrix}\right\}$$

$$E_{-2-3\sqrt{2}} = \text{Null}\left(F - (-2-3\sqrt{2})\cdot I\right) = \text{Null}\left(\begin{bmatrix} 4+3\sqrt{2} & 1 & -2 \\ 4 & 2+3\sqrt{2} & 4 \\ 1 & 3 & -2+3\sqrt{2} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & \frac{-10-3\sqrt{2}}{41} \\ 0 & 1 & \frac{-24+42\sqrt{2}}{41} \\ 0 & 0 & 0 \end{bmatrix}\right)$$

$$= \text{Span}\left\{\begin{bmatrix} \frac{10+3\sqrt{2}}{41} \\ \frac{24-42\sqrt{2}}{41} \\ 1 \end{bmatrix}\right\}$$

? Exercise 3.2.3

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If $\vec{v}$ is an eigenvector of A, then $\{A\vec{v}, A^2\vec{v}\}$ is linearly dependent.

(b) If $\vec{u}$ and $\vec{v}$ are nonzero eigenvectors of A, then $\{A\vec{u}, A\vec{v}\}$ is linearly independent.

(c) An $n \times n$ matrix will have n distinct real eigenvalues.

(d) If A is a 3×3 matrix with 3 linearly independent real eigenvectors, then it will have 3 distinct real eigenvalues.

(e) If a matrix has $\lambda = 0$ as an eigenvalue, then it is invertible.

(f) If A is a 4×4 matrix with characteristic polynomial $p(\lambda) = \lambda^4 + a\lambda^3 + b\lambda^2 + c\lambda + d$ with $d \neq 0$, then it is invertible.

(g) If $\vec{v}$ is an eigenvector of A with eigenvalue λ, then $c\vec{v}$ is an eigenvector of A with eigenvalue $c\lambda$ for every c in $\mathbb{R}$.

(h) If $\vec{v}$ is an eigenvector of A with eigenvalue λ, then $\vec{v}$ is an eigenvector of A^2 with eigenvalue λ^2.

(i) If λ is an eigenvalue of A, then $\det(A^k - \lambda^k I) = 0$.

(j) The matrix A and A^T have the same eigenvalues.

(k) The matrix A and A^T have the same eigenvectors.

Answer

(a) **Always True**

$A\vec{v} = \lambda\vec{v}$ and $A^2\vec{v} = \lambda^2\vec{v}$, so $A^2\vec{v} = \lambda A\vec{v}$ and the set is dependent.

(b) **Sometimes True Sometimes False**

When they have distinct eigenvalues it is assured, but they could already be dependent beforehand and make this false instead.

(c) **Sometimes True Sometimes False**

An easy example where this is true: $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$.

An example of it being false is the 2×2 identity matrix, which only has one eigenvalue.

(d) **Sometimes True Sometimes False**

This is true for an example like $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. For false, consider the 3×3 identity matrix, which only has one eigenvalue but three linearly independent eigenvectors.

(e) **Always False**

If 0 is an eigenvalue, then $p(0) = \det(A) = 0$, and A is not invertible.

(f) **Always True**

Since $p(0) = d \neq 0$ we know $\lambda = 0$ is not an eigenvalue, so A is invertible.

(g) **Sometimes True Sometimes False**

This is true when $\lambda = 0$ this is automatically true. In every other case, the true eigenvalue of $c\vec{v}$ is exactly $\lambda \neq c\lambda$.

(h) **Always True**

$A^2\vec{v} = A(A\vec{v}) = A(\lambda\vec{v}) = \lambda(A\vec{v}) = \lambda^2\vec{v}$.

(i) **Always True**

If $\vec{v}$ is a nonzero eigenvector with eigenvalue λ, then $A^k\vec{v} = \lambda^k\vec{v}$ must hold.

(j) **Always True**

They have the same characteristic polynomial: $\det(A-\lambda I)=\det\left((A-\lambda I)^T\right)=\det\left(A^T-\lambda I\right)$.

(k) **Sometimes True Sometimes False**

When $A=A^T$ it is obviously true, but it can be false, like for the matrix $A=\begin{bmatrix}1 & 1\\ 0 & 2\end{bmatrix}$.

3.2: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.3: Diagonalization and Geometry of Eigenvalues

Objectives

1. Find a diagonal matrix similar to a given matrix A
2. Understand the geometry of diagonalizable 2×2 matrices.
3. *Recipes:* diagonalize a matrix
4. *Vocabulary words:* ***algebraic multiplicity, geometric multiplicity, diagonalization***.

In this section, we connect the concept of a change of basis to that of eigenvalues and eigenvectors in a powerful way called diagonalization.

This page titled 3.3: Diagonalization and Geometry of Eigenvalues is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber.

- **2.8: Bases as Coordinate Systems** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.3: Diagonalization

Recall that two matrices A and B are *similar* if we can find an invertible matrix P such that $A = PBP^{-1}$.

Now we turn to an application that ties together similarity and eigenvalues. This is because of the following observation:

Fact 3.3.1

Let $A = PBP^{-1}$. Then for any $n \geq 1$, we have

$$A^n = PB^nP^{-1}$$

Proof

First note that

$$A^2 = AA = (PBP^{-1})(PBP^{-1}) = PB(P^{-1}P)BP^{-1} = PBIBP^{-1} = PB^2P^{-1}$$

Next we have

$$A^3 = A^2A = (PB^2P^{-1})(PBP^{-1}) = PB^2(P^{-1}P)BP^{-1} = PB^3P^{-1}$$

The pattern is clear.

As we can see, the best matrix to be similar to is a *diagonal matrix*. This is because they are easiest to multiply by, even with themselves.

Note 3.3.1

If D is the diagonal matrix, $D = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_m \end{bmatrix}$, then any power of D is easy: $D^k = \begin{bmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_m^k \end{bmatrix}$

Since similar matrices are just re-coordinatizations of the same linear operation, it makes sense that we would try to use a basis $\mathcal{B}$ such that $A = Q_{\mathcal{B}}BQ_{\mathcal{B}}^{-1}$ yields a "nice" matrix B. We want to make this tedious process have some payoff, like *multiplication by B is easy compared to A*.

Illustration 3.3.1

Let's compute A^{100}, where

$$A = \begin{bmatrix} 5 & 13 \\ -2 & -5 \end{bmatrix} = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}^{-1}$$

By the fact above, we have

$$A^{100} = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^{100} \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}^{-1}$$

The matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is a counterclockwise rotation by 90°. If we rotate by 90° four times, then we end up where we started. Hence rotating by 90° one hundred times is the identity transformation, so

$$A^{100} = \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & 3 \\ 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

So how do we pick $\mathcal{B}$ to become similar to a diagonal matrix? By using eigenvectors!

Definition: Diagonalizable

An $n \times n$ matrix A is ***diagonalizable*** if there are n linearly independent eigenvectors of A.

This is because we can let $\mathcal{B} = \{\vec{b}_1, \ldots, \vec{b}_n\}$ be the basis of $\mathbb{R}^n$ comprised of exactly those eigenvectors, (so that $A\vec{b}_i = \lambda_i \vec{b}_i$ for each $i = 1, \ldots n$)

Then we would have $A = Q_{\mathcal{B}} D Q_{\mathcal{B}}^{-1}$ for the diagonal matrix $D = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$

Example 3.3.1

Find a formula for $\begin{bmatrix} \frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{1}{2} \end{bmatrix}^n$ in which the entries are functions of n, where n is any positive whole number.

Solution

We need to find the eigenvalues and eigenvectors of $A = \begin{bmatrix} \frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{1}{2} \end{bmatrix}$. First we compute the characteristic polynomial:

$$f(\lambda) = \det(A - \lambda I) = \begin{vmatrix} \frac{1}{2} - \lambda & \frac{3}{2} \\ \frac{3}{2} & \frac{1}{2} - \lambda \end{vmatrix} = \lambda^2 - \lambda - 2 = (\lambda + 1)(\lambda - 2)$$

Therefore, the eigenvalues are -1 and 2. We need to compute eigenvectors for each eigenvalue. We start with $\lambda_1 = -1$:

$$(A + 1I)\vec{v} = \vec{0} \iff \vec{v} \text{ is in Null}\left(\begin{bmatrix} \frac{3}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{3}{2} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right)$$

The parametric form is $x = -y$, so $\vec{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue -1.

Now we find an eigenvector with eigenvalue $\lambda_2 = 2$:

$$(A - 2I)\vec{v} = \vec{0} \iff \vec{v} \text{ is in Null}\left(\begin{bmatrix} -\frac{3}{2} & \frac{3}{2} \\ \frac{3}{2} & -\frac{3}{2} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}\right)$$

The parametric form is $x = y$, so $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is an eigenvector with eigenvalue 2.

The eigenvectors $\vec{v}_1, \vec{v}_2$ are linearly independent, so we see that

$$A = PDP^{-1} \quad \text{for} \quad P = \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \quad D = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$$

Alternatively, if we choose 2 as our first eigenvalue, then

$$A = QD'Q^{-1} \quad \text{for} \quad Q = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad D' = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

Using either of these two changes of bases will result in the same final answer, so we will pick Q and D':

$$
\begin{aligned}
A^n &= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}^n \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \\
&= \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2^n & 0 \\ 0 & (-1)^n \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \\
&= \begin{bmatrix} 2^n & (-1)^{n+1} \\ 2^n & (-1)^n \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \\
&= \boxed{\frac{1}{2} \begin{bmatrix} 2^n + (-1)^{n+2} & 2^n + (-1)^{n+1} \\ 2^n + (-1)^{n+1} & 2^n + (-1)^n \end{bmatrix}}
\end{aligned}
$$

This feels like a lot of work, but manually calculating A^n even for a fixed large n is *also a lot of work*. The difference here is that diagonalization takes the same number of steps whether you are finding A^3 or A^{100} or $A^{1035816128}$.

The bad news however, is that not every matrix is diagonalizable.

Illustration 3.3.2: A shear is not diagonalizable

Let $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ so $T(x) = Bx$ is a shear transformation.

The characteristic polynomial of B is $f(\lambda) = (\lambda - 1)^2$, so the only eigenvalue of B is 1. We compute the 1-eigenspace:

$$
(B - I)\vec{v} = \vec{0} \iff \vec{v} \text{ is in } \text{Null}(B - I) = \text{Null}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}
$$

In other words, the 1-eigenspace is exactly the x-axis, so *all* of the eigenvectors of B lie on the x-axis. It follows that B does *not* admit two linearly independent eigenvectors, so it is not diagonalizable.

Notice that in the above example, the eigenvalue $\lambda = 1$ was a double root of the characteristic polynomial $f(\lambda)$. Because of this, there was no second eigenvalue, and therefore no second eigenvector to make a basis with. In general if λ_0 is a double root of the characteristic polynomial, it would need to have two linearly independent eigenvectors attached to that eigenvalue in order to become diagonalizable. This is made more explicit by two definitions of multiplicity:

Definition: Multiplicity

Any real number c has two multiplicities with respect to the $n \times n$ matrix A.

The **algebraic multiplicity** of c is the number of times it is a root of the characteristic polynomial of A. That is if $f(\lambda) = (\lambda - c)^\alpha \cdot p(\lambda)$ for some polynomial $p(\lambda)$ with $p(c) \neq 0$, then the algebraic multiplicity of c is α.

The **geometric multiplicity** of c is the dimension of the eigenspace E_c, or equivalently the nullity of the matrix $A - cI$.

These two definitions are different, and they do not always match each other!

Illustration 3.3.3: Multiplicities in previous examples

We saw in Example 3.3.1 that the matrix $A = \begin{bmatrix} \frac{1}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{1}{2} \end{bmatrix}$ had characteristic polynomial $f(\lambda) = (\lambda + 1)(\lambda - 2)$. The eigenvalue $\lambda = -1$ has algebraic multiplicity 1 since it is a single root of $f(\lambda)$ and geometric multiplicity 1 since E_{-1} was spanned by one vector.

Because these two multiplicities matched for each eigenvalue, we were able to form a basis of eigenvectors and diagonalize A successfully.

Alternatively, the matrix $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ from Illustration 3.3.2 had characteristic polynomial $f(\lambda) = (\lambda - 1)^2$. The eigenvalue $\lambda = 1$ has algebraic multiplicity 2 since it is a double root of $f(\lambda)$ but has geometric multiplicity 1 since E_1 was spanned by one vector.

Because these two multiplicities disagreed for $\lambda = 1$, we were not able to form a basis of eigenvectors or diagonalize B.

Note 3.3.2

A square matrix A will be diagonalizable if and only if for each eigenvalue, its algebraic multiplicity equals its geometric multiplicity.

Example 3.3.2 A diagonalizable 3×3 matrix

Diagonalize $A = \begin{bmatrix} 4 & -3 & 0 \\ 2 & -1 & 0 \\ 1 & -1 & 1 \end{bmatrix}$ or show it is impossible to do so.

Solution

We need to find the eigenvalues and eigenvectors of A. First we compute the characteristic polynomial by expanding cofactors along the third column:

$$f(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 4-\lambda & -3 & 0 \\ 2 & -1-\lambda & 0 \\ 1 & -1 & 1-\lambda \end{vmatrix} = (1-\lambda) \begin{vmatrix} 4-\lambda & -3 \\ 2 & -1-\lambda \end{vmatrix}$$
$$= (1-\lambda)(\lambda^2 - 3\lambda + 2) = -(\lambda - 1)^2(\lambda - 2)$$

Therefore, the eigenvalues are 1 and 2. We need to compute eigenvectors for each eigenvalue. We start with $\lambda_1 = 1$:

$$(A - I)\vec{v} = \vec{0} \iff \vec{v} \text{ is in } \text{Null}(A - I) = \text{Null}\left(\begin{bmatrix} 3 & -3 & 0 \\ 2 & -2 & 0 \\ 1 & -1 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)$$

The parametric vector form is

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

Hence the algebraic multiplicity and geometric multiplicity for $\lambda = 1$ are both two.

Now we compute the eigenspace for $\lambda_2 = 2$:

$$(A - 2I)\vec{v} = \vec{0} \iff \vec{v} \text{ is in}$$
$$\text{Null}(A - 2I) = \text{Null}\left(\begin{bmatrix} 2 & -3 & 0 \\ 2 & -3 & 0 \\ 1 & -1 & -1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}\right)$$

The parametric form is $x = 3z, y = 2z$, so an eigenvector which spans the eigenspace E_2 is

$$\vec{v}_3 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Here the algebraic multiplicity and geometric multiplicity for $\lambda = 2$ are both one.

Since both pairs of multiplicities match, we know the matrix is diagonalizable. We have already found all of the major ingredients needed along the way:

$$A = PDP^{-1} \quad \text{for} \quad P = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Example 3.3.3: A non-diagonalizable 3×3 matrix

Diagonalize $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ or show it is impossible to do so.

Solution

We need to find the eigenvalues and eigenvectors of A. Again we compute the characteristic polynomial by expanding cofactors along the third column:

$$f(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda) \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (2-\lambda)(\lambda-1)^2$$

Therefore, the eigenvalues are 1 and 2. We need to compute eigenvectors for each eigenvalue. We start with $\lambda_1 = 1$:

$$(A - I)\vec{v} = \vec{0} \iff \vec{v} \text{ is in } \text{Null}(A - I) = \text{Null}\left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}\right)$$
$$= \text{Null}\left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right)$$

In other words, the 1-eigenspace is the x-axis. The algebraic multiplicity is 2, but the geometric multiplicity is only 1. Since these do not match, we do not need to check the second eigenvalue, there is no way for its geometric multiplicity to surpass its algebraic multiplicity, so we cannot make a basis of eigenvectors for $\mathbb{R}^3$. We know that $\boxed{A \text{ is not diagonalizable}}$

3.3: Diagonalization is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.3: Similarity** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **5.4: Diagonalization** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.3: Geometry of Diagonalizable Matrices

If we assume that a matrix A is similar to some diagonal matrix D, first we explore the geometric properties of diagonal matrices.

A diagonal matrix is easy to understand geometrically, as it just scales the coordinate axes:

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = a \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = b \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = c \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Therefore, we know that a diagonalizable matrix simply scales the "axes" with respect to a different coordinate system. Indeed, if $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ are linearly independent eigenvectors of an $n \times n$ matrix A, then A scales the $\vec{v}_i$-direction by the eigenvalue λ_i.

In the following examples, we visualize the action of a diagonalizable matrix A in terms of its *dynamics*. In other words, we start with a collection of vectors (drawn as points), and we see where they move when we multiply them by A repeatedly.

Illustration 3.3.4: Eigenvalues $\lambda_1 > 1$, $\lambda_2 = 1$

Let's describe how the matrix $A = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$ acts on the plane.

First we diagonalize A. The characteristic polynomial is

$$f(\lambda) = \begin{vmatrix} 4-\lambda & 3 \\ 1 & 2-\lambda \end{vmatrix} = (4-\lambda)(2-\lambda) - 3 = \lambda^2 - 6\lambda + 5 = (\lambda - 5)(\lambda - 1)$$

We compute the 5-eigenspace:

$$E_5 = \text{Null}(A - 5I) = \text{Null}\left(\begin{bmatrix} -1 & 3 \\ 1 & -3 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}\right)$$

The parametric form of this equation is $x = 3y$, so one eigenvector is $\vec{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

For the 1-eigenspace, we have:

$$E_1 = \text{Null}(A - I) = \text{Null}\left(\begin{bmatrix} 3 & 3 \\ 1 & 1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right)$$

The parametric form of this equation is $x = -y$, so an eigenvector is $\vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. It follows that $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} 5 & 0 \\ 0 & 1 \end{bmatrix}$$

The diagonal matrix D scales the x-coordinate by 5 and the y-coordinate by 1. So any point on the horizontal line $y = c$ will remain on that line and be pushed five times as far away from the line $x = 0$

Figure 3.3.1: Dynamics of the matrix D

The matrix A does the same thing, in the $\vec{v}_1, \vec{v}_2$-coordinate system: multiplying a vector by A scales the $\vec{v}_1$-coordinate by 5 and the $\vec{v}_2$-coordinate by 1. Therefore, A moves vectors five times further along the line spanned by $\vec{v}_1$, and the same amount in the direction of the line spanned by $\vec{v}_2$:

Figure 3.3.2: Dynamics of the matrix A

For $\lambda_1 > 1$, we stretch our vectors away from the origin in the direction of $\vec{v}_1$ by a factor of λ_1.

Since $\lambda_2 = 1$, we have no stretch or squeeze in that direction, which is why we stay the same distance from the origin in the direction of $\vec{v}_2$.

Illustration 3.3.5: Eigenvalues $|\lambda_1| < 1$, $\lambda_2 = -1$

Let's describe how the matrix $A = \begin{bmatrix} \frac{1}{2} & -\frac{3}{2} \\ 0 & -1 \end{bmatrix}$ acts on the plane.

First we diagonalize A. The characteristic polynomial is

$$f(\lambda) = \begin{vmatrix} \frac{1}{2} - \lambda & -\frac{3}{2} \\ 0 & -1 - \lambda \end{vmatrix} = \left(\frac{1}{2} - \lambda\right)(-1 - \lambda)$$

We compute the $\frac{1}{2}$-eigenspace:

$$E_{\frac{1}{2}} = \text{Null}(A - \frac{1}{2}I) = \text{Null}\left(\begin{bmatrix} 0 & -\frac{3}{2} \\ 0 & -\frac{3}{2} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$$

So our first eigenvector is $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

For the -1-eigenspace, we have:

$$E_{-1} = \text{Null}(A + I) = \text{Null}\left(\begin{bmatrix} \frac{3}{2} & -\frac{3}{2} \\ 0 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$$

So our second eigenvector is $\vec{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. It follows that $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & -1 \end{bmatrix}$$

The diagonal matrix D scales the x-coordinate by $\frac{1}{2}$ and the y-coordinate by -1. So any point on the horizontal line $y = c$ will reflect onto the line $y = -c$, and be pushed to be half as far away from the line $x = 0$

Figure 3.3.3: Dynamics of the matrix D

The matrix A does the same thing, in the $\vec{v}_1, \vec{v}_2$-coordinate system: multiplying a vector by A scales the $\vec{v}_1$-coordinate by $\frac{1}{2}$ and the $\vec{v}_2$-coordinate by -1. Therefore, A moves vectors half as far along the line spanned by $\vec{v}_1$, and to the opposite direction on the line spanned by $\vec{v}_2$:

Figure 3.3.4: Dynamics of the matrix A

For $0 < \lambda_1 < 1$, we squeeze our vectors closer to the origin in the direction of $\vec{v}_1$ by a factor of λ_1.

Since $\lambda_2 = -1$, we reflect in that direction and stay the same distance from the origin in the direction of $\vec{v}_2$.

Illustration 3.3.8: Eigenvalues $\lambda_1 < 0,\ \lambda_2 = 0$

Let's Describe how the matrix $A = \begin{bmatrix} -1 & -1 \\ -2 & -2 \end{bmatrix}$ acts on the plane.

First we diagonalize A. The characteristic polynomial is

$$f(\lambda) = \begin{vmatrix} -1-\lambda & -1 \\ -2 & -2-\lambda \end{vmatrix} = (-1-\lambda)(-2-\lambda) - 2 = \lambda^2 + 3\lambda = \lambda(\lambda+3)$$

We compute the -3-eigenspace:

$$E_{-3} = \text{Null}(A+3I) = \text{Null}\left(\begin{bmatrix} 2 & -1 \\ -2 & 1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$$

So our first eigenvector is $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

For the 0-eigenspace, we have:

$$E_0 = \text{Null}(A) = \text{Null}\left(\begin{bmatrix} -1 & -1 \\ -2 & -2 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$$

So our second eigenvector is $\vec{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. It follows that $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \qquad D = \begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}$$

The diagonal matrix D scales the x-coordinate by -3 and the y-coordinate by 0. So any point on the vertical line $x = c$ will reflect and stretch onto the line $x = -3c$, and the y value will be sent to 0

Figure 3.3.5: Dynamics of the matrix D

The matrix A does the same thing, in the $\vec{v}_1, \vec{v}_2$-coordinate system: multiplying a vector by A scales the $\vec{v}_1$-coordinate by -3 and the $\vec{v}_2$-coordinate becomes 0. Therefore, A reflects and stretches vectors three times as far along the line spanned by $\vec{v}_1$, and pulls them onto the $\vec{v}_2$ axis.

Figure 3.3.6: Dynamics of the matrix A

For $\lambda_1 < 0$, we reflect across the $\vec{v}_1$ axis AND stretch or squeeze our vectors in that direction by a factor of $|\lambda_1|$.

Since $\lambda_2 = 0$, we lose all distance in the $\vec{v}_2$ direction, projecting the vector onto the $\vec{v}_1$ axis.

3.3: Geometry of Diagonalizable Matrices is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.4: Diagonalization** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.3: Practice Set

Exercise 3.3.1a

Consider the vectors $\vec{b}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and $\vec{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Find a matrix A which satisfies $A\vec{b}_1 = 7\vec{b}_1$ and $A\vec{b}_2 = 6\vec{b}_2$

Answer

A would have a basis of eigenvectors and be diagonalizable with $A = PDP^{-1}$ for the matrix $P = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 7 & 0 \\ 0 & 6 \end{bmatrix}$

$$A = PDP^{-1} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 7 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7 & -1 \\ 0 & 6 \end{bmatrix}$$

Exercise 3.3.1b

Consider the vectors $\vec{b}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{b}_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. Find a matrix C which satisfies $C\vec{b}_1 = -2\vec{b}_1$ and $C\vec{b}_2 = 3\vec{b}_2$

Answer

C would have a basis of eigenvectors and be diagonalizable with $C = PDP^{-1}$ for the matrix $P = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix}$ above and $D = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$

$$C = PDP^{-1} = \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{bmatrix} = \begin{bmatrix} \frac{16}{7} & -\frac{15}{7} \\ -\frac{10}{7} & -\frac{9}{7} \end{bmatrix}$$

Exercise 3.3.2a

Describe the geometry of the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x+y \\ x-y \end{bmatrix}$

Answer

$T(\vec{x}) = A\vec{x}$ for $A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

A has characteristic polynomial $p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ 1 & -1-\lambda \end{vmatrix} = \lambda^2 - 2$

The roots of this polynomial are our eigenvalues: $\lambda = \pm\sqrt{2}$

$$E_{\sqrt{2}} = \text{Null}\left(A - \sqrt{2}I\right) = \text{Null}\left(\begin{bmatrix} 1-\sqrt{2} & 1 \\ 1 & -1-\sqrt{2} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -1-\sqrt{2} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1+\sqrt{2} \\ 1 \end{bmatrix}\right\}$$

$$E_{-\sqrt{2}} = \text{Null}\left(A + \sqrt{2}I\right) = \text{Null}\left(\begin{bmatrix} 1+\sqrt{2} & 1 \\ 1 & -1+\sqrt{2} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -1+\sqrt{2} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1-\sqrt{2} \\ 1 \end{bmatrix}\right\}$$

So T **stretches by** $\sqrt{2}$ the component of $\vec{x}$ in the direction of $\begin{bmatrix} 1+\sqrt{2} \\ 1 \end{bmatrix}$,

and **reflects** then **stretches by** $\sqrt{2}$ the component of $\vec{x}$ in the direction of $\begin{bmatrix} 1-\sqrt{2} \\ 1 \end{bmatrix}$

Exercise 3.3.2b

Describe the geometry of the transformation $T:\mathbb{R}^2\to\mathbb{R}^2$ with $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)=\begin{bmatrix} y \\ 4x \end{bmatrix}$

Answer

$T(\vec{x})=A\vec{x}$ for $A=\begin{bmatrix} 0 & 1 \\ 4 & 0 \end{bmatrix}$

A has characteristic polynomial $p(\lambda)=\det(A-\lambda I)=\begin{vmatrix} 0-\lambda & 1 \\ 4 & 0-\lambda \end{vmatrix}=\lambda^2-4$

The roots of this polynomial are our eigenvalues: $\lambda=\pm 2$

$E_2=\text{Null}\,(A-2I)=\text{Null}\left(\begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix}\right)=\text{Null}\left(\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}\right)=\text{Span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$

$E_{-2}=\text{Null}\,(A+2I)=\text{Null}\left(\begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix}\right)=\text{Null}\left(\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}\right)=\text{Span}\left\{\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right\}$

So T **stretches by** 2 the component of $\vec{x}$ in the direction of $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$,

and **reflects** then **stretches by** 2 the component of $\vec{x}$ in the direction of $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$

Exercise 3.3.3a

Find a diagonal matrix D and invertible matrix P which satisfies $A=PDP^{-1}$ for the matrix $A=\begin{bmatrix} 9 & 2 \\ 12 & 11 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$p(\lambda)=\det(A-\lambda I)=\begin{vmatrix} 9-\lambda & 2 \\ 12 & 11-\lambda \end{vmatrix}=\lambda^2-20\lambda+75$

Therefore we have eigenvalues $\lambda=5$ with AM=1, and $\lambda=15$ with AM=1.

$E_5=\text{Null}\,(A-5I)=\text{Null}\left(\begin{bmatrix} 4 & 2 \\ 12 & 6 \end{bmatrix}\right)=\text{Null}\left(\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 0 \end{bmatrix}\right)=\text{Span}\left\{\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right\}$

$E_{15}=\text{Null}\,(A-15I)=\text{Null}\left(\begin{bmatrix} -6 & 2 \\ 12 & -4 \end{bmatrix}\right)=\text{Null}\left(\begin{bmatrix} 1 & -\frac{1}{3} \\ 0 & 0 \end{bmatrix}\right)=\text{Span}\left\{\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right\}$

Since AM=GM for all eigenvalues, we know A is diagonalizable with $D=\begin{bmatrix} 5 & 0 \\ 0 & 15 \end{bmatrix}$ and $P=\begin{bmatrix} -1 & 1 \\ 2 & 3 \end{bmatrix}$

Exercise 3.3.3b

Find a diagonal matrix D and invertible matrix P which satisfies $A=PDP^{-1}$ for the matrix $A=\begin{bmatrix} -15 & -20 \\ 12 & 16 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$p(\lambda)=\det(A-\lambda I)=\begin{vmatrix} -15-\lambda & -20 \\ 12 & 16-\lambda \end{vmatrix}=\lambda^2-\lambda$

Therefore we have eigenvalues $\lambda=0$ with AM=1, and $\lambda=1$ with AM=1.

$E_0=\text{Null}\,(A-0I)=\text{Null}\left(\begin{bmatrix} -15 & -20 \\ 12 & 16 \end{bmatrix}\right)=\text{Null}\left(\begin{bmatrix} 1 & \frac{4}{3} \\ 0 & 0 \end{bmatrix}\right)=\text{Span}\left\{\begin{bmatrix} -4 \\ 3 \end{bmatrix}\right\}$

$$E_1 = \text{Null}\,(A - 1I) = \text{Null}\left(\begin{bmatrix} -16 & -20 \\ 12 & 15 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & \frac{5}{4} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -5 \\ 4 \end{bmatrix}\right\}$$

Since AM=GM for all eigenvalues, we know A is diagonalizable with $D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ and $P = \begin{bmatrix} -4 & -5 \\ 3 & 4 \end{bmatrix}$

Exercise 3.3.3c

Find a diagonal matrix D and invertible matrix P which satisfies $A = PDP^{-1}$ for the matrix $A = \begin{bmatrix} -5 & -6 & 6 \\ 1 & 4 & -2 \\ -2 & 0 & 2 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -5-\lambda & -6 & 6 \\ 1 & 4-\lambda & -2 \\ -2 & 0 & 2-\lambda \end{vmatrix} = -\lambda^3 + \lambda^2 + 4\lambda - 4$$

Therefore we have eigenvalues $\lambda = 1$ with AM=1, $\lambda = 2$ with AM=1, and $\lambda = -2$ with AM=1.

$$E_1 = \text{Null}\,(A - 1I) = \text{Null}\left(\begin{bmatrix} -6 & -6 & 6 \\ 1 & 3 & -2 \\ -2 & 0 & 1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}\right\}$$

$$E_2 = \text{Null}\,(A - 2I) = \text{Null}\left(\begin{bmatrix} -7 & -6 & 6 \\ 1 & 2 & -2 \\ -2 & 0 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}$$

$$E_{-2} = \text{Null}\,(A + 2I) = \text{Null}\left(\begin{bmatrix} -3 & -6 & 6 \\ 1 & 6 & -2 \\ -2 & 0 & 4 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}\right\}$$

Since AM=GM for all eigenvalues, we know A is diagonalizable with $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}$

Exercise 3.3.3d

Find a diagonal matrix D and invertible matrix P which satisfies $A = PDP^{-1}$ for the matrix $A = \begin{bmatrix} -11 & 24 & -12 \\ -8 & 17 & -8 \\ -8 & 16 & -7 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -11-\lambda & 24 & -12 \\ -8 & 17-\lambda & -8 \\ -8 & 16 & -7-\lambda \end{vmatrix} = -\lambda^3 - \lambda^2 + 5\lambda - 3$$

Therefore we have eigenvalues $\lambda = 1$ with AM=2, and $\lambda = -3$ with AM=1.

$$E_1 = \text{Null}\,(A - 1I) = \text{Null}\left(\begin{bmatrix} -12 & 24 & -12 \\ -8 & 16 & -8 \\ -8 & 16 & -8 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$$

$$E_{-3} = \text{Null}\,(A + 3I) = \text{Null}\left(\begin{bmatrix} -8 & 24 & -12 \\ -8 & 20 & -8 \\ -8 & 16 & -4 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -\frac{3}{2} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}\right\}$$

Since AM=GM for all eigenvalues, we know A is diagonalizable with $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix}$ and $P = \begin{bmatrix} 2 & -1 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 2 \end{bmatrix}$

Exercise 3.3.3e

Find a diagonal matrix D and invertible matrix P which satisfies $A = PDP^{-1}$ for the matrix $A = \begin{bmatrix} 2 & -2 & 4 \\ 0 & 0 & 0 \\ -1 & 1 & -2 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & -2 & 4 \\ 0 & 0-\lambda & 0 \\ -1 & 1 & -2-\lambda \end{vmatrix} = -\lambda^3$$

Therefore we have eigenvalue $\lambda = 0$ with AM=3.

$$E_0 = \text{Null}\,(A - 0I) = \text{Null}\left(\begin{bmatrix} 2 & -2 & 4 \\ 0 & 0 & 0 \\ -1 & 1 & -2 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}\right\}$$

Since AM≠GM for $\lambda = 0$, we know A is not diagonalizable, so no such D and P exist.

Exercise 3.3.4a

Find the general form of A^k for $A = \begin{bmatrix} -5 & 4 \\ -8 & 7 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -5-\lambda & 4 \\ -8 & 7-\lambda \end{vmatrix} = \lambda^2 - 2\lambda - 3$$

Therefore we have eigenvalues $\lambda = -1$ with AM=1, and $\lambda = 3$ with AM=1.

$$E_{-1} = \text{Null}\,(A + 1I) = \text{Null}\left(\begin{bmatrix} -4 & 4 \\ -8 & 8 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$$

$$E_3 = \text{Null}\,(A - 3I) = \text{Null}\left(\begin{bmatrix} -8 & 4 \\ -8 & 4 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$$

Since AM=GM for all eigenvalues, we know A is diagonalizable with $D = \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$

Therefore $A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} (-1)^k & 0 \\ 0 & 3^k \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2(-1)^k - 3^k & -(-1)^k + 3^k \\ 2(-1)^k - 2\cdot 3^k & -(-1)^k + 2\cdot 3^k \end{bmatrix}$

Exercise 3.3.4b

Find a matrix B satisfying $B^2 = A$ for $A = \begin{bmatrix} -2 & 3 & 3 \\ -4 & 4 & 4 \\ -2 & 3 & 3 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -2-\lambda & 3 & 3 \\ -4 & 4-\lambda & 4 \\ -2 & 3 & 3-\lambda \end{vmatrix} = -\lambda^3 + 5\lambda^2 - 4\lambda$$

Therefore we have eigenvalues $\lambda = 0$ with AM=1, $\lambda = 1$ with AM=1, and $\lambda = 4$ with AM=1.

$$E_0 = \text{Null}\,(A - 0I) = \text{Null}\left(\begin{bmatrix} -2 & 3 & 3 \\ -4 & 4 & 4 \\ -2 & 3 & 3 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right\}$$

$$E_1 = \text{Null}\,(A - 1I) = \text{Null}\left(\begin{bmatrix} -3 & 3 & 3 \\ -4 & 3 & 4 \\ -2 & 3 & 2 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\}$$

$$E_4 = \text{Null}\,(A - 4I) = \text{Null}\left(\begin{bmatrix} -6 & 3 & 3 \\ -4 & 0 & 4 \\ -2 & 3 & -1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$$

Since AM=GM for all eigenvalues, we know A is diagonalizable with $D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix}$ and $P = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

Therefore one solution is $B = A^{\frac{1}{2}} = PD^{\frac{1}{2}}P^{-1} = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 0^{\frac{1}{2}} & 0 & 0 \\ 0 & 1^{\frac{1}{2}} & 0 \\ 0 & 0 & 4^{\frac{1}{2}} \end{bmatrix}\begin{bmatrix} -1 & 0 & 1 \\ 2 & -1 & -1 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ -2 & 2 & 2 \\ 0 & 1 & 1 \end{bmatrix}$

Other possible solutions are $\begin{bmatrix} 0 & -1 & -1 \\ 2 & -2 & -2 \\ 0 & -1 & -1 \end{bmatrix}, \begin{bmatrix} 4 & -3 & -3 \\ 2 & -2 & -2 \\ 4 & -3 & -3 \end{bmatrix}, \begin{bmatrix} -4 & 3 & 3 \\ -2 & 2 & 2 \\ -4 & 3 & 3 \end{bmatrix}$

? Exercise 3.3.4c

Find a matrix C satisfying $C^3 = A$ for $A = \begin{bmatrix} -17 & 0 & 18 \\ -3 & -1 & -1 \\ -9 & 0 & 10 \end{bmatrix}$

Answer

We start with the characteristic polynomial:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -17 - \lambda & 0 & 18 \\ -3 & -1 - \lambda & -1 \\ -9 & 0 & 10 - \lambda \end{vmatrix} = -\lambda^3 - 8\lambda^2 + \lambda + 8$$

Therefore we have eigenvalues $\lambda = 1$ with AM=1, $\lambda = -1$ with AM=1, and $\lambda = -8$ with AM=1.

$$E_1 = \text{Null}\,(A - 1I) = \text{Null}\left(\begin{bmatrix} -18 & 0 & 18 \\ -3 & -2 & -1 \\ -9 & 0 & 9 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right\}$$

$$E_{-1} = \text{Null}\,(A + I) = \text{Null}\left(\begin{bmatrix} -16 & 0 & 18 \\ -3 & 0 & -1 \\ -9 & 0 & 11 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$$

$$E_{-8} = \text{Null}\,(A + 8I) = \text{Null}\left(\begin{bmatrix} -9 & 0 & 18 \\ -3 & 7 & -1 \\ -9 & 0 & 18 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}\right\}$$

Since AM=GM for all eigenvalues, we know A is diagonalizable with $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -8 \end{bmatrix}$ and $P = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$

Therefore the solution is $C = A^{\frac{1}{3}} = PD^{\frac{1}{3}}P^{-1} = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1^{\frac{1}{3}} & 0 & 0 \\ 0 & (-1)^{\frac{1}{3}} & 0 \\ 0 & 0 & (-8)^{\frac{1}{3}} \end{bmatrix}\begin{bmatrix} -1 & 0 & 2 \\ -3 & 1 & 5 \\ 1 & 0 & -1 \end{bmatrix}$

$$= \begin{bmatrix} -5 & 0 & 6 \\ 3 & -1 & -7 \\ -3 & 0 & 4 \end{bmatrix}$$

Exercise 3.3.5

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If A is a 2×2 matrix with two distinct real eigenvalues, then A is diagonalizable.

(b) If A is a 2×2 matrix with characteristic polynomial $p(\lambda) = (\lambda - r)^2$, then A is not diagonalizable.

(c) If D is an $n \times n$ diagonal matrix, then it has n linearly independent eigenvectors.

(d) If A is diagonalizable, then it is invertible.

(e) If A has characteristic polynomial $p(\lambda) = \lambda^2 + a\lambda + b$ and $a^2 > 4b$, then A is diagonalizable.

(f) If A is diagonalizable, then it has n distinct real eigenvalues.

(g) If A has characteristic polynomial $q(\lambda)$, then the matrix $B = 2A$ has characteristic polynomial $p(\lambda) = q\left(\frac{\lambda}{2}\right)$

(h) If A^2 is diagonalizable, then A is diagonalizable.

Answer

(a) **Always True**

Since each eigenvalue must have algebraic multiplicity 1, they must also have geometric multiplicity 1. Thus A is diagonalizable.

(b) **Sometimes True Sometimes False**

When $A = rI$ it is already diagonal, and therefore diagonalizable. However, if $A \neq rI$, then E_r will have dimension 1, and thus will not be diagonalizable.

(c) **Always True**

They are exactly $\{\vec{e}_1, \ldots, \vec{e}_n\}$.

(d) **Sometimes True Sometimes False**

When 0 is not an eigenvalue, this is true. But when 0 is an eigenvalue, we know A will not be invertible.

(e) **Always True**

There will be two distinct real roots, so it must be diagonalizable.

(f) **Sometimes True Sometimes False**

As in many of the examples above we had this exact scenario, so we know it can be true. But, we also know that some matrices can be diagonalizable with repeated eigenvalues, such as the diagonal identity matrix with only one eigenvalue $\lambda = 1$.

(g) **Always False**

The correct equation must be $p(\lambda) = \det(B - \lambda I) = \det(2A - \lambda I) = 2^n \det\left(A - \frac{\lambda}{2} I\right) = 2^n q\left(\frac{\lambda}{2}\right)$

(h) **Sometimes True Sometimes False**

When $A = I$ this holds to be true. But when $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ we see $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ is diagonalizable, but A is not.

3.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.4: Complex Eigenvalues

Objectives

1. Learn how the eigenvalues of a diagonalizable matrix affect the geometry of a linear transformation
2. Learn to find complex eigenvalues and eigenvectors of a matrix.
3. Learn to recognize a rotation-scaling matrix, and compute how much the matrix rotates and scales.
4. Understand the geometry of 2×2 matrices with a complex eigenvalue.
5. *Recipes:* a 2×2 matrix with a complex eigenvalue is similar to a rotation-scaling matrix, the eigenvector trick for 2×2 matrices.
6. *Pictures:* the geometry of matrices with a complex eigenvalue.
7. *Theorem:* the rotation-scaling theorem
8. *Vocabulary word:* ***rotation-scaling matrix***.

In our last section, we saw that an $n \times n$ matrix whose characteristic polynomial has n distinct real roots will have an algebraic multiplicity of 1 for each value λ_i. By definition, each geometric multiplicity will also be exactly 1 in this case, and the matrix is *diagonalizable*: it is similar to a diagonal matrix. Now we explore what the numerical values of each eigenvalue represent geometrically.

This page titled 3.4: Complex Eigenvalues is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **5.5: Complex Eigenvalues** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.4: Complex Eigenvalues and Eigenvectors

Since the characteristic polynomial does not have to have real roots, we discuss the other possibility: *complex* roots. It turns out that such a matrix is similar (in the 2×2 case) to a *rotation-scaling matrix*, which is also relatively easy to understand. As a consequence of the fundamental theorem of algebra applied to the characteristic polynomial, we see that:

Note 3.4.1

Every $n \times n$ matrix has exactly n complex eigenvalues, counted with multiplicity.

We can compute a corresponding (complex) eigenvector in exactly the same way as before: by row reducing the matrix $A - \lambda I$. Now, however, we have to do arithmetic with complex numbers.

Illustration 3.4.1: A 2×2 Matrix

We will find the complex eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

The characteristic polynomial of A is

$$f(\lambda) = \begin{vmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 + 1 = \lambda^2 - 2\lambda + 2$$

The roots of this polynomial are

$$\lambda = \frac{2 \pm \sqrt{4-8}}{2} = 1 \pm i$$

First we compute an eigenvector for $\lambda = 1 + i$. We have

$$A - (1+i)I = \begin{bmatrix} 1-(1+i) & -1 \\ 1 & 1-(1+i) \end{bmatrix} = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}$$

Now we row reduce, noting that the second row is i times the first:

$$\begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \xrightarrow{R_2 = R_2 - iR_1} \begin{bmatrix} -i & -1 \\ 0 & 0 \end{bmatrix} \xrightarrow{R_1 = R_1 \div -i} \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}$$

The parametric form is $x = iy$, so that an eigenvector that spans E_{1+i} is $\vec{v}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}$

Next we compute an eigenvector for $\lambda = 1 - i$. We have

$$A - (1-i)I = \begin{bmatrix} 1-(1-i) & -1 \\ 1 & 1-(1-i) \end{bmatrix} = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix}$$

Now we row reduce, noting that the second row is $-i$ times the first:

$$\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \xrightarrow{R_2 = R_2 + iR_1} \begin{bmatrix} i & -1 \\ 0 & 0 \end{bmatrix} \xrightarrow{R_1 = R_1 \div i} \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$$

The parametric form is $x = -iy$, so that an eigenvector that spans E_{1-i} is $\vec{v}_2 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$

We can verify our answers:

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} i-1 \\ i+1 \end{bmatrix} = (1+i)\begin{bmatrix} i \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -i \\ 1 \end{bmatrix} = \begin{bmatrix} -i-1 \\ -i+1 \end{bmatrix} = (1-i)\begin{bmatrix} -i \\ 1 \end{bmatrix}$$

If A is a matrix with real entries, then its characteristic polynomial has real coefficients, so its complex eigenvalues come in conjugate pairs. In the first example, we notice that

$$\lambda_1 = 1+i \text{ has an eigenvector } \vec{v}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}$$
$$\lambda_2 = 1-i \text{ has an eigenvector } \vec{v}_2 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

In any matrix, an eigenvector for the conjugate eigenvalue is simply the *conjugate eigenvector* (the eigenvector obtained by conjugating each entry of the first eigenvector). This is always true. Indeed, if $A\vec{v} = \lambda\vec{v}$ then if we define $z' = a - bi$ to denote the conjugate of $z = a + bi$:

$$A\vec{v}' = (A\vec{v})' = (\lambda\vec{v})' = (\lambda)'\vec{v}'$$

which exactly says that $\vec{v}'$ is an eigenvector of A with eigenvalue λ'.

Note 3.4.2

Let A be a matrix with real entries. If λ is a complex eigenvalue with eigenvector $\vec{v}$, then λ' is a complex eigenvalue with eigenvector $\vec{v}'$. In other words, both eigenvalues and eigenvectors come in conjugate pairs.

Since it can be tedious to divide by complex numbers while row reducing, it is useful to learn the following trick, which works equally well for matrices with real entries.

Note 3.4.3: Eigenvector Trick for 2×2 Matrices

Let A be a 2×2 matrix, and let λ be a (real or complex) eigenvalue. Then

$$A - \lambda I = \begin{bmatrix} z & w \\ \star & \star \end{bmatrix} \implies \begin{bmatrix} -w \\ z \end{bmatrix} \text{ is an eigenvector with eigenvalue } \lambda$$

assuming the first row of $A - \lambda I$ is nonzero.

Indeed, since λ is an eigenvalue, we know that $A - \lambda I$ is not an invertible matrix. It follows that the rows are collinear (otherwise the determinant is nonzero), so that the second row is automatically a (complex) multiple of the first:

$$\begin{bmatrix} z & w \\ \star & \star \end{bmatrix} = \begin{bmatrix} z & w \\ cz & cw \end{bmatrix}$$

It is obvious that $\begin{bmatrix} -w \\ z \end{bmatrix}$ is in the null space of this matrix, as is $\begin{bmatrix} w \\ -z \end{bmatrix}$, for that matter. Note that we never had to compute the second row of $A - \lambda I$, let alone row reduce!

Example 3.4.1: A 2×2 matrix, the easy way

Find the complex eigenvalues and eigenvectors of the matrix

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Solution

Since the characteristic polynomial of a 2×2 matrix A is $f(\lambda) = \lambda^2 - 2\lambda + 2$, its roots are

$$\lambda = 1 \pm i$$

To find an eigenvector with eigenvalue $1+i$, we compute

$$A-(1+i)I = \begin{bmatrix} -i & -1 \\ \star & \star \end{bmatrix} \xrightarrow{\text{eigenvector}} \boxed{\vec{v}_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix}}$$

The eigenvector for the conjugate eigenvalue is the complex conjugate:

$$\boxed{\vec{v}_2 = \vec{v}_1' = \begin{bmatrix} 1 \\ i \end{bmatrix}}$$

They may look different, but these vectors are simply scalar multiples of our answers when we did this example the hard way:

$$-i\begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -i \end{bmatrix} \qquad i\begin{bmatrix} -i \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

The most important examples of matrices with complex eigenvalues are rotation-scaling matrices, i.e., scalar multiples of rotation matrices.

Definition: Rotation-Scaling matrix

A ***rotation-scaling matrix*** is a 2×2 matrix of the form

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

where a and b are real numbers, not both equal to zero.

The following proposition justifies the name.

Proposition 3.4.1

Let $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ be a rotation-scaling matrix. Then:

1. A is a product of a rotation matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{with a scaling matrix} \quad \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$$

2. The scaling factor r is

$$r = \sqrt{\det(A)} = \sqrt{a^2+b^2}.$$

3. The rotation angle θ is the counterclockwise angle from the positive x-axis to the vector $\begin{bmatrix} a \\ b \end{bmatrix}$:

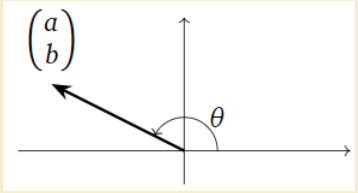


Figure 3.4.1

The eigenvalues of A are $\lambda = a \pm bi$

Proof

Set $r = \sqrt{\det(A)} = \sqrt{a^2+b^2}$. The point $\left(\frac{a}{r}, \frac{b}{r}\right)$ has the property that

$$\left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2 = \frac{a^2+b^2}{r^2} = 1$$

In other words $\left(\frac{a}{r}, \frac{b}{r}\right)$ lies on the unit circle. Therefore, it has the form $(\cos\theta, \sin\theta)$, where θ is the counterclockwise angle from the positive x-axis to the vector $\begin{bmatrix} \frac{a}{r} \\ \frac{b}{r} \end{bmatrix}$, or since it is on the same line, to $\begin{bmatrix} a \\ b \end{bmatrix}$:

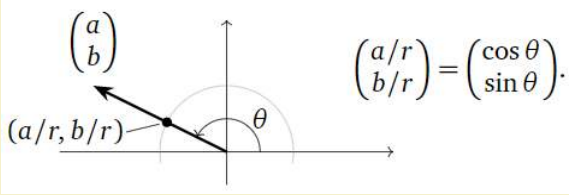


Figure 3.4.2

It follows that

$$A = r\begin{bmatrix} \frac{a}{r} & -\frac{b}{r} \\ \frac{b}{r} & \frac{a}{r} \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

as desired.

For the last statement, we compute the eigenvalues of A as the roots of the characteristic polynomial:

$$f(\lambda) = \begin{vmatrix} a-\lambda & -b \\ b & a-\lambda \end{vmatrix} = (a-\lambda)^2 + b^2 = \lambda^2 - 2a\lambda + (a^2+b^2)$$

Now the quadratic formula yields:

$$\lambda = \frac{2a \pm \sqrt{4a^2 - 4(a^2+b^2)}}{2} = a \pm bi$$

Geometrically, a rotation-scaling matrix does exactly what the name says: it rotates and scales (in either order).

Illustration 3.4.2: A rotation-scaling matrix

What does the matrix $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ do geometrically?

This is a rotation-scaling matrix with $a = b = 1$. Therefore, it scales by a factor of $\sqrt{\det(A)} = \sqrt{2}$ and rotates counterclockwise by $\theta = \tan^{-1}\left(\frac{b}{a}\right) = \tan^{-1}(1) = 45°$:

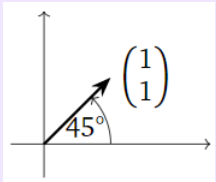


Figure 3.4.3

Here is a picture of A:

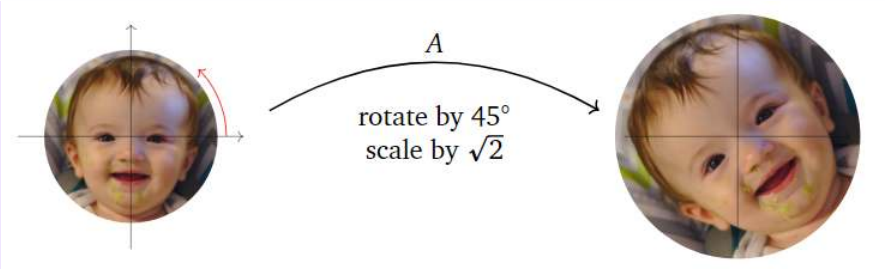


Figure 3.4.4

Illustration 3.4.3: A rotation-scaling matrix

What does the matrix $A = \begin{bmatrix} -\sqrt{3} & -1 \\ 1 & -\sqrt{3} \end{bmatrix}$ do geometrically?

This is a rotation-scaling matrix with $a = -\sqrt{3}$ and $b = 1$. Therefore, it scales by a factor of $\sqrt{\det(A)} = \sqrt{3+1} = 2$ and rotates counterclockwise by the angle θ in the picture:

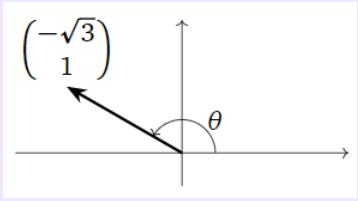


Figure 3.4.5

Since this is in the second quadrant, we have to use some trigonometry to find θ:

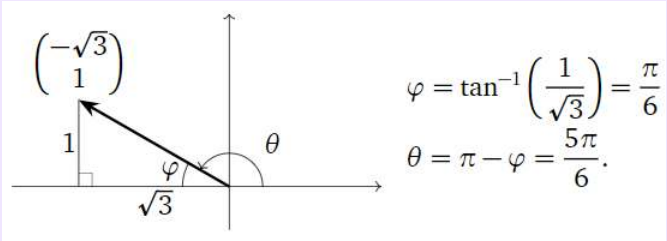


Figure 3.4.6

Therefore, A rotates counterclockwise by $\frac{5\pi}{6}$ and scales by a factor of 2.

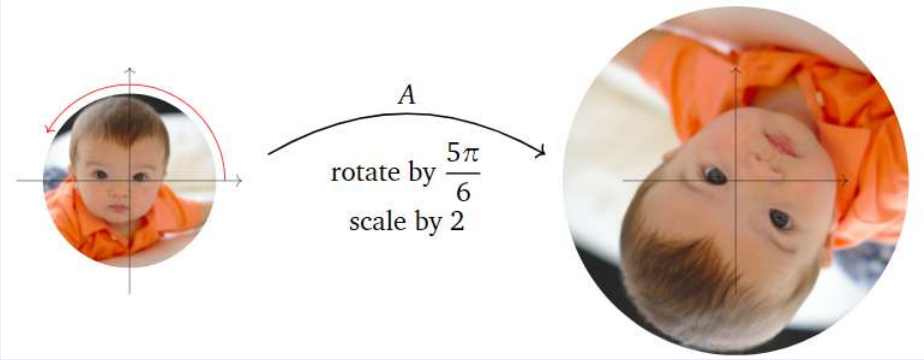


Figure 3.4.7

The matrix in Illustration 3.4.3 has second column $\begin{bmatrix} -\sqrt{3} \\ 1 \end{bmatrix}$, which is rotated counterclockwise from the positive x-axis by an angle of $\frac{5\pi}{6}$. This rotation angle is *not* equal to $\tan^{-1}\left(-\frac{1}{\sqrt{3}}\right) = -\frac{\pi}{6}$ The problem is that the arctan function always outputs values between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$: it does not account for points in the second or third quadrants.

Alternatively, we could have observed that $\begin{bmatrix} -\sqrt{3} \\ 1 \end{bmatrix}$ lies in the second quadrant, so that the angle θ in question is

$$\theta = \tan^{-1}\left(-\frac{1}{\sqrt{3}}\right) + \pi = -\frac{\pi}{6} + \pi = \frac{5\pi}{6}$$

Note 3.4.4

When finding the rotation angle of a vector $\begin{bmatrix} a \\ b \end{bmatrix}$, do not blindly compute $\tan^{-1}\left(\frac{b}{a}\right)$, since this will give the wrong answer when $\begin{bmatrix} a \\ b \end{bmatrix}$ is in the second or third quadrant.

3.4: Complex Eigenvalues and Eigenvectors is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.5: Complex Eigenvalues** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.4: Geometry of Matrices with Complex Eigenvalues

Let A be a 2×2 matrix with a complex, non-real eigenvalue λ. Then A also has the eigenvalue $\lambda' \neq \lambda$. In particular, A has distinct eigenvalues, so it is diagonalizable using the complex numbers. We often like to think of our matrices as describing transformations of $\mathbb{R}^n$ (as opposed to $\mathbb{C}^n$). Because of this, the following construction is useful. It gives something like a diagonalization, except that all matrices involved have *real* entries.

Theorem 3.4.1: Rotation-Scaling Theorem

Let A be a 2×2 real matrix with a complex (non-real) eigenvalue $\lambda = a - bi$ (with $b > 0$), and let $\vec{v} = \begin{bmatrix} \alpha + i\beta \\ \gamma + i\delta \end{bmatrix}$ be a corresponding eigenvector.

Then $A = PBP^{-1}$ for

$$P = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

In particular, A is similar to a rotation-scaling matrix that scales by a factor of $|\lambda| = \sqrt{a^2 + b^2} = \sqrt{\det(B)}$

Proof

This proof is ultimately done by brute force. First we need to show that $\vec{v}_R = \begin{bmatrix} \alpha \\ \gamma \end{bmatrix}$ and $\vec{v}_I = \begin{bmatrix} \beta \\ \delta \end{bmatrix}$ are linearly independent vectors, since otherwise P is not invertible. If not, then there exist real numbers x, y, not both equal to zero, such that $x\vec{v}_R + y\vec{v}_I = \vec{0}$. Then

$$\begin{aligned} (y + ix)\vec{v} &= (y + ix)\left(\vec{v}_R + i\vec{v}_I\right) \\ &= y\vec{v}_R - x\vec{v}_I + i\left(x\vec{v}_R + y\vec{v}_I\right) \\ &= y\vec{v}_R - x\vec{v}_I + i\vec{0} \\ &= y\vec{v}_R - x\vec{v}_I \end{aligned}$$

Now, $(y + ix)\vec{v}$ is also an eigenvector of A with eigenvalue λ, as it is a scalar multiple of $\vec{v}$. But we just showed that $(y + ix)\vec{v}$ is a vector with real entries, and any *real* eigenvector of a real matrix has a real eigenvalue. Therefore, $\vec{v}_R$ and $\vec{v}_I$ must be linearly independent after all.

$$\begin{aligned} A\vec{v} = \lambda\vec{v} &= (a - bi)\begin{bmatrix} \alpha + i\beta \\ \gamma + i\delta \end{bmatrix} \\ &= \begin{bmatrix} (a\alpha + b\beta) + i(a\beta - b\alpha) \\ (a\gamma + b\delta) + i(a\delta - b\gamma) \end{bmatrix} \\ &= \begin{bmatrix} a\alpha + b\beta \\ a\gamma + b\delta \end{bmatrix} + i\begin{bmatrix} a\beta - b\alpha \\ a\delta - b\gamma \end{bmatrix} \end{aligned}$$

On the other hand, we have

$$A\left(\begin{bmatrix} \alpha \\ \gamma \end{bmatrix} + i\begin{bmatrix} \beta \\ \delta \end{bmatrix}\right) = A\vec{v}_R + iA\vec{v}_I$$

Matching real and imaginary parts gives

$$A\vec{v}_R = \begin{bmatrix} a\alpha + b\beta \\ a\gamma + b\delta \end{bmatrix} \qquad A\vec{v}_I = \begin{bmatrix} a\beta - b\alpha \\ a\delta - b\gamma \end{bmatrix}$$

Now we compute $PBP^{-1}\vec{v}_R$ and $PBP^{-1}\vec{v}_I$. Since $P\vec{e}_1 = \vec{v}_R$ and $P\vec{e}_2 = \vec{v}_I$, we have $P^{-1}\vec{v}_R = \vec{e}_1$ and $P^{-1}\vec{v}_I = \vec{e}_2$, so

$$PBP^{-1}\vec{v}_R = PB\vec{e}_1 = P\begin{bmatrix} a \\ b \end{bmatrix} = a\vec{v}_R + b\vec{v}_I$$
$$= a\begin{bmatrix} \alpha \\ \gamma \end{bmatrix} + b\begin{bmatrix} \beta \\ \delta \end{bmatrix} = \begin{bmatrix} a\alpha + b\beta \\ a\gamma + b\delta \end{bmatrix} = A\vec{v}_R$$
$$PBP^{-1}\vec{v}_I = PB\vec{e}_2 = P\begin{bmatrix} -b \\ a \end{bmatrix} = -b\vec{v}_R + a\vec{v}_I$$
$$= -b\begin{bmatrix} \alpha \\ \gamma \end{bmatrix} + a\begin{bmatrix} \beta \\ \delta \end{bmatrix} = \begin{bmatrix} a\beta - b\alpha \\ a\delta - b\gamma \end{bmatrix} = A\vec{v}_I$$

Therefore, $A\vec{v}_R = PBP^{-1}\vec{v}_R$ and $A\vec{v}_I = PBP^{-1}\vec{v}_I$

Since $\vec{v}_R$ and $\vec{v}_I$ are linearly independent, they form a basis for $\mathbb{R}^2$. Let $\vec{w}$ be any vector in $\mathbb{R}^2$, and write $\vec{w} = c\vec{v}_R + d\vec{v}_I$. Then

$$\begin{aligned} A\vec{w} &= A\big(c\vec{v}_R + d\vec{v}_I\big) \\ &= cA\vec{v}_R + dA\vec{v}_I \\ &= cPBP^{-1}\vec{v}_R + dPBP^{-1}\vec{v}_I \\ &= PBP^{-1}\big(c\vec{v}_R + d\vec{v}_I\big) \\ &= PBP^{-1}\vec{w} \end{aligned}$$

This proves that $A = PBP^{-1}$.

So for $\lambda = a - bi$ (with $b > 0$), the rotation-scaling matrix in question is

$$B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

Geometrically, the rotation-scaling theorem says that a 2×2 matrix with a complex eigenvalue behaves similarly to a rotation-scaling matrix. This is the "best" matrix to be similar to when you have complex eigenvalues (instead of a diagonal one when they are real).

Recipe: A 2×2 Matrix with a Complex Eigenvalue

Let A be a 2×2 real matrix.

1. Compute the characteristic polynomial
$$f(\lambda) = \det(A - \lambda I)$$
then compute its roots using the quadratic formula.
2. If the eigenvalues are complex, choose the one with *negative* imaginary part, and call it $\lambda = a - bi$ (with $b > 0$).
3. Find a corresponding (complex) eigenvector $\vec{v} = \begin{bmatrix} \alpha + i\beta \\ \gamma + i\delta \end{bmatrix}$ using the shortcut trick.
4. Then $A = PBP^{-1}$ for
$$P = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$
This scales by a factor of $|\lambda| = \sqrt{a^2 + b^2}$ and rotates counterclockwise by an angle θ where $\cos(\theta) = \dfrac{a}{|\lambda|}$ and $\sin(\theta) = \dfrac{b}{|\lambda|}$

Illustration3.4.4

What does the matrix $A = \begin{bmatrix} 2 & -1 \\ 2 & 0 \end{bmatrix}$ do geometrically?

The characteristic polynomial of A is

$$f(\lambda) = \begin{vmatrix} 2 - \lambda & -1 \\ 2 & 0 - \lambda \end{vmatrix} = -\lambda(2 - \lambda) + 2 = \lambda^2 - 2\lambda + 2$$

Using the quadratic formula, this gives $\lambda = 1 \pm i$

We choose the eigenvalue $\lambda = 1 - i$ (negative imaginary part) and find a corresponding eigenvector, using the trick, Note 3.4.3:

$$A - (1-i)I = \begin{bmatrix} 1+i & -1 \\ \star & \star \end{bmatrix} \xrightarrow{\text{eigenvector}} \vec{v} = \begin{bmatrix} 1 \\ 1+i \end{bmatrix}$$

Therefore, we have $A = PBP^{-1}$ for $P = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -(-1) \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$

The matrix B is the rotation-scaling matrix: it rotates counterclockwise by an angle of 45° and scales by a factor of $\sqrt{2}$. The matrix A does the same thing, but with respect to the $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$-coordinate system:

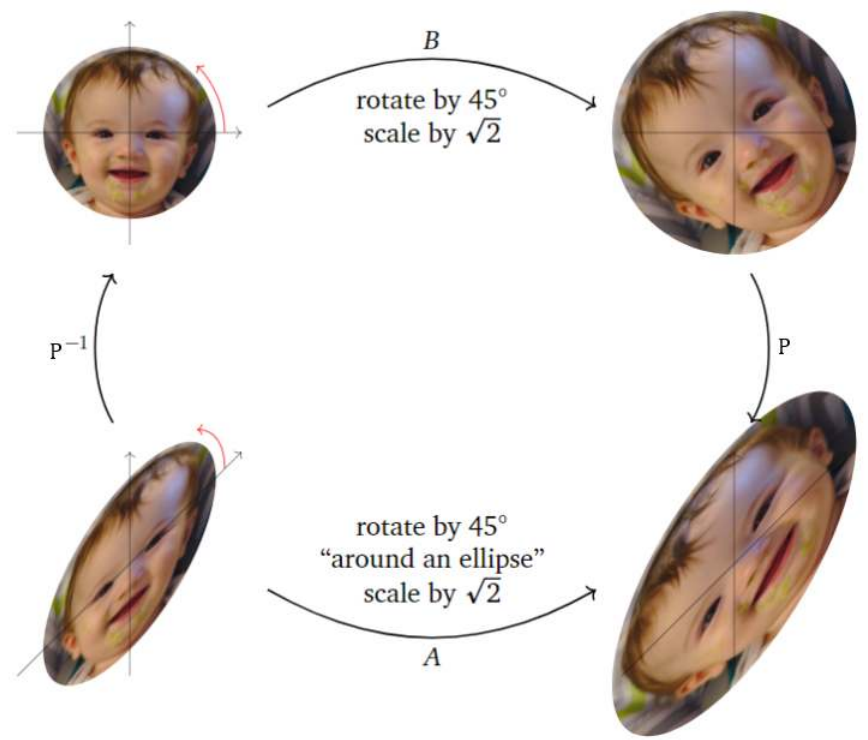


Figure 3.4.8

To summarize:

- B rotates around the circle centered at the origin and passing through $\vec{e}_1$ and $\vec{e}_2$, in the direction from $\vec{e}_1$ to $\vec{e}_2$, then scales by $\sqrt{2}$.
- A rotates around the ellipse centered at the origin and passing through $\vec{v}_R$ and $\vec{v}_I$, in the direction from $\vec{v}_R$ to $\vec{v}_I$, then scales by $\sqrt{2}$.

If instead we had chosen $\lambda' = 1 + i$ as our eigenvalue, then we would have found the eigenvector $\vec{v}' = \begin{bmatrix} 1 \\ 1-i \end{bmatrix}$. In this case we would have $A = P'B'(P')^{-1}$, where

$$P' = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \text{ and } B' = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

So, A is also similar to a *clockwise* rotation by 45°, followed by a scale by $\sqrt{2}$.

Note that this clockwise direction on the new axes is the same as the counterclockwise direction on the first set of axes we defined. The geometric result is the same!

Illustration 3.4.5

What does the matrix $A = \begin{bmatrix} -\sqrt{3}+1 & -2 \\ 1 & -\sqrt{3}-1 \end{bmatrix}$ do geometrically?

The characteristic polynomial of A is

$$f(\lambda) = \begin{vmatrix} -\sqrt{3}+1-\lambda & -2 \\ 1 & -\sqrt{3}-1-\lambda \end{vmatrix} = (-\sqrt{3}+1-\lambda)(-\sqrt{3}-1-\lambda)+2 = \lambda^2+2\sqrt{3}\lambda+4$$

Using the quadratic formula, this gives $\lambda = -\sqrt{3} \pm i$

We choose the eigenvalue $\lambda = -\sqrt{3}-i$ (negative imaginary part) and find a corresponding eigenvector, using the trick, Note 3.4.3:

$$A-(-\sqrt{3}-i)I = \begin{bmatrix} 1+i & -2 \\ \star & \star \end{bmatrix} \xrightarrow{\text{eigenvector}} \vec{v} = \begin{bmatrix} 2 \\ 1+i \end{bmatrix}$$

Therefore, we have $A = PBP^{-1}$ for $P = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} -\sqrt{3} & -(-1) \\ -1 & -\sqrt{3} \end{bmatrix} = \begin{bmatrix} -\sqrt{3} & 1 \\ -1 & -\sqrt{3} \end{bmatrix}$

The matrix B is the rotation-scaling matrix: it rotates counterclockwise by an angle of $\frac{5\pi}{6}$ and scales by a factor of 2. The matrix A does the same thing, but with respect to the $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$-coordinate system:

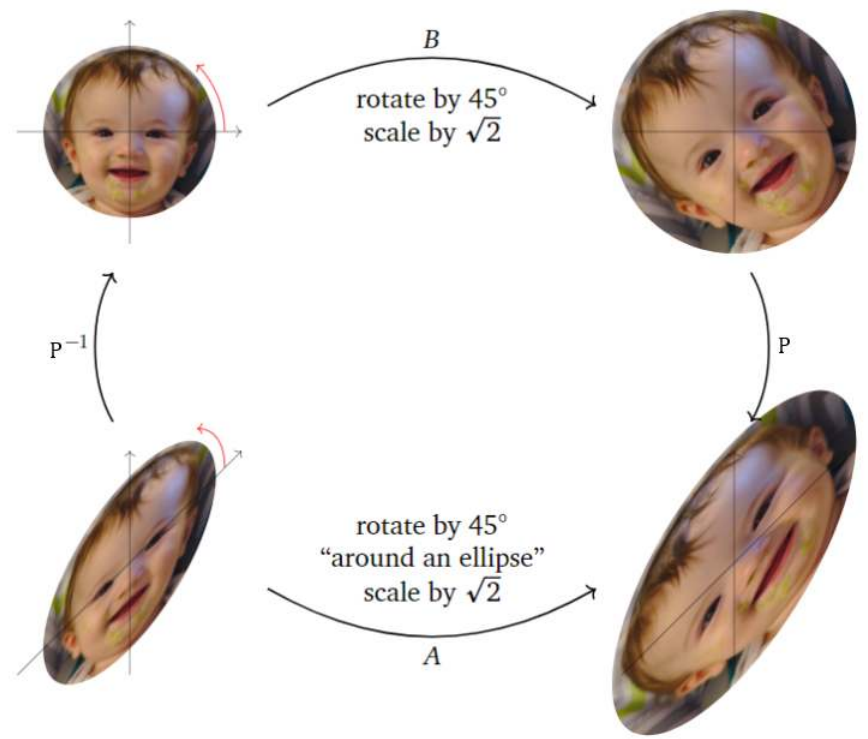


Figure 3.4.9

To summarize:

- B rotates around the circle centered at the origin and passing through $\vec{e}_1$ and $\vec{e}_2$, in the direction from $\vec{e}_1$ to $\vec{e}_2$, then scales by 2.
- A rotates around the ellipse centered at the origin and passing through $\vec{v}_R$ and $\vec{v}_I$, in the direction from $\vec{v}_R$ to $\vec{v}_I$, then scales by 2.

Note 3.4.5: Dynamics of a 2×2 Matrix with a Complex Eigenvalue

Let A be a 2×2 matrix with a complex (non-real) eigenvalue $\lambda = a - bi$ (with $b > 0$). The matrix A is similar to a matrix that rotates counterclockwise by some amount and scales by $|\lambda| = \sqrt{a^2 + b^2}$. Hence, A rotates around an ellipse and scales by $|\lambda|$. There are three different cases.

$|\lambda| = 1$: when the scaling factor is equal to 1, then vectors do not tend to get longer or shorter. In this case, repeatedly multiplying a vector by A simply "rotates around an ellipse". For example,

$$A = \frac{1}{2}\begin{bmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{bmatrix} \qquad \lambda = \frac{\sqrt{3}-i}{2} \qquad |\lambda| = 1$$

gives rise to the following picture:

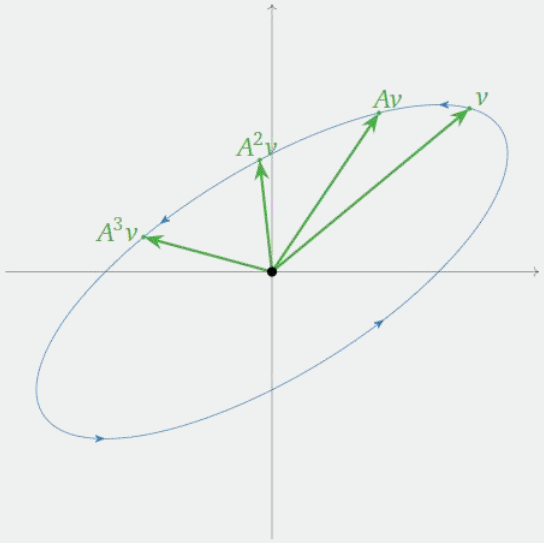


Figure 3.4.10

$|\lambda| < 1$: when the scaling factor is less than 1, then vectors tend to get shorter, i.e., closer to the origin. In this case, repeatedly multiplying a vector by A makes the vector "spiral in". For example,

$$A = \frac{1}{2\sqrt{2}}\begin{bmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{bmatrix} \qquad \lambda = \frac{\sqrt{3}-i}{2\sqrt{2}} \qquad |\lambda| = \frac{1}{\sqrt{2}} < 1$$

gives rise to the following picture:

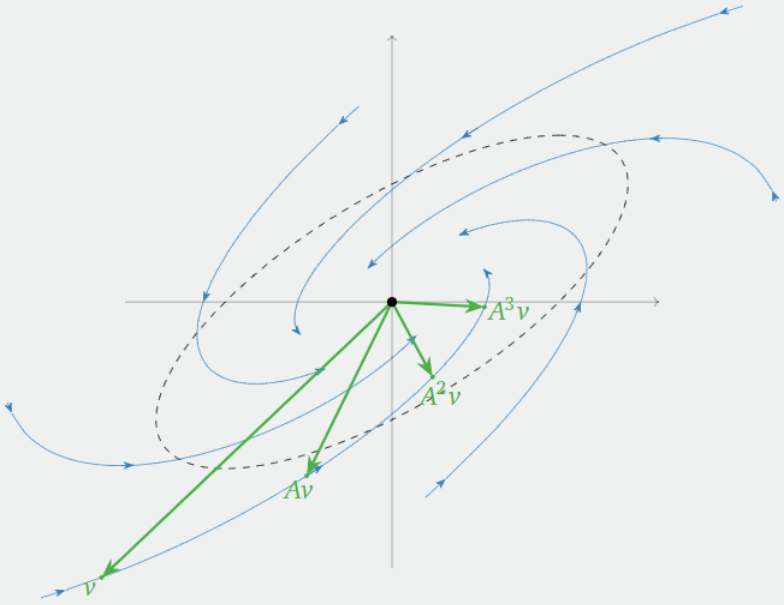


Figure 3.4.11

$|\lambda| > 1$: when the scaling factor is greater than 1, then vectors tend to get longer, i.e., farther from the origin. In this case, repeatedly multiplying a vector by A makes the vector "spiral out". For example,

$$A = \frac{1}{\sqrt{2}}\begin{bmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{bmatrix} \qquad \lambda = \frac{\sqrt{3}-i}{\sqrt{2}} \qquad |\lambda| = \sqrt{2} > 1$$

gives rise to a picture similar to above, but spinning outward (away from the origin) instead of inward.

Remark: Classification of 2×2 matrices up to similarity

At this point, we can write down the "simplest" possible matrix which is similar to any given 2×2 matrix A. There are four cases:

1. A has two real eigenvalues λ_1, λ_2. In this case, A is diagonalizable, so A is similar to the matrix

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

   This representation is unique up to reordering the eigenvalues.
2. A has one real eigenvalue λ of geometric multiplicity 2. In this case, A is *equal* to the dilation matrix

$$\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

3. A has one real eigenvalue λ of geometric multiplicity 1. In this case, A is not diagonalizable, and will be similar to the matrix

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

4. A has no real eigenvalues. In this case, A has a complex eigenvalue λ, and A is similar to a rotation-scaling matrix.

3.4: Geometry of Matrices with Complex Eigenvalues is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.5: Complex Eigenvalues** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.4: Practice Set

Exercise 3.4.1a

Find the eigenvalues and eigenspaces of $A = \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix}$

Answer

We find the characteristic polynomial: $p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 0-\lambda & -4 \\ 1 & 0-\lambda \end{vmatrix} = \lambda^2 + 4$

The roots of this polynomial are our eigenvalues: $\lambda = 0 \pm 2i$

We choose $\lambda = 0 - 2i$ (negative imaginary part). Then

$$E_{0-2i} = \text{Null}\,(A - (0-2i)I) = \text{Null}\left(\begin{bmatrix} 2i & -4 \\ 1 & 2i \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 2i \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -2i \\ 1 \end{bmatrix}\right\}$$

The conjugate space $E_{0+2i} = \text{Span}\left\{\begin{bmatrix} 2i \\ 1 \end{bmatrix}\right\}$

Exercise 3.4.1b

Find the eigenvalues and eigenspaces of $B = \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix}$

Answer

We find the characteristic polynomial: $p(\lambda) = \det(B - \lambda I) = \begin{vmatrix} 1-\lambda & -3 \\ 3 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 10$

The roots of this polynomial are our eigenvalues: $\lambda = 1 \pm 3i$

We choose $\lambda = 1 - 3i$ (negative imaginary part). Then

$$E_{1-3i} = \text{Null}\,(B - (1-3i)I) = \text{Null}\left(\begin{bmatrix} 3i & -3 \\ 3 & 3i \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -i \\ 1 \end{bmatrix}\right\}$$

The conjugate space $E_{1+3i} = \text{Span}\left\{\begin{bmatrix} i \\ 1 \end{bmatrix}\right\}$

Exercise 3.4.1c

Find the eigenvalues and eigenspaces of $C = \begin{bmatrix} 4 & 1 \\ -4 & 4 \end{bmatrix}$

Answer

We find the characteristic polynomial: $p(\lambda) = \det(C - \lambda I) = \begin{vmatrix} 4-\lambda & 1 \\ -4 & 4-\lambda \end{vmatrix} = \lambda^2 - 8\lambda + 20$

The roots of this polynomial are our eigenvalues: $\lambda = 4 \pm 2i$

We choose $\lambda = 4 - 2i$ (negative imaginary part). Then

$$E_{4-2i} = \text{Null}\,(C - (4-2i)I) = \text{Null}\left(\begin{bmatrix} 2i & 1 \\ -4 & 2i \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -\frac{i}{2} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} i \\ 2 \end{bmatrix}\right\}$$

The conjugate space $E_{4+2i} = \text{Span}\left\{\begin{bmatrix} -i \\ 2 \end{bmatrix}\right\}$

Exercise 3.4.1d

Find the eigenvalues and eigenspaces of $D = \begin{bmatrix} 0 & 2 \\ -2 & -2 \end{bmatrix}$

Answer

We find the characteristic polynomial: $p(\lambda) = \det(D - \lambda I) = \begin{vmatrix} 0-\lambda & 2 \\ -2 & -2-\lambda \end{vmatrix} = \lambda^2 + 2\lambda + 4$

The roots of this polynomial are our eigenvalues: $\lambda = -1 \pm i\sqrt{3}$

We choose $\lambda = -1 - i\sqrt{3}$ (negative imaginary part). Then

$$E_{-1-i\sqrt{3}} = \text{Null}\left(D - (-1-i\sqrt{3})I\right) = \text{Null}\left(\begin{bmatrix} 1+i\sqrt{3} & 2 \\ -2 & -1+i\sqrt{3} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & \frac{1-i\sqrt{3}}{2} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -1+i\sqrt{3} \\ 2 \end{bmatrix}\right\}$$

The conjugate space $E_{-1+i\sqrt{3}} = \text{Span}\left\{\begin{bmatrix} -1-i\sqrt{3} \\ 2 \end{bmatrix}\right\}$

Exercise 3.4.1e

Find the eigenvalues and eigenspaces of $E = \begin{bmatrix} 4 & 0 & -4 \\ 0 & 3 & 0 \\ 2 & 0 & 2 \end{bmatrix}$

Answer

We find the characteristic polynomial: $p(\lambda) = \det(E - \lambda I) = \begin{vmatrix} 4-\lambda & 0 & -4 \\ 0 & 3-\lambda & 0 \\ 2 & 0 & 2-\lambda \end{vmatrix} = -\lambda^3 + 9\lambda^2 - 34\lambda + 48$

The roots of this polynomial are our eigenvalues: $\lambda = 3$ and $\lambda = 3 \pm i\sqrt{7}$

We choose $\lambda = 3 - i\sqrt{7}$ (negative imaginary part). Then

$$E_{3-i\sqrt{7}} = \text{Null}\left(E - (3-i\sqrt{7})I\right) = \text{Null}\left(\begin{bmatrix} 1+i\sqrt{7} & 0 & -4 \\ 0 & i\sqrt{7} & 0 \\ 2 & 0 & -1+i\sqrt{7} \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & \frac{-1+i\sqrt{7}}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right)$$

$$= \text{Span}\left\{\begin{bmatrix} 1-i\sqrt{7} \\ 0 \\ 2 \end{bmatrix}\right\}$$

The conjugate space $E_{3+i\sqrt{7}} = \text{Span}\left\{\begin{bmatrix} 1+i\sqrt{7} \\ 0 \\ 2 \end{bmatrix}\right\}$

$$E_3 = \text{Null}\,(E - 3I) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -4 \\ 0 & 0 & 0 \\ 2 & 0 & -1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right\}$$

Exercise 3.4.1f

Find the eigenvalues and eigenspaces of $F = \begin{bmatrix} 1 & 1 & -2 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{bmatrix}$

Answer

We find the characteristic polynomial: $p(\lambda) = \det(F - \lambda I) = \begin{vmatrix} 1-\lambda & 1 & -2 \\ 1 & 2-\lambda & -1 \\ 1 & 0 & 1-\lambda \end{vmatrix} = -\lambda^3 + 4\lambda^2 - 6\lambda + 4$

The roots of this polynomial are our eigenvalues: $\lambda = 2$ and $\lambda = 1 \pm i$

We choose $\lambda = 1 - i$ (negative imaginary part). Then

$$E_{1-i} = \text{Null}\,(F - (1-i)I) = \text{Null}\left(\begin{bmatrix} i & 1 & -2 \\ 1 & 1+i & -1 \\ 1 & 0 & i \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & i \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -i \\ 1 \\ 1 \end{bmatrix}\right\}$$

The conjugate space $E_{1+i} = \text{Span}\left\{\begin{bmatrix} i \\ 1 \\ 1 \end{bmatrix}\right\}$

$$E_2 = \text{Null}\,(F - 2I) = \text{Null}\left(\begin{bmatrix} -1 & 1 & -2 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}\right\}$$

Exercise 3.4.2a

Describe the geometry of the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x - y \\ x + y \end{bmatrix}$

Answer

$T(\vec{x}) = A\vec{x}$ for $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

A has characteristic polynomial $p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 2$

The roots of this polynomial are our eigenvalues: $\lambda = 1 \pm i$

We choose $\lambda = 1 - i$ (negative imaginary part). Then

$$E_{1-i} = \text{Null}\,(A - (1-i)I) = \text{Null}\left(\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} i \\ 1 \end{bmatrix}\right\}$$

Therefore $A = PBP^{-1}$ for $P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = \sqrt{2}\begin{bmatrix} \cos\left(\frac{\pi}{4}\right) & -\sin\left(\frac{\pi}{4}\right) \\ \sin\left(\frac{\pi}{4}\right) & \cos\left(\frac{\pi}{4}\right) \end{bmatrix}$

So T **stretches by** $\sqrt{2}$ and **rotates counterclockwise by** $\theta = \frac{\pi}{4}$ **radians** around the ellipse with positive x-axis in the direction of $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and positive y-axis in the direction of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (which is actually a circle since the axes are perpendicular and equal length).

The conjugate space $E_{1+i} = \text{Span}\left\{\begin{bmatrix} -i \\ 1 \end{bmatrix}\right\}$ represents the same action (clockwise rotation).

Exercise 3.4.2b

Describe the geometry of the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y \\ x \end{bmatrix}$

Answer

$T(\vec{x}) = A\vec{x}$ for $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

A has characteristic polynomial $p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 0-\lambda & -1 \\ 1 & 0-\lambda \end{vmatrix} = \lambda^2 + 1$

The roots of this polynomial are our eigenvalues: $\lambda = \pm i$

We choose $\lambda = 0 - i$ (negative imaginary part). Then

$$E_{0-i} = \text{Null}\,(A - (0-i)I) = \text{Null}\left(\begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -i \\ 1 \end{bmatrix}\right\}$$

Therefore $A = PBP^{-1}$ for $P = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = 1 \cdot \begin{bmatrix} \cos\left(\frac{\pi}{2}\right) & -\sin\left(\frac{\pi}{2}\right) \\ \sin\left(\frac{\pi}{2}\right) & \cos\left(\frac{\pi}{2}\right) \end{bmatrix}$

So T **rotates counterclockwise by** $\theta = \frac{\pi}{2}$ **radians** (no scaling) around the ellipse with positive x-axis in the direction of $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and positive y-axis in the direction of $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$ (which is a circle).

The conjugate space $E_i = \text{Span}\left\{\begin{bmatrix} i \\ 1 \end{bmatrix}\right\}$ represents the same action (clockwise rotation).

Exercise 3.4.2c

Describe the geometry of the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -2x - 3y \\ 15x + 10y \end{bmatrix}$

Answer

$T(\vec{x}) = A\vec{x}$ for $A = \begin{bmatrix} -2 & -3 \\ 15 & 10 \end{bmatrix}$

A has characteristic polynomial $p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} -2-\lambda & -3 \\ 15 & 10-\lambda \end{vmatrix} = \lambda^2 - 8\lambda + 25$

The roots of this polynomial are our eigenvalues: $\lambda = 4 \pm 3i$

We choose $\lambda = 4 - 3i$ (negative imaginary part). Then

$E_{4-3i} = \text{Null}\left(A - (4-3i)I\right) = \text{Null}\left(\begin{bmatrix} -6+3i & -3 \\ 15 & 6+3i \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & \frac{2+i}{5} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -2-i \\ 5 \end{bmatrix}\right\}$

Therefore $A = PBP^{-1}$ for $P = \begin{bmatrix} -2 & -1 \\ 5 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 \\ -3 & 4 \end{bmatrix} = 5\begin{bmatrix} \cos\left(\tan^{-1}\left(\frac{3}{4}\right)\right) & -\sin\left(\tan^{-1}\left(\frac{3}{4}\right)\right) \\ \sin\left(\tan^{-1}\left(\frac{3}{4}\right)\right) & \cos\left(\tan^{-1}\left(\frac{3}{4}\right)\right) \end{bmatrix}$

So T **stretches by** 5 and **rotates counterclockwise by** $\theta = \tan^{-1}\left(\frac{3}{4}\right)$ **radians** around the ellipse with positive x-axis in the direction of $\begin{bmatrix} -2 \\ 5 \end{bmatrix}$ and positive y-axis in the direction of $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$.

The conjugate space $E_{4+3i} = \text{Span}\left\{\begin{bmatrix} -2+i \\ 5 \end{bmatrix}\right\}$ represents the same action (clockwise rotation).

Exercise 3.4.2d

Describe the geometry of the transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ with $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y \\ 4x \end{bmatrix}$

Answer

$T(\vec{x}) = A\vec{x}$ for $A = \begin{bmatrix} 0 & -1 \\ 4 & 0 \end{bmatrix}$

A has characteristic polynomial $p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 0-\lambda & -1 \\ 4 & 0-\lambda \end{vmatrix} = \lambda^2 + 4$

The roots of this polynomial are our eigenvalues: $\lambda = \pm 2i$

We choose $\lambda = 0 - 2i$ (negative imaginary part). Then

$E_{0-2i} = \text{Null}\left(A - (0-2i)I\right) = \text{Null}\left(\begin{bmatrix} 2i & -1 \\ 4 & 2i \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & \frac{i}{2} \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -i \\ 2 \end{bmatrix}\right\}$

Therefore $A = PBP^{-1}$ for $P = \begin{bmatrix} 0 & -1 \\ 2 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix} = 2\begin{bmatrix} \cos\left(\frac{\pi}{2}\right) & -\sin\left(\frac{\pi}{2}\right) \\ \sin\left(\frac{\pi}{2}\right) & \cos\left(\frac{\pi}{2}\right) \end{bmatrix}$

So T **stretches by** 2 and **rotates counterclockwise by** $\theta = \frac{\pi}{2}$ **radians** around the ellipse with positive x-axis in the direction of $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ and positive y-axis in the direction of $\begin{bmatrix} -1 \\ 0 \end{bmatrix}$.

The conjugate space $E_{2i} = \text{Span}\left\{\begin{bmatrix} i \\ 2 \end{bmatrix}\right\}$ represents the same action (clockwise rotation).

Exercise 3.4.2e

Describe the geometry of the transformation $T:\mathbb{R}^2\to\mathbb{R}^2$ with $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)=\begin{bmatrix} (1+\sqrt{3})x+y\sqrt{3} \\ -2x\sqrt{3}+(1-\sqrt{3})y \end{bmatrix}$

Answer

$T(\vec{x})=A\vec{x}$ for $A=\begin{bmatrix} 1+\sqrt{3} & \sqrt{3} \\ -2\sqrt{3} & 1-\sqrt{3} \end{bmatrix}$

A has characteristic polynomial $p(\lambda)=\det(A-\lambda I)=\begin{vmatrix} 1+\sqrt{3}-\lambda & \sqrt{3} \\ -2\sqrt{3} & 1-\sqrt{3}-\lambda \end{vmatrix}=\lambda^2-2\lambda+4$

The roots of this polynomial are our eigenvalues: $\lambda=1\pm i\sqrt{3}$

We choose $\lambda=1-i\sqrt{3}$ (negative imaginary part). Then

$$E_{1-i\sqrt{3}}=\text{Null}\left(A-(1-i\sqrt{3})I\right)=\text{Null}\left(\begin{bmatrix} \sqrt{3}+i\sqrt{3} & \sqrt{3} \\ -2\sqrt{3} & -\sqrt{3}+i\sqrt{3} \end{bmatrix}\right)=\text{Null}\left(\begin{bmatrix} 1 & \frac{1-i}{2} \\ 0 & 0 \end{bmatrix}\right)=\text{Span}\left\{\begin{bmatrix} -1+i \\ 2 \end{bmatrix}\right\}$$

Therefore $A=PBP^{-1}$ for $P=\begin{bmatrix} -1 & 1 \\ 2 & 0 \end{bmatrix}$ and $B=\begin{bmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{bmatrix}=2\begin{bmatrix} \cos\left(\frac{\pi}{3}\right) & -\sin\left(\frac{\pi}{3}\right) \\ \sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right) \end{bmatrix}$

So T **stretches by** 2 and **rotates counterclockwise by** $\theta=\frac{\pi}{3}$ **radians** around the ellipse with positive x-axis in the direction of $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and positive y-axis in the direction of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

The conjugate space $E_{1+i\sqrt{3}}=\text{Span}\left\{\begin{bmatrix} -1-i \\ 2 \end{bmatrix}\right\}$ represents the same action (clockwise rotation).

Exercise 3.4.3a

Find a linear transformation $T:\mathbb{R}^2\to\mathbb{R}^2$ which rotates each vector $\vec{x}$ by an angle of $\frac{\pi}{6}$ radians counterclockwise in a circle.

Answer

This rotation matrix should be $A=\begin{bmatrix} \cos\left(\frac{\pi}{6}\right) & -\sin\left(\frac{\pi}{6}\right) \\ \sin\left(\frac{\pi}{6}\right) & \cos\left(\frac{\pi}{6}\right) \end{bmatrix}=\begin{bmatrix} \frac{\sqrt{3}}{2} & \frac{-1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$

Exercise 3.4.3b

Find a linear transformation $T:\mathbb{R}^2\to\mathbb{R}^2$ which rotates each vector $\vec{x}$ by an angle of $\frac{\pi}{12}$ radians counterclockwise around a circle and stretches its length by a factor of 4

Answer

This rotation-scaling matrix should be $A=4\begin{bmatrix} \cos\left(\frac{\pi}{12}\right) & -\sin\left(\frac{\pi}{12}\right) \\ \sin\left(\frac{\pi}{12}\right) & \cos\left(\frac{\pi}{12}\right) \end{bmatrix}=\begin{bmatrix} \sqrt{6}+\sqrt{2} & -\sqrt{6}+\sqrt{2} \\ \sqrt{6}-\sqrt{2} & \sqrt{6}+\sqrt{2} \end{bmatrix}$

Exercise 3.4.3c

Find a linear transformation $T:\mathbb{R}^2\to\mathbb{R}^2$ which rotates each vector $\vec{x}$ by an angle of $\frac{\pi}{2}$ radians counterclockwise around an ellipse of the form $\frac{x^2}{4}+\frac{y^2}{9}=c$

Answer

The rotation-scaling matrix for a counterclockwise rotation by $\frac{\pi}{2}$ (no scaling) is $B=\begin{bmatrix} \cos\left(\frac{\pi}{2}\right) & -\sin\left(\frac{\pi}{2}\right) \\ \sin\left(\frac{\pi}{2}\right) & \cos\left(\frac{\pi}{2}\right) \end{bmatrix}=\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

The positive x-axis of such an ellipse is $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and the positive y-axis is $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$

This means the change of basis matrix is $P = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

Therefore we must have $A = PBP^{-1} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{2}{3} \\ \frac{3}{2} & 0 \end{bmatrix}$

This transformation rotates vectors counterclockwise by $\frac{\pi}{2}$ around the ellipse $\frac{x^2}{4} + \frac{y^2}{9} = c$.

3.4: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.4: Dot Products and Orthogonal Complements

Learning Objectives

1. Understand the relationship between the dot product, length, and distance.
2. Understand the relationship between the dot product and orthogonality.
3. Understand the basic properties of orthogonal complements.
4. Learn to compute the orthogonal complement of a subspace.
5. *Recipe:* shortcuts for computing the orthogonal complements of common subspaces.
6. *Picture:* orthogonal complements in $\mathbb{R}^2$ and $\mathbb{R}^3$.
7. *Vocabulary words:* ***dot product, length, distance, unit vector, unit vector in the direction of $\vec{x}$, row space***.
8. *Essential vocabulary words:* ***orthogonal, orthogonal complement***.

In the next subchapter, it will be necessary to find the *closest* point on a subspace to a given point, like so:

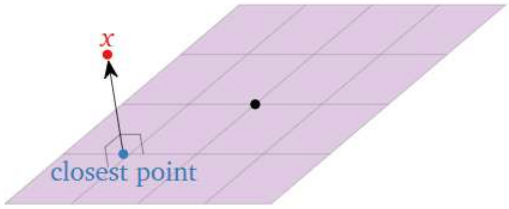


Figure 3.4.1

The closest point has the property that the difference between the two points is *orthogonal*, or *perpendicular*, to the subspace. For this reason, we need to develop notions of orthogonality, length, and distance.

This page titled 3.4: Dot Products and Orthogonal Complements is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **6.1: Dot Products and Orthogonality** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **6.2: Orthogonal Complements** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.5: Dot Products and Orthogonality

The basic construction in this section is the *dot product*, which measures angles between vectors and computes the length of a vector.

Definition: Dot Product

The ***dot product*** of two vectors $\vec{x}, \vec{y}$ in $\mathbb{R}^n$ is

$$\vec{x} \cdot \vec{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

Thinking of $\vec{x}, \vec{y}$ as matrices, this is the same as a row "multiplied" by a column in the matrix product x^Ty.

For example,

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32$$

Notice that the dot product of two *vectors* is a *scalar*.

You can do arithmetic with dot products mostly as usual, as long as you remember you can only take the dot product of two vectors together, and that the result is a scalar.

Note 3.5.1: Properties of the Dot Product

Let $\vec{x}, \vec{y}, \vec{z}$ be vectors in $\mathbb{R}^n$ and let c be a scalar.

1. *Commutativity:* $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$
2. *Distributivity with addition:* $(\vec{x} + \vec{y}) \cdot \vec{z} = \vec{x} \cdot \vec{z} + \vec{y} \cdot \vec{z}$
3. *Distributivity with scalar multiplication:* $(c\vec{x}) \cdot \vec{y} = c(\vec{x} \cdot \vec{y})$

The dot product of a vector with itself is an important special case:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1^2 + x_2^2 + \cdots + x_n^2$$

Therefore, for any vector $\vec{x}$, we have:

- $\vec{x} \cdot \vec{x} \geq 0$
- $\vec{x} \cdot \vec{x} = 0 \iff \vec{x} = \vec{0}$

This leads to a good definition of *length*.

Fact 3.5.1

The ***length*** of a vector $\vec{x}$ in $\mathbb{R}^n$ is the number

$$\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

In $\mathbb{R}^2$, this means for $\vec{v} = \begin{bmatrix} a \\ b \end{bmatrix}$, we have $\|\vec{v}\| = \sqrt{a^2 + b^2}$. This is just the Pythagorean Theorem:

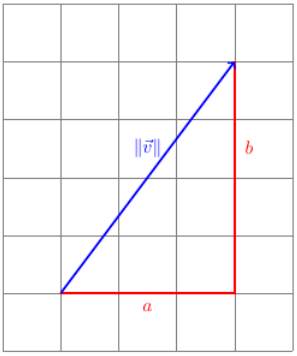


Figure 3.5.1

For vectors in $\mathbb{R}^3$, one can check that $\|\vec{x}\|$ really is the length of $\vec{x}$, although now this requires *two* applications of the Pythagorean theorem.

Note that the length of a vector is the length of the *arrow*; if we think in terms of points, then the length is its distance from the origin.

Example 3.5.1

Suppose that $\|\vec{x}\| = 2,\ \|\vec{y}\| = 3$, and $\vec{x} \cdot \vec{y} = -4$

Find $\|2\vec{x} + 3\vec{y}\|$

Solution

We compute

$$\begin{aligned} \|2\vec{x} + 3\vec{y}\|^2 &= (2\vec{x} + 3\vec{y}) \cdot (2\vec{x} + 3\vec{y}) \\ &= 4\vec{x} \cdot \vec{x} + 6\vec{x} \cdot \vec{y} + 6\vec{x} \cdot \vec{y} + 9\vec{y} \cdot \vec{y} \\ &= 4\|\vec{x}\|^2 + 9\|\vec{y}\|^2 + 12\vec{x} \cdot \vec{y} \\ &= 4 \cdot 4 + 9 \cdot 9 - 12 \cdot 4 = 49 \end{aligned}$$

Hence $\|2\vec{x} + 3\vec{y}\| = \sqrt{49} = \boxed{7}$

Fact 3.5.2

If $\vec{x}$ is a vector and c is a scalar, then $\|c\vec{x}\| = |c| \cdot \|\vec{x}\|$

This says that scaling a vector by c scales its length by $|c|$. For example,

$$\left\| \begin{bmatrix} 6 \\ 8 \end{bmatrix} \right\| = \left\| 2 \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\| = 2 \left\| \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\| = 2 \cdot \sqrt{3^2 + 4^2} = 10$$

Now that we have a good notion of length, we can define the *distance* between points in $\mathbb{R}^n$. Recall that the difference between two points $\vec{x}, \vec{y}$ is naturally a vector, namely, the vector $\vec{y} - \vec{x}$ pointing from $\vec{x}$ to $\vec{y}$.

Definition: Distance

The ***distance*** between two points $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ in $\mathbb{R}^n$ is the length of the vector from $\vec{x}$ to $\vec{y}$, and is written as a function of these vectors:

$$\text{dist}(\vec{x}, \vec{y}) = \|\vec{y} - \vec{x}\|$$

Example 3.5.2

Find the distance from $(1,2)$ to $(4,4)$

Solution

Let $\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$. Then

$$\text{dist}(\vec{x},\vec{y}) = \|\vec{y}-\vec{x}\| = \left\| \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\| = \sqrt{3^2+2^2} = \boxed{\sqrt{13}}$$

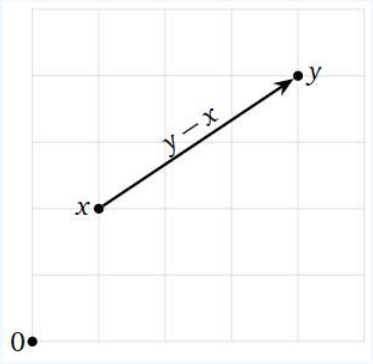


Figure 3.5.2

Vectors with length 1 are very common in applications, so we give them a name.

Definition: Unit Vector

A ***unit vector*** is a vector $\vec{x}$ with length $\|\vec{x}\| = \sqrt{\vec{x}\cdot\vec{x}} = 1$

The standard coordinate vectors, $\vec{e}_1, \vec{e}_2, \vec{e}_3, \ldots$ are unit vectors:

$$\|\vec{e}_1\| = \left\| \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\| = \sqrt{1^2+0^2+0^2} = 1$$

For any nonzero vector $\vec{x}$, there is a unique unit vector pointing in the same direction. It is obtained by dividing by the length of $\vec{x}$.

Fact 3.5.3

Let $\vec{x}$ be a nonzero vector in $\mathbb{R}^n$. The ***unit vector in the direction of*** $\vec{x}$ is the vector $\dfrac{1}{\|\vec{x}\|}\vec{x}$

This is in fact a unit vector (noting that $\|\vec{x}\|$ is a positive number), so $\left|\dfrac{1}{\|\vec{x}\|}\right| = \dfrac{1}{\|\vec{x}\|}$:

$$\left\| \frac{1}{\|\vec{x}\|}\vec{x} \right\| = \frac{1}{\|\vec{x}\|}\|\vec{x}\| = 1$$

Example 3.5.3

Find the unit vector $\vec{u}$ in the direction of $\vec{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

Solution

We divide by the length of $\vec{x}$:

$$\vec{u} = \frac{1}{\|\vec{x}\|}\vec{x} = \frac{1}{\sqrt{3^2+4^2}}\begin{bmatrix}3\\4\end{bmatrix} = \frac{1}{5}\begin{bmatrix}3\\4\end{bmatrix} = \boxed{\begin{bmatrix}\frac{3}{5}\\\frac{4}{5}\end{bmatrix}}$$

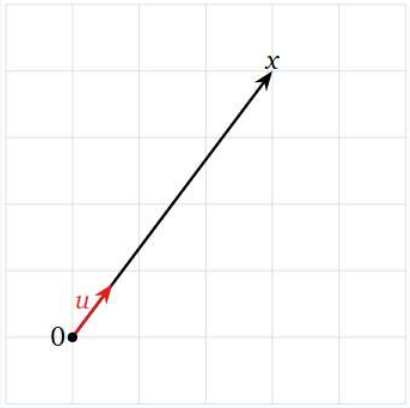


Figure 3.5.3

The dot product can be used to define *orthogonality*, i.e., when two vectors are perpendicular to each other.

Definition: Orthogonal and Perpendicular

Two vectors $\vec{x}, \vec{y}$ in $\mathbb{R}^n$ are ***orthogonal*** or ***perpendicular*** if $\vec{x} \cdot \vec{y} = 0$.

Notation: $\vec{x} \perp \vec{y}$ means $\vec{x} \cdot \vec{y} = 0$.

Note 3.5.2

Since $\vec{0} \cdot \vec{x} = \vec{0}$ for any vector $\vec{x}$, the zero vector is orthogonal to every vector in $\mathbb{R}^n$.

We motivate the above definition using the *law of cosines* in $\mathbb{R}^2$. In our language, the law of cosines asserts that if $\vec{x}, \vec{y}$ are two nonzero vectors, and if $\alpha > 0$ is the angle between them, then

$$\|\vec{y} - \vec{x}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2 - 2\|\vec{x}\|\,\|\vec{y}\|\cos(\alpha)$$

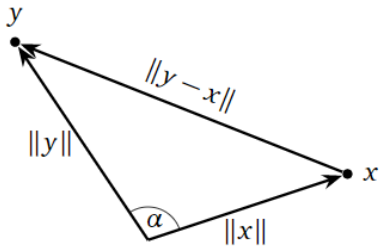


Figure 3.5.4

In particular, $\alpha = 90^\circ$ if and only if $\cos(\alpha) = 0$, which happens if and only if $\|\vec{y} - \vec{x}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$. Therefore,

$$\begin{aligned}\vec{x} \text{ and } \vec{y} \text{ are perpendicular} &\iff \|\vec{x}\|^2 + \|\vec{y}\|^2 = \|\vec{y} - \vec{x}\|^2\\ &\iff \vec{x}\cdot\vec{x} + \vec{y}\cdot\vec{y} = (\vec{y} - \vec{x})\cdot(\vec{y} - \vec{x})\\ &\iff \vec{x}\cdot\vec{x} + \vec{y}\cdot\vec{y} = \vec{y}\cdot\vec{y} + \vec{x}\cdot\vec{x} - 2\vec{x}\cdot\vec{y}\\ &\iff \vec{x}\cdot\vec{y} = 0\end{aligned}$$

To reiterate:

Note 3.5.3

$$\vec{x} \perp \vec{y} \quad \iff \quad \vec{x} \cdot \vec{y} = 0 \quad \iff \quad \|\vec{y} - \vec{x}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$$

Example 3.5.4

Find all vectors orthogonal to $\vec{v} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

Solution

We have to find all vectors $\vec{x}$ such that $\vec{x} \cdot \vec{v} = 0$. This means solving the equation

$$0 = \vec{x} \cdot \vec{v} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = x_1 + x_2 - x_3$$

The parametric form for the solution set is $x_1 = -x_2 + x_3$, so the parametric vector form of the general solution is

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Therefore, the answer is the *plane*

$$\boxed{\text{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}}$$

For instance,

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \perp \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \quad \text{because} \quad \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = 0$$

Example 3.5.5

Find all vectors orthogonal to both $\vec{v} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ and $\vec{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

Solution

We have to solve the system of two homogeneous equations

$$0 = \vec{x} \cdot \vec{v} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = x_1 + x_2 - x_3$$

$$0 = \vec{x} \cdot \vec{w} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = x_1 + x_2 + x_3$$

In matrix form:

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The parametric vector form of the solution set is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

Therefore, the answer is the *line*

$$\boxed{\text{Span}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right\}}$$

For instance,

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = 0 \quad \text{and} \quad \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 0$$

Remark: Angle between two vectors

More generally, the law of cosines gives a formula for the angle α between two nonzero vectors:

$$\begin{aligned} 2\|\vec{x}\|\,\|\vec{y}\|\cos(\alpha) &= \|\vec{x}\|^2 + \|\vec{y}\|^2 - \|\vec{y}-\vec{x}\|^2 \\ &= \vec{x}\cdot\vec{x} + \vec{y}\cdot\vec{y} - (\vec{y}-\vec{x})\cdot(\vec{y}-\vec{x}) \\ &= \vec{x}\cdot\vec{x} + \vec{y}\cdot\vec{y} - \vec{y}\cdot\vec{y} - \vec{x}\cdot\vec{x} + 2\vec{x}\cdot\vec{y} \\ &= 2\vec{x}\cdot\vec{y} \\ \implies \alpha &= \cos^{-1}\left(\frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\,\|\vec{y}\|}\right) \end{aligned}$$

In higher dimensions, we take this to be the *definition* of the angle between $\vec{x}$ and $\vec{y}$.

It will be important to compute the set of *all* vectors that are orthogonal to a given set of vectors. It turns out that a vector is orthogonal to a set of vectors if and only if it is orthogonal to the span of those vectors, which is a subspace, so we restrict ourselves to the case of subspaces.

3.5: Dot Products and Orthogonality is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.1: Dot Products and Orthogonality** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **6.2: Orthogonal Complements** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.5: Orthogonal Complements

Taking the orthogonal complement is an operation that is performed on *subspaces*.

Definition: Orthogonal Complement

Let W be a subspace of $\mathbb{R}^n$. Its ***orthogonal complement*** is the subspace

$$W^\perp = \{\vec{v} \text{ in } \mathbb{R}^n : \vec{v} \cdot \vec{w} = 0 \text{ for all } \vec{w} \text{ in } W\}$$

The symbol $W^\perp$ is sometimes read "W perp."

This is the set of all vectors $\vec{v}$ in $\mathbb{R}^n$ that are orthogonal to all of the vectors in W. We will show below that $W^\perp$ is indeed a subspace.

Note 3.5.4

We now have two similar-looking pieces of notation:

$$A^T \text{ is the transpose of a matrix } A.$$
$$W^\perp \text{ is the orthogonal complement of a subspace } W.$$

Try not to confuse the two.

The orthogonal complement of a line W through the origin in $\mathbb{R}^2$ is the perpendicular line $W^\perp$.

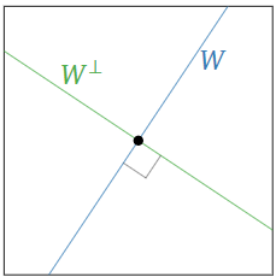


Figure 3.5.5

The orthogonal complement of a line W in $\mathbb{R}^3$ is the perpendicular plane $W^\perp$.

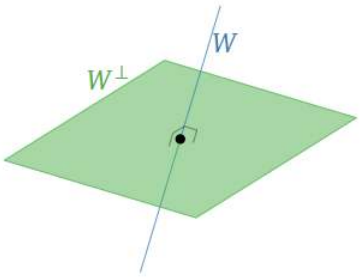


Figure 3.5.6

The orthogonal complement of a plane W in $\mathbb{R}^3$ is the perpendicular line $W^\perp$.

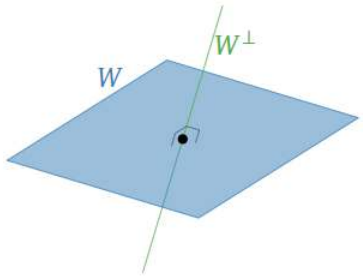


Figure 3.5.7

We see in the above pictures that $(W^\perp)^\perp = W$.

Illustration 3.5.1

The orthogonal complement of $\mathbb{R}^n$ is $\{\vec{0}\}$, since the zero vector is the only vector that is orthogonal to all of the vectors in $\mathbb{R}^n$.

For the same reason, we have $\{\vec{0}\}^\perp = \mathbb{R}^n$

Since any subspace is a span, the following proposition gives a recipe for computing the orthogonal complement of any subspace. However, below we will give several shortcuts for computing the orthogonal complements of other common kinds of subspaces–in particular, null spaces. To compute the orthogonal complement of a general subspace, usually it is best to rewrite the subspace as the column space or null space of a matrix.

Proposition 3.5.1: The Orthogonal Complement of a Column Space

Let A be a matrix and let $W = \text{Col}(A)$. Then

$$W^\perp = \text{Null}(A^T)$$

Proof

To justify the first equality, we need to show that a vector $\vec{x}$ is perpendicular to the all of the vectors in W if and only if it is perpendicular only to $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$. Since the $\vec{v}_i$ are contained in W, we really only have to show that if $\vec{x} \cdot \vec{v}_1 = \vec{x} \cdot \vec{v}_2 = \cdots = \vec{x} \cdot \vec{v}_m = 0$, then $\vec{x}$ is perpendicular to every vector $\vec{v}$ in W. Indeed, any vector in W has the form $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$ for suitable scalars $c_1, c_2, \dots, c_m$, so

$$\begin{aligned}\vec{x} \cdot \vec{v} &= \vec{x} \cdot (c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m) \\ &= c_1(\vec{x} \cdot \vec{v}_1) + c_2(\vec{x} \cdot \vec{v}_2) + \cdots + c_m(\vec{x} \cdot \vec{v}_m) \\ &= c_1(0) + c_2(0) + \cdots + c_m(0) = 0\end{aligned}$$

Therefore, $\vec{x}$ is in $W^\perp$.

To prove the second equality, we let

$$A = \begin{bmatrix} —\vec{v}_1^T— \\ —\vec{v}_2^T— \\ \vdots \\ —\vec{v}_m^T— \end{bmatrix}$$

By the row-column rule for matrix multiplication, for any vector $\vec{x}$ in $\mathbb{R}^n$ we have

$$A\vec{x} = \begin{bmatrix} \vec{v}_1^T\vec{x} \\ \vec{v}_2^T\vec{x} \\ \vdots \\ \vec{v}_m^T\vec{x} \end{bmatrix} = \begin{bmatrix} \vec{v}_1 \cdot \vec{x} \\ \vec{v}_2 \cdot \vec{x} \\ \vdots \\ \vec{v}_m \cdot \vec{x} \end{bmatrix}$$

Therefore, $\vec{x}$ is in $\text{Null}(A)$ if and only if $\vec{x}$ is perpendicular to each vector $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$.

Since column spaces are the same as spans, we can rephrase the proposition as follows. Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ be vectors in $\mathbb{R}^n$, and let $W = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$. Then

$$W^\perp = \{\text{all vectors orthogonal to each } \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\} = \text{Null}\left(\begin{bmatrix} -\vec{v}_1- \\ -\vec{v}_2- \\ \vdots \\ -\vec{v}_m- \end{bmatrix}\right)$$

Again, it is important to be able to go easily back and forth between spans and column spaces. If you are handed a span, you can apply the proposition once you have rewritten your span as a column space. By the proposition, computing the orthogonal complement of a span means *solving a system of linear equations*.

Example 3.5.6

Compute $W^\perp$, where $W = \text{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}\right\}$

Solution

Let $\vec{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$, then $\text{Span}\{\vec{v}_1, \vec{v}_2\}^\perp$ is the solution set of the homogeneous linear system associated to the matrix

$$\begin{bmatrix} -\vec{v}_1- \\ -\vec{v}_2- \end{bmatrix} = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 3 & 1 \end{bmatrix}$$

This is the solution set of the system of equations

$$\begin{cases} x_1 - x_2 + 2x_3 = 0 \\ -2x_1 + 3x_2 + x_3 = 0 \end{cases}$$

$$\begin{bmatrix} 1 & -1 & 2 \\ -2 & 3 & 1 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & 7 \\ 0 & 1 & 5 \end{bmatrix}$$

The free variable is x_3, so the parametric form of the solution set is $x_1 = -7x_3$, $x_2 = -5x_3$, and the parametric vector form is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -7 \\ -5 \\ 1 \end{bmatrix}$$

Therefore, we have $\boxed{W^\perp = \text{Span}\left\{\begin{bmatrix} -7 \\ -5 \\ 1 \end{bmatrix}\right\}}$

We can check our work:

$$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} -7 \\ -5 \\ 1 \end{bmatrix} = 0 \qquad \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -7 \\ -5 \\ 1 \end{bmatrix} = 0$$

In order to find shortcuts for computing orthogonal complements, we need the following basic facts. Looking back at the above examples, all of these facts should be believable.

Fact 3.5.4: Facts about Orthogonal Complements

Let W be a subspace of $\mathbb{R}^n$. Then:

1. $W^\perp$ is also a subspace of $\mathbb{R}^n$

2. $(W^\perp)^\perp = W$
3. $\dim(W) + \dim(W^\perp) = n$

Proof

For the first assertion, we verify the three defining properties of subspaces:

1. The zero vector is in $W^\perp$ because the zero vector is orthogonal to every vector in $\mathbb{R}^n$.
2. Let $\vec{u}, \vec{v}$ be in $W^\perp$, so $\vec{u} \cdot \vec{x} = 0$ and $\vec{v} \cdot \vec{x} = 0$ for every vector $\vec{x}$ in W. We must verify that $(\vec{u} + \vec{v}) \cdot \vec{x} = 0$ for every $\vec{x}$ in W. Indeed, we have
$$(\vec{u} + \vec{v}) \cdot \vec{x} = \vec{u} \cdot \vec{x} + \vec{v} \cdot \vec{x} = 0 + 0 = 0$$
3. Let $\vec{u}$ be in $W^\perp$, so $\vec{u} \cdot \vec{x} = 0$ for every $\vec{x}$ in W, and let c be a scalar. We must verify that $(c\vec{u}) \cdot \vec{x} = 0$ for every x in W. Indeed, we have
$$(c\vec{u}) \cdot \vec{x} = c(\vec{u} \cdot \vec{x}) = c0 = 0$$

Next we prove the third assertion. Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ be a basis for W, so $m = \dim(W)$, and let $\vec{v}_{m+1}, \vec{v}_{m+2}, \dots, \vec{v}_k$ be a basis for $W^\perp$, so $k - m = \dim(W^\perp)$. We need to show $k = n$. First we claim that $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m, \vec{v}_{m+1}, \vec{v}_{m+2}, \dots, \vec{v}_k\}$ is linearly independent. Suppose that $c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$. Let $\vec{w} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m$ and $\vec{u} = c_{m+1}\vec{v}_{m+1} + c_{m+2}\vec{v}_{m+2} + \cdots + c_k\vec{v}_k$, so $\vec{w}$ is in W, $\vec{u}$ is in $W^\perp$, and $\vec{w} + \vec{u} = \vec{0}$. Then $\vec{w} = -\vec{u}$ is in both W and $W^\perp$, which implies $\vec{w}$ is perpendicular to *itself*. In particular, $\vec{w} \cdot \vec{w} = 0$, so $\vec{w} = 0$, and hence $\vec{u} = 0$. Therefore, all coefficients c_i are equal to zero, because $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ and $\{\vec{v}_{m+1}, \vec{v}_{m+2}, \dots, \vec{v}_k\}$ are linearly independent.

It follows from the previous paragraph that $k \leq n$. Suppose that $k < n$. Then the matrix
$$A = \begin{bmatrix} —\vec{v}_1^T— \\ —\vec{v}_2^T— \\ \vdots \\ —\vec{v}_k^T— \end{bmatrix}$$
has more columns than rows (it is "wide"), so its null space is nonzero. Let $\vec{x}$ be a nonzero vector in $\text{Null}(A)$. Then
$$\vec{0} = A\vec{x} = \begin{bmatrix} \vec{v}_1 \cdot \vec{x} \\ \vec{v}_2 \cdot \vec{x} \\ \vdots \\ \vec{v}_k \cdot \vec{x} \end{bmatrix}$$
Since $\vec{v}_1 \cdot \vec{x} = \vec{v}_2 \cdot \vec{x} = \cdots = \vec{v}_m \cdot \vec{x} = 0$, it follows from that $\vec{x}$ is in $W^\perp$, and since $\vec{v}_{m+1} \cdot \vec{x} = \vec{v}_{m+2} \cdot \vec{x} = \cdots = \vec{v}_k \cdot \vec{x} = 0$, we know $\vec{x}$ is in $(W^\perp)^\perp$. As above, this implies $\vec{x}$ is orthogonal to itself, and must therefore be the zero vector. This contradicts our assumption that $\vec{x}$ is nonzero. Therefore, $k = n$, as desired.

Finally, we prove the second assertion. Clearly W is contained in $(W^\perp)^\perp$: this says that everything in W is perpendicular to the set of all vectors perpendicular to everything in W. Let $m = \dim(W)$. By 3, we have $\dim(W^\perp) = n - m$, so $\dim((W^\perp)^\perp) = n - (n - m) = m$. The only m-dimensional subspace of $(W^\perp)^\perp$ is all of $(W^\perp)^\perp$, so $(W^\perp)^\perp = W$.

For an example confirming the dimension rule, look at the Figures 5 6 and 7 above. The dimension of W plus the dimension of $W^\perp$ always adds up to n, the dimension of $\mathbb{R}^n$ they live in.

Definition: Row Space

The ***row space*** of a matrix A is the span of the rows of A, and is denoted $\text{Row}(A)$.

If A is an $m \times n$ matrix, then the rows of A are vectors with n entries, so $\text{Row}(A)$ is a subspace of $\mathbb{R}^n$. Equivalently, since the rows of A are the columns of A^T, the row space of A is the column space of A^T:
$$\text{Row}(A) = \text{Col}(A^T)$$

We showed in the above proposition, if A has rows $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$, then $\text{Row}(A)^\perp = \text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}^\perp = \text{Null}(A)$

Taking orthogonal complements of both sides gives

$$\text{Row}(A) = \text{Null}(A)^\perp$$

Replacing A by A^T and using $\text{Row}(A) = \text{Col}(A^T)$ gives

$$\text{Col}(A)^\perp = \text{Null}(A^T) \quad \text{and} \quad \text{Col}(A) = \text{Null}(A^T)^\perp$$

To summarize:

Recipe: Shortcuts for Computing Orthogonal Complements

For any vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$, we have

$$\text{Span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}^\perp = \text{Null}\left(\begin{bmatrix} —\vec{v}_1— \\ —\vec{v}_2— \\ \vdots \\ —\vec{v}_m— \end{bmatrix}\right)$$

For any matrix A, we have

$$\begin{aligned} \text{Row}(A)^\perp &= \text{Null}(A) & \qquad \text{Null}(A)^\perp &= \text{Row}(A) \\ \text{Col}(A)^\perp &= \text{Null}(A^T) & \qquad \text{Null}(A^T)^\perp &= \text{Col}(A) \end{aligned}$$

As mentioned in the beginning of this subsection, in order to compute the orthogonal complement of a general subspace, usually it is best to rewrite the subspace as the column space or null space of a matrix.

Example 3.5.7: Orthogonal complement of a subspace

Compute the orthogonal complement of the subspace

$$W = \{(x, y, z) \text{ in } \mathbb{R}^3 : 3x + 2y = z\}$$

Solution

Rewriting, we see that W is the solution set of the system of equations $3x + 2y - z = 0$, i.e., the null space of the matrix $A = \begin{bmatrix} 3 & 2 & -1 \end{bmatrix}$ Therefore,

$$\boxed{W^\perp = \text{Row}(A) = \text{Span}\left\{\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}\right\}}$$

No row reduction was needed! In fact, this shortcut is widely used in multivariable calculus when finding a vector orthogonal to a plane!

Example 3.5.8: Orthogonal complement of an eigenspace

Find the orthogonal complement of the 5-eigenspace of the matrix

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

Solution

The 5-eigenspace is

$$E_5 = \text{Null}(A - 5I) = \text{Null}\left(\begin{bmatrix} -3 & 4 & -1 \\ 3 & -3 & 0 \\ -2 & 4 & -2 \end{bmatrix}\right)$$

so

$$\boxed{E_5^{\perp} = \text{Row}\left(\begin{bmatrix} -3 & 4 & -1 \\ 3 & -3 & 0 \\ -2 & 4 & -2 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -3 \\ 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 4 \\ -2 \end{bmatrix}\right\}}$$

These vectors would have to be linearly independent to be a basis for $E_5^{\perp}$, and we know they are not! This is because for $\lambda = 5$ to be an eigenvalue of A, we know E_5 must contain a non-zero vector. This means $A - 5I$ is not invertible, and neither will be $(A - 5I)^T = A^T - 5I$. Thus the columns of $A^T - 5I$, which are the rows we are using above must be linearly dependent.

Suppose that A is an $m \times n$ matrix. Let us refer to the dimensions of $\text{Col}(A)$ and $\text{Row}(A)$ as the *row rank* and the *column rank* of A (note that the column rank of A is the same as the rank of A). The next theorem says that the row and column ranks are the same. This is surprising for a couple of reasons. First, $\text{Row}(A)$ lies in $\mathbb{R}^n$ and $\text{Col}(A)$ lies in $\mathbb{R}^m$. Also, the theorem implies that A and A^T have the same number of pivots, even though the reduced row echelon forms of A and A^T have nothing to do with each other otherwise.

Theorem 3.5.1

Let A be a matrix. Then the row rank of A is equal to the column rank of A.

Proof

We know for any $m \times n$ matrix A: $\dim \text{Col}(A) + \dim \text{Null}(A) = n$

On the other hand, we have found that $\dim \text{Null}(A)^{\perp} + \dim \text{Null}(A) = n$ also holds

which implies $\dim \text{Col}(A) = \dim \text{Null}(A)^{\perp}$

Since $\text{Null}(A)^{\perp} = \text{Row}(A)$, we have $\dim \text{Col}(A) = \dim \text{Row}(A)$ as desired.

In particular, both the row rank and the column rank are equal to the number of pivots of A, which is why we just say "rank".

3.5: Orthogonal Complements is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.2: Orthogonal Complements** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.5: Practice Set

Exercise 3.5.1

Let $\vec{x} = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 6 \\ -8 \end{bmatrix}$, and $\vec{z} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$

(a) Evaluate $\vec{x} \cdot \vec{y}$

(b) Evaluate $\vec{x} \cdot \vec{z}$

(c) Evaluate $\|\vec{y}\|$

(d) Find a unit vector in the direction of $\vec{z}$

(e) What is the angle between $\vec{x}$ and $\vec{z}$

(f) What is the angle between $\vec{x} - \vec{y}$ and $\vec{z} - \vec{x}$?

(g) What is the orthogonal complement of $W = \text{Span}\{\vec{x}\}$?

(h) What is the orthogonal complement of $W = \text{Span}\{\vec{y}, \vec{z}\}$?

Answer

(a) $\vec{x} \cdot \vec{y} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ -8 \end{bmatrix} = (-2) \cdot 6 + 0 \cdot (-8) = -12$

(b) $\vec{x} \cdot \vec{z} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 3 \end{bmatrix} = (-2) \cdot (-2) + 0 \cdot 3 = 4$

(c) $\|\vec{y}\| = \sqrt{\vec{y} \cdot \vec{y}} = \sqrt{6^2 + (-8)^2} = \sqrt{100} = 10$

(d) $\|\vec{z}\| = \sqrt{\vec{z} \cdot \vec{z}} = \sqrt{(-2)^2 + 3^2} = \sqrt{13}$

The unit vector in the direction of $\vec{z}$ is $\frac{1}{\|\vec{z}\|}\vec{z} = \frac{1}{\sqrt{13}} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{2}{\sqrt{13}} \\ \frac{3}{\sqrt{13}} \end{bmatrix}$

(e) $\vec{x} \cdot \vec{z} = 4$ from part (b) and $\|\vec{z}\| = \sqrt{13}$ from part (d),

$\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{(-2)^2 + 0^2} = \sqrt{4} = 2$

Thus the angle between these vectors is $\cos^{-1}\left(\frac{\vec{x} \cdot \vec{z}}{\|\vec{x}\|\|\vec{z}\|}\right) = \cos^{-1}\left(\frac{4}{2\sqrt{13}}\right) = \cos^{-1}\left(\frac{2}{\sqrt{13}}\right)$

(f) First we find $\vec{a} = \vec{x} - \vec{y} = \begin{bmatrix} -2 \\ 0 \end{bmatrix} - \begin{bmatrix} 6 \\ -8 \end{bmatrix} = \begin{bmatrix} -8 \\ 8 \end{bmatrix}$ and $\vec{b} = \vec{z} - \vec{x} = \begin{bmatrix} -2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$

$\|\vec{a}\| = \sqrt{\vec{a} \cdot \vec{a}} = \sqrt{(-8)^2 + 8^2} = \sqrt{128} = 8\sqrt{2}$

$\|\vec{b}\| = \sqrt{\vec{b} \cdot \vec{b}} = \sqrt{0^2 + 3^2} = \sqrt{9} = 3$

$\vec{a} \cdot \vec{b} = \begin{bmatrix} -8 \\ 8 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 3 \end{bmatrix} = (-8) \cdot 0 + 8 \cdot 3 = 24$

Thus the angle between these vectors is $\cos^{-1}\left(\frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|\|\vec{b}\|}\right) = \cos^{-1}\left(\frac{24}{8\sqrt{2} \cdot 3}\right) = \cos^{-1}\left(\frac{1}{\sqrt{2}}\right) = \frac{\pi}{4}$ radians

(g) If $W = \text{Span}\left\{\begin{bmatrix} -2 \\ 0 \end{bmatrix}\right\}$, then $W^{\perp} = \text{Null}\left(\begin{bmatrix} -2 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$

(h) If $W = \text{Span}\left\{\begin{bmatrix} 6 \\ -8 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \end{bmatrix}\right\}$, then $W^{\perp} = \text{Null}\left(\begin{bmatrix} 6 & -8 \\ -2 & 3 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \{\vec{0}\}$

Exercise 3.5.2

Let $\vec{u} = \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} 2 \\ 11 \\ 17 \end{bmatrix}$, and $\vec{w} = \begin{bmatrix} -6 \\ 12 \\ 12 \end{bmatrix}$

(a) Evaluate $\vec{v} \cdot \vec{w}$

(b) Evaluate $(\vec{u} + \vec{v}) \cdot (\vec{v} - \vec{w})$

(c) Evaluate $\|2\vec{v} - 3\vec{w}\|$

(d) Find a unit vector in the direction of $\vec{u}$

(e) What is the angle between $\vec{u}$ and $\vec{w}$

(f) What is the angle between $\vec{w}$ and $\vec{v} - \vec{w}$?

(g) What is the orthogonal complement of $W = \text{Span}\{\vec{u}\}$?

(h) What is the orthogonal complement of $W = \text{Span}\{\vec{u}, \vec{w}\}$?

Answer

(a) $\vec{v} \cdot \vec{w} = \begin{bmatrix} 2 \\ 11 \\ 17 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ 12 \\ 12 \end{bmatrix} = 2 \cdot (-6) + 11 \cdot 12 + 17 \cdot 12 = 324$

(b) $\vec{u} + \vec{v} = \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix} + \begin{bmatrix} 2 \\ 11 \\ 17 \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ 15 \end{bmatrix}$ and $\vec{v} - \vec{w} = \begin{bmatrix} 2 \\ 11 \\ 17 \end{bmatrix} - \begin{bmatrix} -6 \\ 12 \\ 12 \end{bmatrix} = \begin{bmatrix} 8 \\ -1 \\ 5 \end{bmatrix}$

So $(\vec{u} + \vec{v}) \cdot (\vec{v} - \vec{w}) = \begin{bmatrix} 3 \\ 9 \\ 15 \end{bmatrix} \cdot \begin{bmatrix} 8 \\ -1 \\ 5 \end{bmatrix} = 3 \cdot 8 + 9 \cdot (-1) + 15 \cdot 5 = 90$

(c) $2\vec{v} - 3\vec{w} = \begin{bmatrix} 4 \\ 22 \\ 34 \end{bmatrix} - \begin{bmatrix} -18 \\ 36 \\ 36 \end{bmatrix} = \begin{bmatrix} 22 \\ -14 \\ -2 \end{bmatrix}$

So $\|2\vec{v} - 3\vec{w}\| = \sqrt{(2\vec{v} - 3\vec{w}) \cdot (2\vec{v} - 3\vec{w})} = \sqrt{22^2 + (-14)^2 + (-2)^2} = \sqrt{684} = 6\sqrt{19}$

(d) $\|\vec{u}\| = \sqrt{\vec{u} \cdot \vec{u}} = \sqrt{1^2 + (-2)^2 + (-2)^2} = \sqrt{9} = 3$

The unit vector in the direction of $\vec{u}$ is $\frac{1}{\|\vec{u}\|}\vec{u} = \frac{1}{3}\begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ -\frac{2}{3} \end{bmatrix}$

(e) $\|\vec{u}\| = 3$ from part (d), and $\|\vec{w}\| = \sqrt{\vec{w} \cdot \vec{w}} = \sqrt{(-6)^2 + 12^2 + 12^2} = \sqrt{324} = 18$

$\vec{u} \cdot \vec{w} = \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix} \cdot \begin{bmatrix} -6 \\ 12 \\ 12 \end{bmatrix} = 1 \cdot (-6) + (-2) \cdot 12 + (-2) \cdot 12 = -54$

Thus the angle between these vectors is $\cos^{-1}\left(\frac{\vec{u} \cdot \vec{w}}{\|\vec{u}\|\|\vec{w}\|}\right) = \cos^{-1}\left(\frac{-54}{3 \cdot 18}\right) = \cos^{-1}(-1) = \pi$ radians

(f) First we find $\vec{a} = \vec{v} - \vec{w} = \begin{bmatrix} 2 \\ 11 \\ 17 \end{bmatrix} - \begin{bmatrix} -6 \\ 12 \\ 12 \end{bmatrix} = \begin{bmatrix} 8 \\ -1 \\ 5 \end{bmatrix}$

We know $\|\vec{w}\| = 18$ from part (e), and $\|\vec{a}\| = \sqrt{\vec{a} \cdot \vec{a}} = \sqrt{8^2 + (-1)^2 + 5^2} = \sqrt{90} = 3\sqrt{10}$

$\vec{w} \cdot \vec{a} = \begin{bmatrix} -6 \\ 12 \\ 12 \end{bmatrix} \cdot \begin{bmatrix} 8 \\ -1 \\ 5 \end{bmatrix} = (-6) \cdot 8 + 12 \cdot (-1) + 12 \cdot 5 = 0$

Thus the angle between these vectors is $\cos^{-1}\left(\frac{\vec{w}\cdot\vec{a}}{\|\vec{w}\|\|\vec{a}\|}\right)=\cos^{-1}\left(\frac{0}{18\cdot 3\sqrt{10}}\right)=\cos^{-1}(0)=\frac{\pi}{2}$ radians

(g) If $W=\text{Span}\left\{\begin{bmatrix}1\\-2\\-2\end{bmatrix}\right\}$, then $W^{\perp}=\text{Null}\left(\begin{bmatrix}1&-2&-2\end{bmatrix}\right)=\text{Span}\left\{\begin{bmatrix}2\\0\\1\end{bmatrix},\begin{bmatrix}2\\1\\0\end{bmatrix}\right\}$

(h) If $W=\text{Span}\left\{\begin{bmatrix}1\\-2\\-2\end{bmatrix},\begin{bmatrix}-6\\12\\12\end{bmatrix}\right\}$, then $W^{\perp}=\text{Null}\left(\begin{bmatrix}1&-2&-2\\-6&12&12\end{bmatrix}\right)=\text{Null}\left(\begin{bmatrix}1&-2&-2\\0&0&0\end{bmatrix}\right)$
$=\text{Span}\left\{\begin{bmatrix}2\\0\\1\end{bmatrix},\begin{bmatrix}2\\1\\0\end{bmatrix}\right\}$

Exercise 3.5.3

Let $\vec{x}=\begin{bmatrix}1\\-2\end{bmatrix}$, $\vec{y}=\begin{bmatrix}3\\1\end{bmatrix}$, and $\vec{z}=\begin{bmatrix}h\\h+1\end{bmatrix}$

(a) What value(s) of h make $\vec{x}$ orthogonal to $\vec{z}$?

(b) What value(s) of h make $\vec{y}$ orthogonal to $\vec{z}$?

(c) What value(s) of h make $\vec{z}$ a unit vector?

(d) What value(s) of h make the angle between $\vec{y}$ and $\vec{z}$ equal to $\frac{\pi}{4}$ radians?

(e) What value(s) of h make the angle between $\vec{x}$ and $\vec{z}$ equal to $\frac{\pi}{3}$ radians?

Answer

(a) $\vec{x}\cdot\vec{z}=\begin{bmatrix}1\\-2\end{bmatrix}\cdot\begin{bmatrix}h\\h+1\end{bmatrix}=1\cdot h+(-2)\cdot(h+1)=-h-2$

These are orthogonal when $\vec{x}\cdot\vec{z}=-h-2=0$ so our solution is $h=-2$

(b) $\vec{y}\cdot\vec{z}=\begin{bmatrix}3\\1\end{bmatrix}\cdot\begin{bmatrix}h\\h+1\end{bmatrix}=3\cdot h+1\cdot(h+1)=4h+1$

These are orthogonal when $\vec{y}\cdot\vec{z}=4h+1=0$ so our solution is $h=-\frac{1}{4}$

(c) $\|\vec{z}\|=\sqrt{\vec{z}\cdot\vec{z}}=\sqrt{h^2+(h+1)^2}=\sqrt{2h^2+2h+1}$

So $\vec{z}$ will be a unit vector when $\sqrt{2h^2+2h+1}=1$, which has solutions $h=0,-1$

(d) $\|\vec{y}\|=\sqrt{\vec{y}\cdot\vec{y}}=\sqrt{3^2+1^2}=\sqrt{10}$

We already have seen that $\|\vec{z}\|=\sqrt{2h^2+2h+1}$ and $\vec{y}\cdot\vec{z}=4h+1$

Therefore we want to satisfy $\cos\left(\frac{\pi}{4}\right)=\frac{4h+1}{\sqrt{10}\cdot\sqrt{2h^2+2h+1}}$

This is equivalent to $\frac{1}{\sqrt{2}}=\frac{4h+1}{\sqrt{20h^2+20h+10}}\rightarrow\sqrt{10h^2+10h+5}=4h+1$

If we square both sides this becomes $10h^2+10h+5=16h^2+8h+1\rightarrow 6h^2-2h-4=0\rightarrow h=-\frac{2}{3},1$

Only $h=1$ actually satisfies the original equation with the square roots, so the answer is $h=1$.

(e) $\|\vec{x}\|=\sqrt{\vec{x}\cdot\vec{x}}=\sqrt{1^2+(-2)^2}=\sqrt{5}$

We already have seen that $\|\vec{z}\|=\sqrt{2h^2+2h+1}$ and $\vec{x}\cdot\vec{z}=-h-2$

Therefore we want to satisfy $\cos\left(\frac{\pi}{3}\right)=\frac{-h-2}{\sqrt{5}\cdot\sqrt{2h^2+2h+1}}$

This is equivalent to $\frac{1}{2}=\frac{-h-2}{\sqrt{10h^2+10h+5}}\rightarrow\sqrt{10h^2+10h+5}=-2h-4$

If we square both sides this becomes $10h^2+10h+5=4h^2+16h+16 \rightarrow 6h^2-6h-11=0 \rightarrow h=\dfrac{3\pm 5\sqrt{3}}{6}$

Neither of these solutions actually satisfy the original equation with the square roots, so there are **no solutions**.

Exercise 3.5.4

Let $\vec{u}=\begin{bmatrix}3\\-2\\-1\end{bmatrix}$, $\vec{v}=\begin{bmatrix}h\\1\\-1\end{bmatrix}$, and $\vec{w}=\begin{bmatrix}h\\h\\h\end{bmatrix}$

(a) What value(s) of h make $\vec{u}$ orthogonal to $\vec{v}$?

(b) What value(s) of h make $\vec{u}$ orthogonal to $\vec{w}$?

(c) What value(s) of h make $\vec{v}$ orthogonal to $\vec{w}$?

(d) What value(s) of h make $\vec{w}$ a unit vector?

(e) What value(s) of h make $\vec{v}$ a unit vector?

(f) What value(s) of h make the angle between $\vec{v}$ and $\vec{w}$ equal to $\dfrac{\pi}{3}$ radians?

Answer

(a) $\vec{u}\cdot\vec{v}=\begin{bmatrix}3\\-2\\-1\end{bmatrix}\cdot\begin{bmatrix}h\\1\\-1\end{bmatrix}=3\cdot h+(-2)\cdot 1+(-1)\cdot(-1)=3h-1$

These vectors are orthogonal when $\vec{u}\cdot\vec{v}=3h-1=0$, so our solution is $h=\dfrac{1}{3}$

(b) $\vec{u}\cdot\vec{w}=\begin{bmatrix}3\\-2\\-1\end{bmatrix}\cdot\begin{bmatrix}h\\h\\h\end{bmatrix}=3\cdot h+(-2)\cdot h+(-1)\cdot h=0$

These vectors are orthogonal when $\vec{u}\cdot\vec{w}=0$, which holds **for all possible values of** h

(c) $\vec{v}\cdot\vec{w}=\begin{bmatrix}h\\1\\-1\end{bmatrix}\cdot\begin{bmatrix}h\\h\\h\end{bmatrix}=h\cdot h+1\cdot h+(-1)\cdot h=h^2$

These vectors are orthogonal when $\vec{v}\cdot\vec{w}=h^2=0$, so our solution is $h=0$.

(d) $\|\vec{w}\|=\sqrt{\vec{w}\cdot\vec{w}}=\sqrt{h^2+h^2+h^2}=\sqrt{3h^2}$

This will be a unit vector when $\|\vec{w}\|=\sqrt{3h^2}=1$, so our solutions are $h=\pm\dfrac{1}{\sqrt{3}}$

(e) $\|\vec{v}\|=\sqrt{\vec{v}\cdot\vec{v}}=\sqrt{h^2+1^2+(-1)^2}=\sqrt{h^2+2}$

This will be a unit vector when $\|\vec{v}\|=\sqrt{h^2+2}=1$, which has no real solutions.

There is no h **possible** that makes $\vec{v}$ a unit vector.

(f) We already have seen that $\|\vec{v}\|=\sqrt{h^2+2}$, $\|\vec{w}\|=\sqrt{3h^2}$, and $\vec{v}\cdot\vec{w}=h^2$

Therefore we want to satisfy $\cos\left(\dfrac{\pi}{3}\right)=\dfrac{h^2}{\sqrt{h^2+2}\cdot\sqrt{3h^2}}$

This is equivalent to $\dfrac{1}{2}=\dfrac{h^2}{\sqrt{3h^4+6h^2}}\rightarrow\sqrt{3h^4+6h^2}=2h^2$

If we square both sides this becomes $3h^4+6h^2=4h^4\rightarrow h^4-6h^2=0\rightarrow h=0,\pm\sqrt{6}$

$h=0$ does not satisfy the original equation (since a zero vector cannot make an angle anyway!) but $h=\pm\sqrt{6}$ are our solutions.

Exercise 3.5.5

Let $A = \begin{bmatrix} 3 & 2 & -1 \\ -2 & 4 & -3 \\ 7 & -6 & 5 \end{bmatrix}$

(a) Find a basis for $\text{Row}(A)$.

(b) Find a basis for $\text{Col}(A)$.

(c) Find a basis for $\text{Row}(A^T)$.

(d) Find a basis for $\text{Col}(A^T)$.

(e) Find a basis for the orthogonal complement of $\text{Row}(A)$.

(f) Find a basis for the orthogonal complement of $\text{Col}(A)$.

Answer

$$A = \begin{bmatrix} 3 & 2 & -1 \\ -2 & 4 & -3 \\ 7 & -6 & 5 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & \frac{1}{8} \\ 0 & 1 & -\frac{11}{16} \\ 0 & 0 & 0 \end{bmatrix}$$

(a) $\text{Row}(A) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ \frac{1}{8} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -\frac{11}{16} \end{bmatrix}\right\}$ so one such basis is $\left\{\begin{bmatrix} 8 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 16 \\ -11 \end{bmatrix}\right\}$

(b) $\text{Col}(A) = \text{Span}\left\{\begin{bmatrix} 3 \\ -2 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ -6 \end{bmatrix}\right\}$ so one such basis is $\left\{\begin{bmatrix} 3 \\ -2 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ -6 \end{bmatrix}\right\}$

$$A^T = \begin{bmatrix} 3 & -2 & 7 \\ 2 & 4 & -6 \\ -1 & -3 & 5 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$$

(c) $\text{Row}(A^T) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}\right\}$ so one such basis is $\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}\right\}$

(another possible answer would be your answer from part (b) since $\text{Row}(A^T) = \text{Col}(A)$)

(d) $\text{Col}(A^T) = \text{Span}\left\{\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 4 \\ -3 \end{bmatrix}\right\}$ so one such basis is $\left\{\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 4 \\ -3 \end{bmatrix}\right\}$

(another possible answer would be your answer from part (a) since $\text{Col}(A^T) = \text{Row}(A)$)

(e) $(\text{Row}(A))^\perp = \text{Null}(A) = \text{Span}\left\{\begin{bmatrix} -2 \\ 11 \\ 16 \end{bmatrix}\right\}$ so $\left\{\begin{bmatrix} -2 \\ 11 \\ 16 \end{bmatrix}\right\}$ is one such basis.

(f) $(\text{Col}(A))^\perp = (\text{Row}(A^T))^\perp = \text{Null}(A^T) = \text{Span}\left\{\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}\right\}$ so $\left\{\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}\right\}$ is one such basis.

Exercise 3.5.6

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If $\vec{v} \in \mathbb{R}^n$ then $\vec{v}$ is orthogonal to $\vec{0}$.

(b) If $\vec{v}$ is orthogonal to all vectors $\vec{w} \in \mathbb{R}^n$, then $\vec{v} = \vec{0}$.

(c) If $\vec{v}, \vec{w}$ are non-zero orthogonal vectors, then they are linearly independent.

(d) If W is a subspace of $\mathbb{R}^n$ with dimension m, then there is an $n-m$ dimensional subspace containing the vectors which are orthogonal to every vector in W.

(e) If W_2 is a subspace of W_1, then $W_2^\perp$ is a subspace of $W_1^\perp$

(f) If W_1 and W_2 are subspaces with the same orthogonal complement, then $W_1 = W_2$

(g) If $\vec{u}$ and $\vec{v}$ are unit vectors satisfying $\vec{u} \cdot \vec{v} = -1$, then $\{\vec{u}, \vec{v}\}$ is a linearly independent set.

Answer

(a) **Always True**

$\vec{0} \cdot \vec{v} = 0$ always holds for any vector $\vec{v}$.

(b) **Always True**

$\vec{v} \cdot \vec{v} = 0$ means that $\vec{v} = \vec{0}$.

(c) **Always True**

If $a\vec{v} + b\vec{w} = \vec{0}$, then $\vec{v} \cdot (a\vec{v} + b\vec{w}) = \vec{v} \cdot \vec{0}$, which simplifies to $a\|\vec{v}\| = 0$, and since $\vec{v} \neq \vec{0}$, we must have $a = 0$, now this becomes $b\vec{w} = \vec{0}$, and since $\vec{w} \neq \vec{0}$ we must have $b = 0$ and these vectors are linearly independent.

(d) **Always True**

If $W = \text{Row}(A)$ has dimension m, then $W^\perp = \text{Null}(A)$ must have dimension $n - m$ by the rank-nullity equation.

(e) **Sometimes True Sometimes False**

This is true easily $W_1 = W_2$, but in any other case, $W_1^\perp$ is a subspace of $W_2^\perp$ instead.

(f) **Always True**

$W_1 = (W_1^\perp)^\perp$ and $W_2 = (W_2^\perp)^\perp$ so the same orthogonal complement forces the subspaces to be the same.

(g) **Always False**

The angle between these vectors satisfies $\cos(\theta) = \dfrac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|} = \dfrac{-1}{1 \cdot 1} = -1$, so it must be $\theta = \pi$ radians, and thus $\vec{u} = -\vec{v}$ must hold.

3.5: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.5: Orthogonal Projection and Least Squares

Learning Objectives

1. Understand the orthogonal decomposition of a vector with respect to a subspace.
2. Understand the relationship between orthogonal decomposition and orthogonal projection.
3. Understand the relationship between orthogonal decomposition and the closest vector on / distance to a subspace.
4. Learn the basic properties of orthogonal projections as linear transformations and as matrix transformations.
5. Examples of best fit problems.
6. *Recipes:* orthogonal projection onto a line, orthogonal decomposition by solving a system of equations, orthogonal projection via a complicated matrix product.
7. *Pictures:* orthogonal decomposition, orthogonal projection, geometry of a least-squares solution.
8. *Vocabulary words:* ***orthogonal decomposition***, ***orthogonal projection, least-squares solution***.

Let W be a subspace of $\mathbb{R}^n$ and let $\vec{x}$ be a vector in $\mathbb{R}^n$. In this section, we will learn to compute the *closest vector* $\vec{x}_W$ to $\vec{x}$ in W. The vector $\vec{x}_W$ is called the *orthogonal projection* of $\vec{x}$ onto W. This is exactly what we will use to almost solve matrix equations, and solve using the method of least squares.

This page titled 3.5: Orthogonal Projection and Least Squares is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **6.3: Orthogonal Projection** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.6: Orthogonal Decomposition

We begin by fixing some notation.

Definition: $\vec{x}_W$

Let W be a subspace of $\mathbb{R}^n$ and let $\vec{x}$ be a vector in $\mathbb{R}^n$. We denote the closest vector to $\vec{x}$ on W by $\vec{x}_W$.

To say that $\vec{x}_W$ is the closest vector to $\vec{x}$ on W means that the difference $\vec{x} - \vec{x}_W$ is *orthogonal* to the vectors in W:

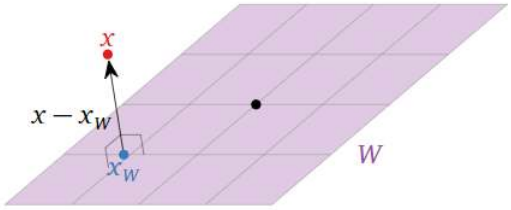


Figure 3.6.1

In other words, if $\vec{x}_{W^\perp} = \vec{x} - \vec{x}_W$, then we have $\vec{x} = \vec{x}_W + \vec{x}_{W^\perp}$, where $\vec{x}_W$ is in W and $\vec{x}_{W^\perp}$ is in $W^\perp$. The first order of business is to prove that the closest vector always exists.

Theorem 3.6.1: Orthogonal Decomposition

Let W be a subspace of $\mathbb{R}^n$ and let $\vec{x}$ be a vector in $\mathbb{R}^n$. Then we can write $\vec{x}$ uniquely as

$$\vec{x} = \vec{x}_W + \vec{x}_{W^\perp}$$

where $\vec{x}_W$ is the closest vector to $\vec{x}$ on W and $\vec{x}_{W^\perp}$ is in $W^\perp$.

Proof

Let $m = \dim(W)$, so $n - m = \dim(W^\perp)$. Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ be a basis for W and let $\vec{v}_{m+1}, \vec{v}_{m+2}, \dots, \vec{v}_n$ be a basis for $W^\perp$. We know $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m, \vec{v}_{m+1}, \vec{v}_{m+2}, \dots, \vec{v}_n\}$ is linearly independent, so it forms a basis for $\mathbb{R}^n$. Therefore, we can write

$$\vec{x} = (c_1\vec{v}_1 + \cdots + c_m\vec{v}_m) + (c_{m+1}\vec{v}_{m+1} + \cdots + c_n\vec{v}_n) = \vec{x}_W + \vec{x}_{W^\perp},$$

where $\vec{x}_W = c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$ and $\vec{x}_{W^\perp} = c_{m+1}\vec{v}_{m+1} + \cdots + c_n\vec{v}_n$. Since $\vec{x}_{W^\perp}$ is orthogonal to W, the vector $\vec{x}_W$ is the closest vector to $\vec{x}$ on W, so this proves that such a decomposition exists.

As for uniqueness, suppose that

$$\vec{x} = \vec{x}_W + \vec{x}_{W^\perp} = \vec{y}_W + \vec{y}_{W^\perp}$$

for $\vec{x}_W, \vec{y}_W$ in W and $\vec{x}_{W^\perp}, \vec{y}_{W^\perp}$ in $W^\perp$. Rearranging gives

$$\vec{x}_W - \vec{y}_W = \vec{y}_{W^\perp} - \vec{x}_{W^\perp}$$

Since W and $W^\perp$ are subspaces, the left side of the equation is in W and the right side is in $W^\perp$. Therefore, $\vec{x}_W - \vec{y}_W$ is in W and in $W^\perp$, so it is orthogonal to itself, which implies $\vec{x}_W - \vec{y}_W = \vec{0}$. Hence $\vec{x}_W = \vec{y}_W$ and $\vec{x}_{W^\perp} = \vec{y}_{W^\perp}$, which proves uniqueness.

Definition: Orthogonal Decomposition and Orthogonal Projection

Let W be a subspace of $\mathbb{R}^n$ and let $\vec{x}$ be a vector in $\mathbb{R}^n$. The expression

$$\vec{x} = \vec{x}_W + \vec{x}_{W^\perp}$$

for $\vec{x}_W$ in W and $\vec{x}_{W^\perp}$ in $W^\perp$, is called the ***orthogonal decomposition*** of $\vec{x}$ with respect to W, and the closest vector $\vec{x}_W$ is the ***orthogonal projection*** of $\vec{x}$ onto W.

Since $\vec{x}_W$ is the closest vector on W to $\vec{x}$, the distance from $\vec{x}$ to the subspace W is the length of the vector from $\vec{x}_W$ to $\vec{x}$, i.e., the length of $\vec{x}_{W^\perp}$. To restate:

Note 3.6.1: Closest Vector and Distance

Let W be a subspace of $\mathbb{R}^n$ and let $\vec{x}$ be a vector in $\mathbb{R}^n$.

- The orthogonal projection $\vec{x}_W$ is the closest vector to $\vec{x}$ in W
- The distance from $\vec{x}$ to W is $\|\vec{x}_{W^\perp}\|$

Illustration 3.6.1: Orthogonal decomposition with respect to the xy-plane

Let W be the xy-plane in $\mathbb{R}^3$, so $W^\perp$ is the z-axis. It is easy to compute the orthogonal decomposition of a vector with respect to this W:

$$\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \implies \vec{u}_W = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \quad \vec{u}_{W^\perp} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$
$$\vec{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \implies \vec{v}_W = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \quad \vec{v}_{W^\perp} = \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix}$$

We see that the orthogonal decomposition in this case expresses a vector in terms of a "horizontal" component (in the xy-plane) and a "vertical" component (on the z-axis).

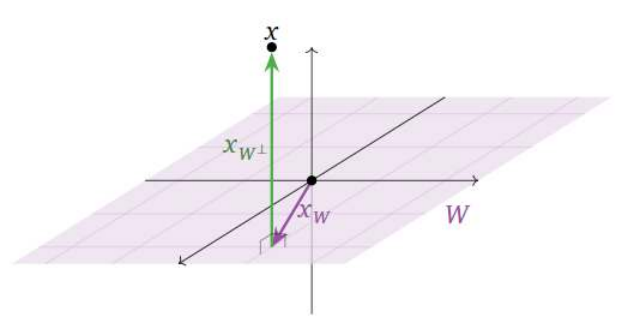


Illustration 3.6.2: Orthogonal decomposition of a vector in W

If $\vec{x}$ is in a subspace W, then the closest vector to $\vec{x}$ in W is itself, so $\vec{x} = \vec{x}_W$ and $\vec{x}_{W^\perp} = \vec{0}$. Conversely, if $\vec{x} = \vec{x}_W$ then $\vec{x}$ is contained in W because $\vec{x}_W$ is contained in W.

Illustration 3.6.3: Orthogonal decomposition of a vector in $W^\perp$

If W is a subspace and $\vec{x}$ is in $W^\perp$, then the orthogonal decomposition of $\vec{x}$ is $\vec{x} = \vec{0} + \vec{x}$, where $\vec{0}$ is in W and $\vec{x}$ is in $W^\perp$. It follows that $\vec{x}_W = \vec{0}$. Conversely, if $\vec{x}_W = \vec{0}$ then the orthogonal decomposition of $\vec{x}$ is $\vec{x} = \vec{x}_W + \vec{x}_{W^\perp} = \vec{0} + \vec{x}_{W^\perp}$, so $\vec{x} = \vec{x}_{W^\perp}$ is in $W^\perp$.

Now we turn to the problem of computing $\vec{x}_W$ and $\vec{x}_{W^\perp}$. Of course, since $\vec{x}_{W^\perp} = \vec{x} - \vec{x}_W$, really all we need is to compute $\vec{x}_W$. The following theorem gives a method for computing the orthogonal projection onto a column space. To compute the orthogonal projection onto a general subspace, usually it is best to rewrite the subspace as the column space of a matrix.

Theorem 3.6.2

Let A be an $m \times n$ matrix, let $W = \text{Col}(A)$, and let $\vec{x}$ be a vector in $\mathbb{R}^m$. Then the matrix equation

$$A^T A\vec{c} = A^T\vec{x}$$

in the unknown vector $\vec{c}$ is consistent, and $\vec{x}_W$ is equal to $A\vec{c}$ for any solution $\vec{c}$.

Proof

Let $\vec{x} = \vec{x}_W + \vec{x}_{W^\perp}$ be the orthogonal decomposition with respect to W. By definition $\vec{x}_W$ lies in $W = \text{Col}(A)$ and so there is a vector $\vec{c}$ in $\mathbb{R}^n$ with $A\vec{c} = \vec{x}_W$. Choose any such vector $\vec{c}$. We know that $\vec{x} - \vec{x}_W = \vec{x} - A\vec{c}$ lies in $W^\perp$, which is equal to $\text{Null}(A^T)$. We thus have

$$\vec{0} = A^T(\vec{x} - A\vec{c}) = A^T\vec{x} - A^T A\vec{c}$$

and so

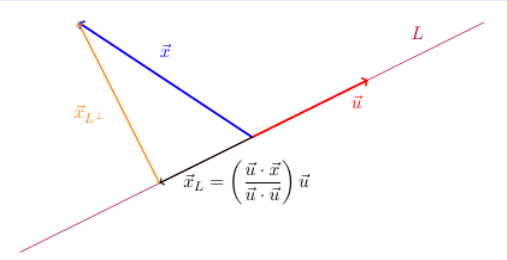
$$A^T A\vec{c} = A^T\vec{x}$$

This exactly means that $A^T A\vec{c} = A^T\vec{x}$ is consistent. If $\vec{c}$ is any solution to $A^T A\vec{c} = A^T\vec{x}$ then by reversing the above logic, we conclude that $\vec{x}_W = A\vec{c}$.

Illustration 3.6.4: Orthogonal Projection onto a Line

Let $L = \text{Span}\{\vec{u}\}$ be a line in $\mathbb{R}^n$ and let $\vec{x}$ be a vector in $\mathbb{R}^n$. By Theorem 3.6.2, to find $\vec{x}_L$ we must solve the matrix equation $u^T u\vec{c} = u^T\vec{x}$, where we regard u as an $n \times 1$ matrix with column vector $\vec{u}$ (the column space of this matrix is exactly L!). But $u^T u = \vec{u} \cdot \vec{u}$ and $u^T x = \vec{u} \cdot \vec{x}$, so $\vec{c} = \dfrac{\vec{u} \cdot \vec{x}}{\vec{u} \cdot \vec{u}}$ is a solution of $u^T u\vec{c} = u^T\vec{x}$, and hence

$$\vec{x}_L = u\vec{c} = \frac{\vec{u} \cdot \vec{x}}{\vec{u} \cdot \vec{u}}\vec{u}$$



To reiterate:

Recipe: Orthogonal Projection onto a Line

If $L = \text{Span}\{\vec{u}\}$ is a line, then $\vec{x}_L = \dfrac{\vec{u} \cdot \vec{x}}{\vec{u} \cdot \vec{u}}\vec{u}$ and $\vec{x}_{L^\perp} = \vec{x} - \vec{x}_L$ for any vector $\vec{x}$.

Example 3.6.1: Projection onto a line in $\mathbb{R}^2$

Compute the orthogonal projection of $\vec{x} = \begin{bmatrix} -6 \\ 4 \end{bmatrix}$ onto the line L spanned by $\vec{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and find the distance from $\vec{x}$ to L.

Solution

First we find

$$\vec{x}_L = \frac{\vec{x} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}\vec{u} = \frac{-6-4}{1+1}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = -5\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -5 \\ 5 \end{bmatrix}$$

and so $\vec{x}_{L^\perp} = \vec{x} - \vec{x}_L = \begin{bmatrix} -6 \\ 4 \end{bmatrix} - \begin{bmatrix} -5 \\ 5 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$

The distance from $\vec{x}$ to L is

$$\|\vec{x}_{L^\perp}\| = \sqrt{(-1)^2 + (-1)^2} = \boxed{\sqrt{2}}$$

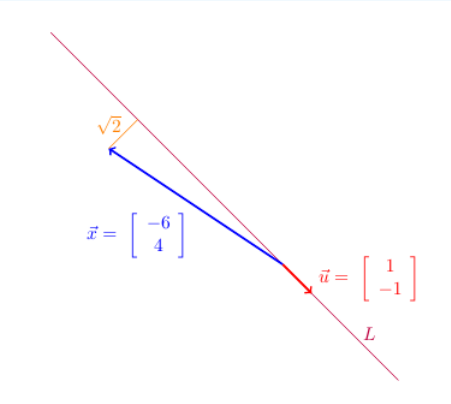


Illustration 3.6.5: Projection onto a line in $\mathbb{R}^3$

Let $\vec{x} = \begin{bmatrix} -2 \\ 3 \\ 1 \end{bmatrix}$ and $\vec{u} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, and let L be the line spanned by $\vec{u}$.

We will compute $\vec{x}_L$ and $\vec{x}_L^\perp$:

$$\vec{x}_L = \frac{\vec{x} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u} = \frac{2+3+1}{1+1+1} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 2 \end{bmatrix} \qquad \vec{x}_{L^\perp} = \vec{x} - \vec{x}_L = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

It is easy to see that $\vec{x}_L$ is on the line L and that $\vec{x}_L \cdot \vec{x}_{L^\perp} = 0$ so they are indeed perpendicular.

When A is a matrix with more than one column, computing the orthogonal projection of $\vec{x}$ onto $W = \text{Col}(A)$ means solving the matrix equation $A^TA\vec{c} = A^T\vec{x}$. In other words, we can compute the closest vector by *solving a system of linear equations*. To be explicit, we state Theorem 3.6.2 as a recipe:

Recipe: Compute an Orthogonal Decomposition

Let W be a subspace of $\mathbb{R}^m$. Here is a method to compute the orthogonal decomposition of a vector $\vec{x}$ with respect to W:

0. Rewrite W as the column space of a matrix A. In other words, find a spanning set for W, and let A be the matrix with those columns.
1. Compute the matrix A^TA and the vector $A^T\vec{x}$.
2. Form the augmented matrix for the matrix equation $A^TA\vec{c} = A^T\vec{x}$ in the unknown vector $\vec{c}$, and row reduce.
3. This equation is always consistent; choose one solution $\vec{c}$. Then

$$\vec{x}_W = A\vec{c} \qquad \vec{x}_{W^\perp} = \vec{x} - \vec{x}_W$$

Example 3.6.2: Projection onto a plane in $\mathbb{R}^3$

Let

$$W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\} \qquad \vec{x} = \begin{bmatrix} 6 \\ -7 \\ 2 \end{bmatrix}$$

Compute $\vec{x}_W$ and the distance from $\vec{x}$ to W.

Solution

We have to solve the matrix equation $A^TA\vec{c} = A^T\vec{x}$, where $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}$

We have

$$A^TA = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad A^T\vec{x} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 6 \\ -7 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \end{bmatrix}$$

We form an augmented matrix and row reduce:

$$\left[\begin{array}{cc|c} 2 & 1 & 4 \\ 1 & 2 & -1 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & -2 \end{array}\right] \implies \vec{c} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

It follows that

$$\vec{x}_W = A\vec{c} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \boxed{\begin{bmatrix} 1 \\ -2 \\ -3 \end{bmatrix}} \qquad \vec{x}_{W^\perp} = \vec{x} - \vec{x}_W = \begin{bmatrix} 5 \\ -5 \\ 5 \end{bmatrix}$$

The distance from $\vec{x}$ to W is

$$\|\vec{x}_{W^\perp}\| = \sqrt{5^2 + (-5)^2 + 5^2} = \boxed{5\sqrt{3}}$$

Illustration 3.6.6: Projection onto another plane in $\mathbb{R}^3$

Let

$$W = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : x_1 - 2x_2 = x_3 \right\} \quad \text{and} \quad \vec{x} = \begin{bmatrix} 18 \\ 19 \\ 4 \end{bmatrix}$$

We attempt to compute $\vec{x}_W$ two ways:

Method 1: First we need to find a spanning set for W. We notice that W is the solution set of the homogeneous equation $x_1 - 2x_2 - x_3 = 0$, so $W = \text{Null}\left(\begin{bmatrix} 1 & -2 & -1 \end{bmatrix}\right)$. We know how to compute a basis for a null space: we row reduce and find the parametric vector form. The matrix $\begin{bmatrix} 1 & -2 & -1 \end{bmatrix}$ is already in reduced row echelon form. The parametric form is $x_1 = 2x_2 + x_3$, so the parametric vector form is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

and hence a basis for V is given by

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

We let A be the matrix whose columns are our basis vectors: $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

Hence $\text{Col}(A) = \text{Null}\left(\begin{bmatrix} 1 & -2 & -1 \end{bmatrix}\right) = W$.

Now we can continue with step 1 of the recipe. We compute

$$A^TA = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix} \qquad A^T\vec{x} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 18 \\ 19 \\ 4 \end{bmatrix} = \begin{bmatrix} 55 \\ 22 \end{bmatrix}$$

We write the linear system $A^TA\vec{c} = A^T\vec{x}$ as an augmented matrix and row reduce:

$$\left[\begin{array}{cc|c} 5 & 2 & 55 \\ 2 & 2 & 22 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & 11 \\ 0 & 1 & 0 \end{array}\right]$$

Hence we can take $\vec{c} = \begin{bmatrix} 11 \\ 0 \end{bmatrix}$, so

$$\vec{x}_W = A\vec{c} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 11 \\ 0 \end{bmatrix} = \begin{bmatrix} 22 \\ 11 \\ 0 \end{bmatrix}$$

Method 2: In this case, it is easier to compute $\vec{x}_{W^\perp}$. Indeed, since $W = \text{Null}\left(\begin{bmatrix} 1 & -2 & -1 \end{bmatrix}\right)$, the orthogonal complement is the line

$$L = W^\perp = \text{Col}\left(\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}\right\}$$

Using the formula for projection onto a line spanned by $\vec{u} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$ gives:

$$\vec{x}_{W^\perp} = \vec{x}_V = \left(\frac{\vec{u}\cdot\vec{x}}{\vec{u}\cdot\vec{u}}\right)\vec{u} = \frac{-24}{6}\vec{u} = -4\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ 8 \\ 4 \end{bmatrix}$$

Hence we have

$$\vec{x}_W = \vec{x} - \vec{x}_{W^\perp} = \begin{bmatrix} 18 \\ 19 \\ 4 \end{bmatrix} - \begin{bmatrix} -4 \\ 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 22 \\ 11 \\ 0 \end{bmatrix}$$

as above.

3.6: Orthogonal Decomposition is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.3: Orthogonal Projection** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.6: Least Squares Solutions

We begin by clarifying exactly what we will mean by a "best-approximate solution" to an inconsistent matrix equation $A\vec{x} = \vec{b}$.

Definition: Least-Squares Solution

Let A be an $m \times n$ matrix and let $\vec{b}$ be a vector in $\mathbb{R}^m$. A ***least-squares solution*** of the matrix equation $A\vec{x} = \vec{b}$ is a vector $\hat{x}$ in $\mathbb{R}^n$ such that

$$\text{dist}(\vec{b},\, A\hat{x}) \leq \text{dist}(\vec{b},\, A\vec{x})$$

for all other vectors $\vec{x}$ in $\mathbb{R}^n$.

Recall that $\text{dist}(\vec{v}, \vec{w}) = \|\vec{v} - \vec{w}\|$ is the distance between the vectors $\vec{v}$ and $\vec{w}$. The term "least-squares" comes from the fact that $\text{dist}(\vec{b}, A\hat{x}) = \|\vec{b} - A\hat{x}\|$ is the square root of the sum of the squares of the entries of the vector $\vec{b} - A\hat{x}$. So a least-squares solution minimizes the sum of the squares of the differences between the entries of $A\hat{x}$ and $\vec{b}$. In other words, a least-squares solution solves the equation $A\vec{x} = \vec{b}$ as closely as possible, in the sense that the sum of the squares of the difference $\vec{b} - A\vec{x}$ is minimized.

Suppose that the equation $A\vec{x} = \vec{b}$ is inconsistent. Recall that the column space of A is the set of all other vectors, $\vec{c}$, such that $A\vec{x} = \vec{c}$ is consistent. In other words, $\text{Col}(A)$ is the set of all vectors of the form $A\vec{x}$. Hence, the closest vector (of the form $A\vec{x}$) to $\vec{b}$ is the orthogonal projection of $\vec{b}$ onto $\text{Col}(A)$. This is denoted $\vec{b}_{\text{Col}(A)}$:

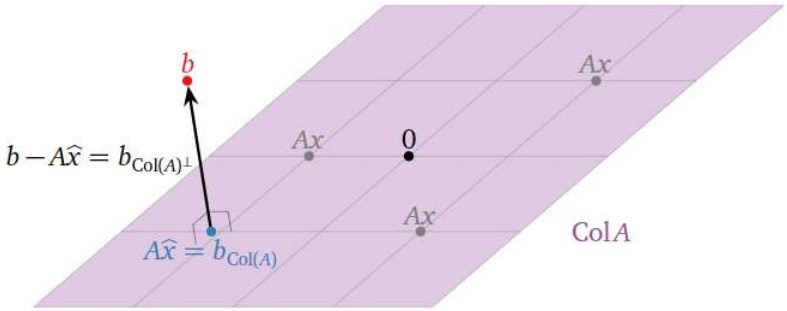


Figure 3.6.2

Note 3.6.2

A least-squares solution of $A\vec{x} = \vec{b}$ is a solution $\hat{x}$ of the consistent equation $A\vec{x} = \vec{b}_{\text{Col}(A)}$

Note 3.6.3

If $A\vec{x} = \vec{b}$ is consistent, then $\vec{b}_{\text{Col}(A)} = \vec{b}$, so that a least-squares solution is the same as a usual solution.

Where is $\hat{x}$ in this picture? If $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ are the columns of A, then

$$A\hat{x} = A\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \vdots \\ \hat{x}_n \end{bmatrix} = \hat{x}_1\vec{v}_1 + \hat{x}_2\vec{v}_2 + \cdots + \hat{x}_n\vec{v}_n$$

Hence the entries of $\hat{x}$ are the "coordinates" of $\vec{b}_{\text{Col}(A)}$ with respect to the spanning set $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ of $\text{Col}(A)$. (They are honest $\mathcal{B}$-coordinates if the columns of A are linearly independent.)

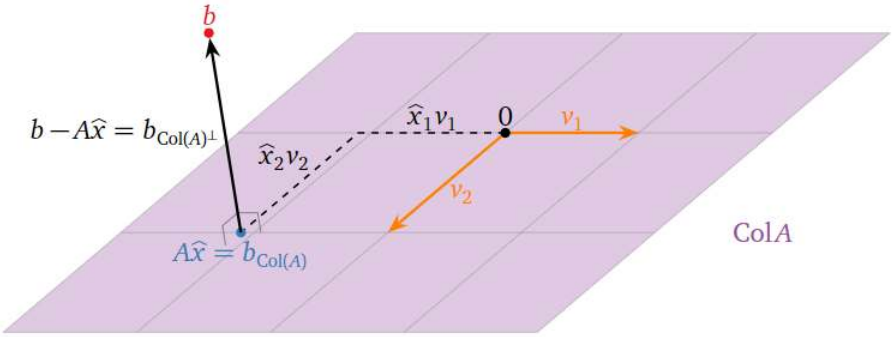


Figure 3.6.3

Theorem 3.6.3

Let A be an $m \times n$ matrix and let $\vec{b}$ be a vector in $\mathbb{R}^m$. The least-squares solutions of $A\vec{x} = \vec{b}$ are the solutions of the matrix equation

$$A^T A\vec{x} = A^T\vec{b}$$

Proof

This is an immediate result of our work on orthogonal projections above. (In fact, you could argue it was the main reason we learned how to do orthogonal projections in the first place!)

To reiterate: once you have found a least-squares solution $\hat{x}$ of $A\vec{x} = \vec{b}$, then $\vec{b}_{\mathrm{Col}(A)}$ is equal to $A\hat{x}$.

Illustration 3.6.7

We will find the least-squares solutions of $A\vec{x} = \vec{b}$ where:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

We have

$$A^T A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix}$$

and

$$A^T\vec{b} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$$

We form an augmented matrix and row reduce:

$$\left[\begin{array}{cc|c} 5 & 3 & 0 \\ 3 & 3 & 6 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 5 \end{array}\right]$$

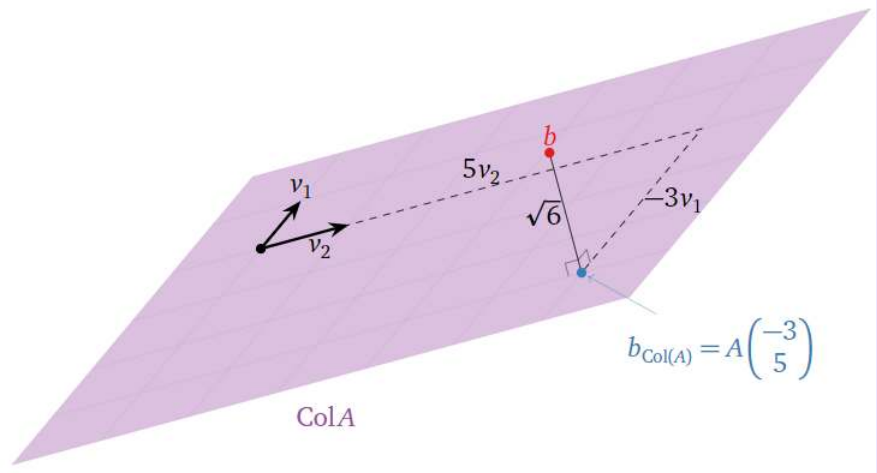
Therefore, the only least-squares solution is $\hat{x} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$

This solution minimizes the distance from $A\hat{x}$ to $\vec{b}$, i.e., the sum of the squares of the entries of $\vec{b} - A\hat{x} = \vec{b} - \vec{b}_{\mathrm{Col}(A)} = \vec{b}_{\mathrm{Col}(A)^\perp}$. In this case, we have

$$||\vec{b} - A\hat{x}|| = \left|\left|\begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}\right|\right| = \left|\left|\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right|\right| = \sqrt{1^2 + (-2)^2 + 1^2} = \sqrt{6}$$

Therefore, $\vec{b}_{\text{Col}(A)} = A\hat{x}$ is $\sqrt{6}$ units from $\vec{b}$

In the following picture, $\vec{v}_1, \vec{v}_2$ are the columns of A:



Example 3.6.3

Find the least-squares solutions of $A\vec{x} = \vec{b}$ where:

$$A = \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix}$$

Solution

We have

$$A^T A = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ -1 & 5 \end{bmatrix}$$

and

$$A^T\vec{b} = \begin{bmatrix} 2 & -1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

We form an augmented matrix and row reduce:

$$\left[\begin{array}{cc|c} 5 & -1 & 6 \\ -1 & 5 & -6 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -1 \end{array}\right]$$

Therefore, the only least-squares solution is $\boxed{\hat{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}}$

The reader may have noticed that we have been careful to say "the least-squares solutions" in the plural, and "a least-squares solution" using the indefinite article. This is because a least-squares solution need not be unique: indeed, if the columns of A are linearly dependent, then $A\vec{x} = \vec{b}_{\text{Col}(A)}$ has infinitely many solutions. The following theorem gives equivalent criteria for uniqueness.

Theorem 3.6.4

Let A be an $m \times n$ matrix and let $\vec{b}$ be a vector in $\mathbb{R}^m$. The following are equivalent:

1. $A\vec{x} = \vec{b}$ has a unique least-squares solution.
2. The columns of A are linearly independent.
3. A^TA is invertible.

In this case, the least-squares solution is

$$\hat{x} = (A^TA)^{-1}A^T\vec{b}$$

Proof

The set of least-squares solutions of $A\vec{x} = \vec{b}$ is the solution set of the consistent equation $A^TA\vec{x} = A^T\vec{b}$, which is a translate of the solution set of the homogeneous equation $A^TA\vec{x} = \vec{0}$. Since A^TA is a square matrix, the equivalence of 1 and 3 follows from the invertible matrix theorem. The set of least-squares solutions is also the solution set of the consistent equation $A\vec{x} = \vec{b}_{\text{Col}(A)}$, which has a unique solution if and only if the columns of A are linearly independent.

Example 3.6.4: Infinitely many least-squares solutions

Find the least-squares solutions of $A\vec{x} = \vec{b}$ where:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 2 & -3 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

Solution

We have

$$A^TA = \begin{bmatrix} 3 & 3 & -3 \\ 3 & 5 & -7 \\ -3 & -7 & 11 \end{bmatrix} \qquad A^T\vec{b} = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix}$$

We form an augmented matrix and row reduce:

$$\left[\begin{array}{ccc|c} 3 & 3 & -3 & 6 \\ 3 & 5 & -7 & 0 \\ -3 & -7 & 11 & 6 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

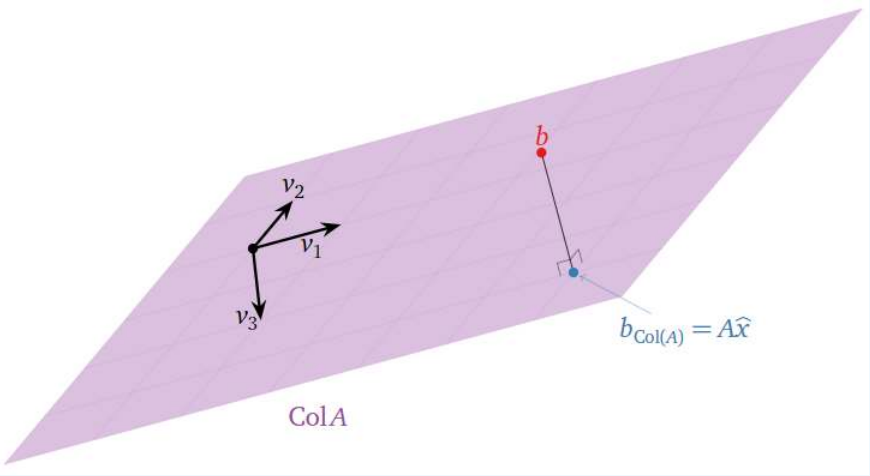
The free variable is x_3, so the solution set is

$$\begin{cases} x_1 = -x_3 + 5 \\ x_2 = 2x_3 - 3 \\ x_3 = x_3 \end{cases} \xrightarrow[\text{vector form}]{\text{parametric}} \boxed{\hat{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix}}$$

For example, taking $x_3 = 0$ and $x_3 = 1$ gives the least-squares solutions

$$\hat{x}_0 = \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix} \quad \text{and} \quad \hat{x}_1 = \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix}$$

Geometrically, we see that the columns $\vec{v}_1, \vec{v}_2, \vec{v}_3$ of A are coplanar:



Therefore, there are many ways of writing $b_{\mathrm{Col}(A)}$ as a linear combination of $\vec{v}_1, \vec{v}_2, \vec{v}_3$:

$$b_{\mathrm{Col}(A)} = A\hat{x}_0 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 5 \\ -3 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

$$b_{\mathrm{Col}(A)} = A\hat{x}_1 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 2 & -3 \end{bmatrix} \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

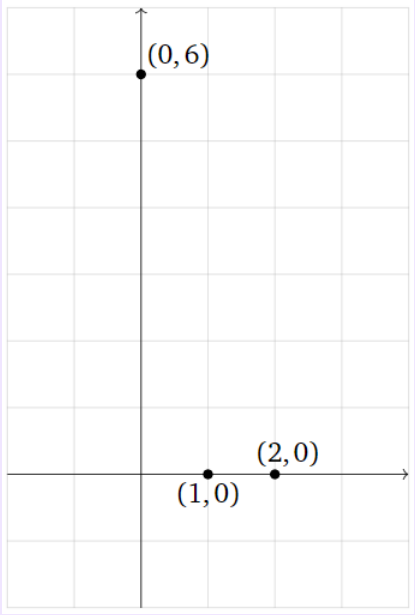
Now we give an application of the method of least-squares to data modeling. We begin with a basic example.

Illustration 3.6.8: Best-Fit Line

Suppose that we have measured three data points

$$(0,6), \quad (1,0), \quad (2,0)$$

and that our model for these data asserts that the points should lie on a line. Of course, these three points do not actually lie on a single line, but this could be due to errors in our measurement. How do we predict which line they are supposed to lie on?



The general equation for a (non-vertical) line is

$$y = Mx + B$$

If our three data points were to lie on this line, then the following equations would be satisfied:

$$\begin{aligned} 6 &= M \cdot 0 + B \\ 0 &= M \cdot 1 + B \\ 0 &= M \cdot 2 + B \end{aligned}$$

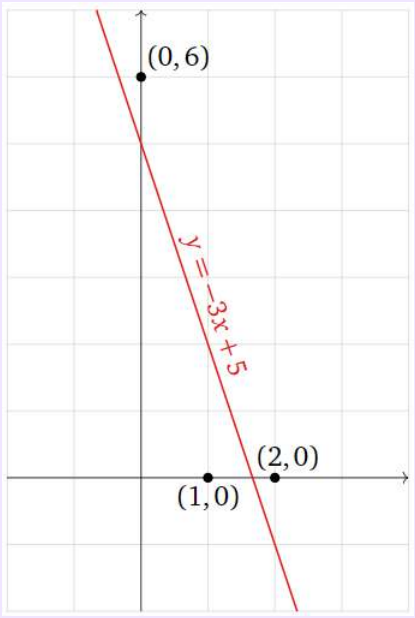
In order to find the best-fit line, we try to solve the above equations in the unknowns M and B. As the three points do not actually lie on a line, there is no actual solution, so instead we compute a least-squares solution.

Putting our linear equations into matrix form, we are trying to solve $A\vec{x} = \vec{b}$ for

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \qquad x = \begin{bmatrix} M \\ B \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

We solved this least-squares problem in Illustration 3.6.7: the only least-squares solution to $A\vec{x} = \vec{b}$ is $\hat{x} = \begin{bmatrix} M \\ B \end{bmatrix} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$, so the best-fit line is

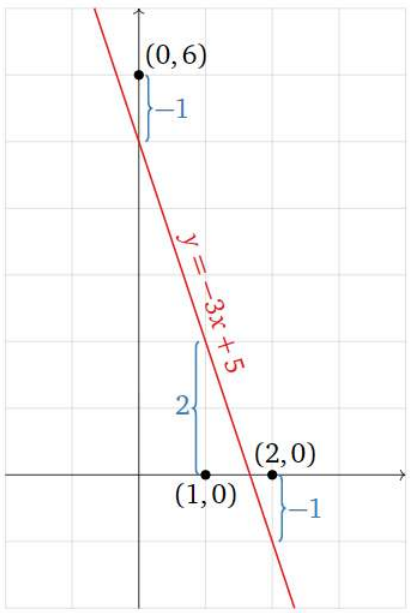
$$y = -3x + 5$$



What exactly is the line $y = f(x) = -3x + 5$ minimizing? The least-squares solution $\hat{x}$ minimizes the sum of the squares of the entries of the vector $\vec{b} - A\hat{x}$. The vector $\vec{b}$ is the left-hand side of the original system, and

$$A \begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} -3(0) + 5 \\ -3(1) + 5 \\ -3(2) + 5 \end{bmatrix} = \begin{bmatrix} f(0) \\ f(1) \\ f(2) \end{bmatrix}$$

In other words, $A\hat{x}$ is the vector whose entries are the y-coordinates of the graph of the line at the values of x we specified in our data points, and $\vec{b}$ is the vector whose entries are the y-coordinates of those data points. The difference $\vec{b} - A\hat{x}$ is the vertical distance of the graph from the data points:



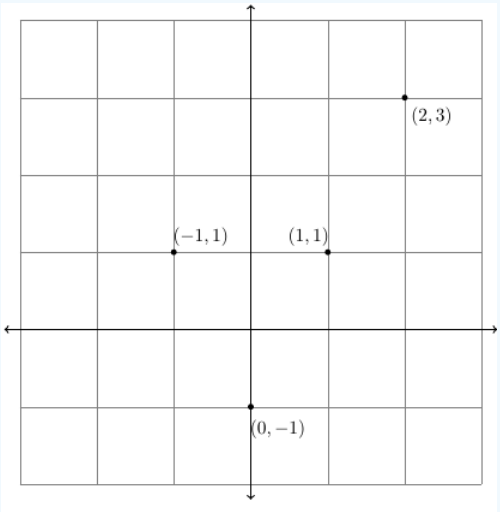
$$\vec{b} - A\hat{x} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} - A \begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

The best-fit line minimizes the sum of the squares of these vertical distances.

Example 3.6.5: Best-fit parabola

Find the parabola that best approximates the data points

$$(-1,\ 1),\quad (0,\ -1),\quad (1,\ 1),\quad (2,\ 3)$$



What quantity is being minimized?

Solution

The general equation for a parabola is

$$y = Bx^2 + Cx + D$$

If the four points were to lie on this parabola, then the following equations would be satisfied:

$$\begin{aligned} 1 &= B(-1)^2 + C(-1) + D \\ -1 &= B(0)^2 + C(0) + D \\ 1 &= B(1)^2 + C(1) + D \\ 3 &= B(2)^2 + C(2) + D \end{aligned}$$

We treat this as a system of equations in the unknowns B, C, D. In matrix form, we can write this as $A\vec{x} = \vec{b}$ for

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} B \\ C \\ D \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 3 \end{bmatrix}$$

We find a least-squares solution by multiplying both sides by the transpose:

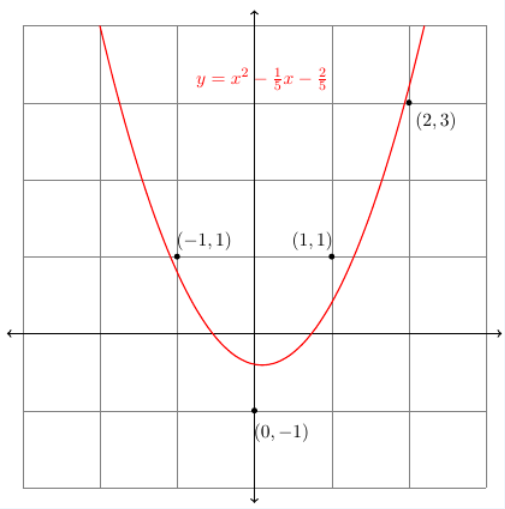
$$A^TA = \begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix} \qquad A^T\vec{b} = \begin{bmatrix} 14 \\ 6 \\ 4 \end{bmatrix}$$

then forming an augmented matrix and row reducing:

$$\left[\begin{array}{ccc|c} 18 & 8 & 6 & 14 \\ 8 & 6 & 2 & 6 \\ 6 & 2 & 4 & 4 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{1}{5} \\ 0 & 0 & 1 & -\frac{2}{5} \end{array}\right] \implies \hat{x} = \begin{bmatrix} 1 \\ -\frac{1}{5} \\ -\frac{2}{5} \end{bmatrix}$$

The best-fit parabola is

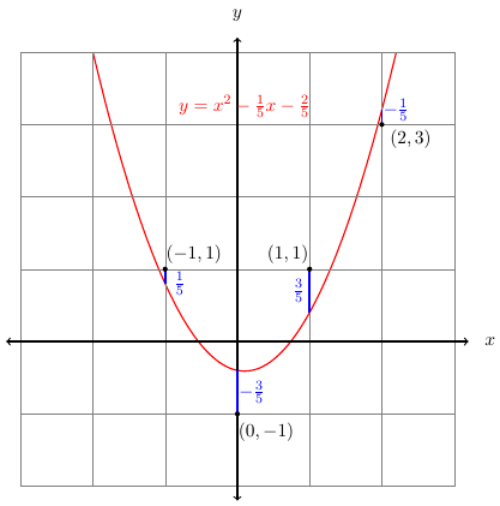
$$\boxed{y = 1x^2 - \frac{1}{5}x - \frac{2}{5}}$$



Now we consider what exactly the parabola $y = f(x)$ is minimizing. The least-squares solution $\hat{x}$ minimizes the sum of the squares of the entries of the vector $\vec{b} - A\hat{x}$. The vector $\vec{b}$ is the left-hand side of the original system, and

$$A\hat{x} = \begin{bmatrix} (-1)^2 - \frac{1}{5}(-1) - \frac{2}{5} \\ (0)^2 - \frac{1}{5}(0) - \frac{2}{5} \\ (1)^2 - \frac{1}{5}(1) - \frac{2}{5} \\ (2)^2 - \frac{1}{5}(2) - \frac{2}{5} \end{bmatrix} = \begin{bmatrix} \frac{4}{5} \\ -\frac{2}{5} \\ \frac{2}{5} \\ \frac{16}{5} \end{bmatrix} = \begin{bmatrix} f(-1) \\ f(0) \\ f(1) \\ f(2) \end{bmatrix}$$

In other words, $A\hat{x}$ is the vector whose entries are the y-coordinates of the graph of the parabola at the values of x we specified in our data points, and $\vec{b}$ is the vector whose entries are the y-coordinates of those data points. The difference $\vec{b} - A\hat{x}$ is the vertical distance of the graph from the data points:



$$\vec{b} - A\hat{x} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} \frac{4}{5} \\ -\frac{2}{5} \\ \frac{2}{5} \\ \frac{16}{5} \end{bmatrix} = \begin{bmatrix} \frac{1}{5} \\ -\frac{3}{5} \\ \frac{3}{5} \\ -\frac{1}{5} \end{bmatrix}$$

The best-fit parabola minimizes the sum of the squares of these vertical distances

Note 3.6.4

Gauss invented the method of least-squares to find a best-fit ellipse: he correctly predicted the (elliptical) orbit of the asteroid Ceres as it passed behind the sun in 1801. People struggle to do this (even with fancy calculators) more than two centuries later.

3.6: Least Squares Solutions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.5: The Method of Least Squares** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.6: Practice Set

Exercise 3.6.1a

Find the projection of $\vec{x} = \begin{bmatrix} 0 \\ 10 \end{bmatrix}$ onto the subspace $W = \text{Span}\left\{\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right\}$

Answer

$$\vec{x}_W = \left(\frac{\vec{x}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\right)\vec{w} = \left(\frac{20}{5}\right)\vec{w} = 4\vec{w} = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$$

Exercise 3.6.1b

Find the projection of $\vec{x} = \begin{bmatrix} -5 \\ -1 \\ 0 \end{bmatrix}$ onto the subspace $W = \text{Span}\left\{\begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}\right\}$

Answer

$$\vec{x}_W = \left(\frac{\vec{x}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\right)\vec{w} = \left(\frac{-14}{14}\right)\vec{w} = -1\vec{w} = \begin{bmatrix} -3 \\ 1 \\ -2 \end{bmatrix}$$

Exercise 3.6.1c

Find the projection of $\vec{x} = \begin{bmatrix} -2 \\ 5 \\ 7 \end{bmatrix}$ onto the subspace $W = \text{Span}\left\{\begin{bmatrix} 4 \\ 3 \\ -1 \end{bmatrix}\right\}$

Answer

$$\vec{x}_W = \left(\frac{\vec{x}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\right)\vec{w} = \left(\frac{0}{26}\right)\vec{w} = 0\vec{w} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Exercise 3.6.1d

Find the projection of $\vec{x} = \begin{bmatrix} 6 \\ 1 \\ -2 \\ 5 \end{bmatrix}$ onto the subspace $W = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right\}$

Answer

$$\vec{x}_W = \left(\frac{\vec{x}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\right)\vec{w} = \left(\frac{10}{4}\right)\vec{w} = \frac{5}{2}\vec{w} = \begin{bmatrix} \frac{5}{2} \\ \frac{5}{2} \\ \frac{5}{2} \\ \frac{5}{2} \end{bmatrix}$$

Exercise 3.6.2a

Find the projection of $\vec{x} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$ onto the subspace $W = \text{Span}\left\{\begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right\}$

Answer

$\begin{bmatrix} 3 & 1 \end{bmatrix}$

$A = \begin{bmatrix} 3 & 1 \\ 1 & 1 \\ -1 & 0 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix}$ and $A^T\vec{x} = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$ and $A^TA = \begin{bmatrix} 3 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 11 & 4 \\ 4 & 2 \end{bmatrix}$

We solve $\left[\begin{array}{cc|c} 11 & 4 & 5 \\ 4 & 2 & 4 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 4 \end{array}\right] \rightarrow \vec{c} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$

So our answer is $\vec{x}_W = A\vec{c} = \begin{bmatrix} 3 & 1 \\ 1 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}$

Exercise 3.6.2b

Find the projection of $\vec{x} = \begin{bmatrix} 7 \\ 1 \\ 3 \end{bmatrix}$ onto the subspace $W = \text{Span}\left\{\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}\right\}$

Answer

$A = \begin{bmatrix} 1 & 1 \\ 4 & 2 \\ 2 & 4 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 1 & 4 & 2 \\ 1 & 2 & 4 \end{bmatrix}$

$A^T\vec{x} = \begin{bmatrix} 1 & 4 & 2 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 7 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 17 \\ 21 \end{bmatrix}$ and $A^TA = \begin{bmatrix} 1 & 4 & 2 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 4 & 2 \\ 2 & 4 \end{bmatrix} = \begin{bmatrix} 21 & 17 \\ 17 & 21 \end{bmatrix}$

We solve $\left[\begin{array}{cc|c} 21 & 17 & 17 \\ 17 & 21 & 21 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 1 \end{array}\right] \rightarrow \vec{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

So our answer is $\vec{x}_W = A\vec{c} = \begin{bmatrix} 1 & 1 \\ 4 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$

Exercise 3.6.2c

Find the projection of $\vec{x} = \begin{bmatrix} 7 \\ 5 \\ 18 \\ -10 \end{bmatrix}$ onto the subspace $W = \text{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ 5 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ -2 \end{bmatrix}\right\}$

Answer

$A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 5 & 0 \\ -3 & -2 \end{bmatrix} \rightarrow A^T = \begin{bmatrix} 1 & -1 & 5 & -3 \\ 2 & 1 & 0 & -2 \end{bmatrix}$

$A^T\vec{x} = \begin{bmatrix} 1 & -1 & 5 & -3 \\ 2 & 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 7 \\ 5 \\ 18 \\ -10 \end{bmatrix} = \begin{bmatrix} 122 \\ 39 \end{bmatrix}$ and $A^TA = \begin{bmatrix} 1 & -1 & 5 & -3 \\ 2 & 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 5 & 0 \\ -3 & -2 \end{bmatrix} = \begin{bmatrix} 36 & 7 \\ 7 & 9 \end{bmatrix}$

We solve $\left[\begin{array}{cc|c} 36 & 7 & 122 \\ 7 & 9 & 39 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 2 \end{array}\right] \rightarrow \vec{c} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$

So our answer is $\vec{x}_W = A\vec{c} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 5 & 0 \\ -3 & -2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \\ 15 \\ -13 \end{bmatrix}$

Exercise 3.6.2d

Find the projection of $\vec{x} = \begin{bmatrix} 0 \\ -\frac{2}{3} \\ -\frac{2}{3} \\ \frac{4}{3} \end{bmatrix}$ onto the subspace $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ -6 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} \right\}$

Answer

$$A = \begin{bmatrix} 1 & 3 & 1 \\ 2 & -1 & 1 \\ -6 & 1 & -1 \\ -2 & 0 & 0 \end{bmatrix} \to A^T = \begin{bmatrix} 1 & 2 & -6 & -2 \\ 3 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{bmatrix}$$

$$A^T\vec{x} = \begin{bmatrix} 1 & 2 & -6 & -2 \\ 3 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{2}{3} \\ -\frac{2}{3} \\ \frac{4}{3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 1 & 2 & -6 & -2 \\ 3 & -1 & 1 & 0 \\ 1 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & 1 \\ 2 & -1 & 1 \\ -6 & 1 & -1 \\ -2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 45 & -5 & 9 \\ -5 & 11 & 1 \\ 9 & 1 & 3 \end{bmatrix}$$

We solve $\left[\begin{array}{ccc|c} 45 & -5 & 9 & 0 \\ -5 & 11 & 1 & 0 \\ 9 & 1 & 3 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right] \to \vec{c} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

So our answer is $\vec{x}_W = A\vec{c} = \begin{bmatrix} 1 & 3 & 1 \\ 2 & -1 & 1 \\ -6 & 1 & -1 \\ -2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

(This is because $\vec{x}$ is orthogonal to all vectors in W and resides entirely in $W^\perp$.)

Exercise 3.6.2e

Find the projection of $\vec{x} = \begin{bmatrix} 9 \\ 32 \\ -7 \\ -11 \end{bmatrix}$ onto the subspace $W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 3 \\ 4 \\ -5 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 7 \\ -3 \\ -2 \end{bmatrix} \right\}$

Answer

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 3 & 4 & 7 \\ 4 & -1 & -3 \\ -5 & 0 & -2 \end{bmatrix} \to A^T = \begin{bmatrix} 2 & 3 & 4 & -5 \\ 2 & 4 & -1 & 0 \\ 1 & 7 & -3 & -2 \end{bmatrix}$$

$$A^T\vec{x} = \begin{bmatrix} 2 & 3 & 4 & -5 \\ 2 & 4 & -1 & 0 \\ 1 & 7 & -3 & -2 \end{bmatrix} \begin{bmatrix} 9 \\ 32 \\ -7 \\ -11 \end{bmatrix} = \begin{bmatrix} 141 \\ 153 \\ 276 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 2 & 3 & 4 & -5 \\ 2 & 4 & -1 & 0 \\ 1 & 7 & -3 & -2 \end{bmatrix} \begin{bmatrix} 2 & 2 & 1 \\ 3 & 4 & 7 \\ 4 & -1 & -3 \\ -5 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 54 & 12 & 21 \\ 12 & 21 & 33 \\ 21 & 33 & 63 \end{bmatrix}$$

We solve $\left[\begin{array}{ccc|c} 54 & 12 & 21 & 141 \\ 12 & 21 & 33 & 153 \\ 21 & 33 & 63 & 276 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right] \to \vec{c} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

So our answer is $\vec{x}_W = A\vec{c} = \begin{bmatrix} 2 & 2 & 1 \\ 3 & 4 & 7 \\ 4 & -1 & -3 \\ -5 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 32 \\ -7 \\ -11 \end{bmatrix}$

(This is because $\vec{x}$ is entirely contained inside of W and must equal its own projection.)

Exercise 3.6.3a

Find the line of best fit for the points $(1,4),(2,-1)$ and $(3,3)$

Answer

We want $y=mx+b$, so we set $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = m\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix} = m\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix}$$

The matrix $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$ has $A^T = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}$

$$A^T\vec{y} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 6 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & 6 \\ 6 & 3 \end{bmatrix}$$

Solving $\left[\begin{array}{cc|c} 14 & 6 & 11 \\ 6 & 3 & 6 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 3 \end{array}\right]$ tells us that $\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 3 \end{bmatrix}$

Therefore the line of best fit is $y = -\frac{1}{2}x + 3$

Exercise 3.6.3b

Find the line of best fit for the points $(1,3),(0,3)$ and $(3,-4)$

Answer

We want $y=mx+b$, so we set $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = m\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 3 \\ 3 \\ -4 \end{bmatrix} = m\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ -4 \end{bmatrix}$$

The matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 3 & 1 \end{bmatrix}$ has $A^T = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \end{bmatrix}$

$$A^T\vec{y} = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 3 \\ -4 \end{bmatrix} = \begin{bmatrix} -9 \\ 2 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 4 \\ 4 & 3 \end{bmatrix}$$

Solving $\left[\begin{array}{cc|c} 10 & 4 & -9 \\ 4 & 3 & 2 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} 1 & 0 & -\frac{5}{2} \\ 0 & 1 & 4 \end{array}\right]$ tells us that $\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} -\frac{5}{2} \\ 4 \end{bmatrix}$

Therefore the line of best fit is $y = -\frac{5}{2}x + 4$

Exercise 3.6.3c

Find the line of best fit for the points $(1,2),(1,3),(2,-1)$ and $(2,-3)$

Answer

We want $y=mx+b$, so we set $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = m\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 2 \\ 3 \\ -1 \\ -3 \end{bmatrix} = m \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix} + b \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -1 \\ -3 \end{bmatrix}$$

The matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 1 \\ 2 & 1 \end{bmatrix}$ has $A^T = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

$$A^T\vec{y} = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ -1 \\ -3 \end{bmatrix} = \begin{bmatrix} -3 \\ 1 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 2 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 6 \\ 6 & 4 \end{bmatrix}$$

Solving $\left[\begin{array}{cc|c} 10 & 6 & -3 \\ 6 & 4 & 1 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{cc|c} 1 & 0 & -\frac{9}{2} \\ 0 & 1 & 7 \end{array}\right]$ tells us that $\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} -\frac{9}{2} \\ 7 \end{bmatrix}$

Therefore the line of best fit is $y = -\frac{9}{2}x + 7$

Exercise 3.6.4a

Find the plane of best fit for the points $(0, 2, 10), (-1, 0, 20), (1, 4, -30)$ and $(1, -1, 10)$

Answer

We want $z = ax + by + c$, so we set $\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = a \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + b \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 10 \\ 20 \\ -30 \\ 10 \end{bmatrix} = a \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 2 \\ 0 \\ 4 \\ -1 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 0 & 2 & 1 \\ -1 & 0 & 1 \\ 1 & 4 & 1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 10 \\ 20 \\ -30 \\ 10 \end{bmatrix}$$

The matrix $A = \begin{bmatrix} 0 & 2 & 1 \\ -1 & 0 & 1 \\ 1 & 4 & 1 \\ 1 & -1 & 1 \end{bmatrix}$ has $A^T = \begin{bmatrix} 0 & -1 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

$$A^T\vec{z} = \begin{bmatrix} 0 & -1 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 10 \\ 20 \\ -30 \\ 10 \end{bmatrix} = \begin{bmatrix} -40 \\ -110 \\ 10 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 0 & -1 & 1 & 1 \\ 2 & 0 & 4 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ -1 & 0 & 1 \\ 1 & 4 & 1 \\ 1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 1 \\ 3 & 21 & 5 \\ 1 & 5 & 4 \end{bmatrix}$$

Solving $\left[\begin{array}{ccc|c} 3 & 3 & 1 & -40 \\ 3 & 21 & 5 & -110 \\ 1 & 5 & 4 & 10 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -11 \\ 0 & 1 & 0 & -7 \\ 0 & 0 & 1 & 14 \end{array}\right]$ tells us that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -11 \\ -7 \\ 14 \end{bmatrix}$

Therefore the plane of best fit is $z = -11x - 7y + 14$

Exercise 3.6.4b

Find the plane of best fit for the points $(0, 1, 0), (-1, 0, 0), (1, 1, 0)$ and $(0, -1, 18)$

Answer

We want $z = ax + by + c$, so we set $\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = a \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + b \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix}0\\0\\0\\18\end{bmatrix} = a\begin{bmatrix}0\\-1\\1\\0\end{bmatrix} + b\begin{bmatrix}1\\0\\1\\-1\end{bmatrix} + c\begin{bmatrix}1\\1\\1\\1\end{bmatrix} \longleftrightarrow \begin{bmatrix}0&1&1\\-1&0&1\\1&1&1\\0&-1&1\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}0\\0\\0\\18\end{bmatrix}$$

The matrix $A = \begin{bmatrix}0&1&1\\-1&0&1\\1&1&1\\0&-1&1\end{bmatrix}$ has $A^T = \begin{bmatrix}0&-1&1&0\\1&0&1&-1\\1&1&1&1\end{bmatrix}$

$$A^T\vec{z} = \begin{bmatrix}0&-1&1&0\\1&0&1&-1\\1&1&1&1\end{bmatrix}\begin{bmatrix}0\\0\\0\\18\end{bmatrix} = \begin{bmatrix}0\\-18\\18\end{bmatrix} \text{ and } A^TA = \begin{bmatrix}0&-1&1&0\\1&0&1&-1\\1&1&1&1\end{bmatrix}\begin{bmatrix}0&1&1\\-1&0&1\\1&1&1\\0&-1&1\end{bmatrix} = \begin{bmatrix}2&1&0\\1&3&1\\0&1&4\end{bmatrix}$$

Solving $\left[\begin{array}{ccc|c}2&1&0&0\\1&3&1&-18\\0&1&4&18\end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c}1&0&0&5\\0&1&0&-10\\0&0&1&7\end{array}\right]$ tells us that $\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}5\\-10\\7\end{bmatrix}$

Therefore the plane of best fit is $z = 5x - 10y + 7$

Exercise 3.6.4c

Find the plane of best fit for the points $(1,1,-3), (1,-1,-3), (0,1,0), (1,0,3)$ and $(0,-2,0)$

Answer

We want $z = ax + by + c$, so we set $\begin{bmatrix}z_1\\z_2\\z_3\\z_4\\z_5\end{bmatrix} = a\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix} + b\begin{bmatrix}y_1\\y_2\\y_3\\y_4\\y_5\end{bmatrix} + c\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}$

$$\begin{bmatrix}-3\\-3\\0\\3\\0\end{bmatrix} = a\begin{bmatrix}1\\1\\0\\1\\0\end{bmatrix} + b\begin{bmatrix}1\\-1\\1\\0\\-2\end{bmatrix} + c\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix} \longleftrightarrow \begin{bmatrix}1&1&1\\1&-1&1\\0&1&1\\1&0&1\\0&-2&1\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}-3\\-3\\0\\3\\0\end{bmatrix}$$

The matrix $A = \begin{bmatrix}1&1&1\\1&-1&1\\0&1&1\\1&0&1\\0&-2&1\end{bmatrix}$ has $A^T = \begin{bmatrix}1&1&0&1&0\\1&-1&1&0&-2\\1&1&1&1&1\end{bmatrix}$

$$A^T\vec{z} = \begin{bmatrix}1&1&0&1&0\\1&-1&1&0&-2\\1&1&1&1&1\end{bmatrix}\begin{bmatrix}-3\\-3\\0\\3\\0\end{bmatrix} = \begin{bmatrix}-3\\0\\-3\end{bmatrix} \text{ and } A^TA = \begin{bmatrix}1&1&0&1&0\\1&-1&1&0&-2\\1&1&1&1&1\end{bmatrix}\begin{bmatrix}1&1&1\\1&-1&1\\0&1&1\\1&0&1\\0&-2&1\end{bmatrix} = \begin{bmatrix}3&0&3\\0&7&-1\\3&-1&5\end{bmatrix}$$

Solving $\left[\begin{array}{ccc|c}3&0&3&-3\\0&7&-1&0\\3&-1&5&-3\end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c}1&0&0&-1\\0&1&0&0\\0&0&1&0\end{array}\right]$ tells us that $\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}-1\\0\\0\end{bmatrix}$

Therefore the plane of best fit is $z = -1x + 0y + 0$, or $z = -x$

Exercise 3.6.5a

Find the parabola of best fit for the points $(-1,-3), (0,-3), (1,17)$ and $(2,-3)$

Answer

We want $y = ax^2 + bx + c$, so we set $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = a \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \\ x_4^2 \end{bmatrix} + b \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} -3 \\ -3 \\ 17 \\ -3 \end{bmatrix} = a \begin{bmatrix} (-1)^2 \\ 0^2 \\ 1^2 \\ 2^2 \end{bmatrix} + b \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \\ 17 \\ -3 \end{bmatrix}$$

The matrix $A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$ has $A^T = \begin{bmatrix} 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

$$A^T\vec{y} = \begin{bmatrix} 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -3 \\ 17 \\ -3 \end{bmatrix} = \begin{bmatrix} 2 \\ 14 \\ 8 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 18 & 8 & 6 \\ 8 & 6 & 2 \\ 6 & 2 & 4 \end{bmatrix}$$

Solving $\left[\begin{array}{ccc|c} 18 & 8 & 6 & 2 \\ 8 & 6 & 2 & 14 \\ 6 & 2 & 4 & 8 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -5 \\ 0 & 1 & 0 & 7 \\ 0 & 0 & 1 & 6 \end{array}\right]$ tells us that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} -5 \\ 7 \\ 6 \end{bmatrix}$

Therefore the parabola of best fit is $y = -5x^2 + 7x + 6$

Exercise 3.6.5b

Find the parabola of best fit for the points $(-1, 16), (-1, 3), (0, 3), (1, 3)$ and $(2, 29)$

Answer

We want $y = ax^2 + bx + c$, so we set $\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = a \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \\ x_4^2 \\ x_5^2 \end{bmatrix} + b \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 16 \\ 3 \\ 3 \\ 3 \\ 29 \end{bmatrix} = a \begin{bmatrix} (-1)^2 \\ (-1)^2 \\ 0^2 \\ 1^2 \\ 2^2 \end{bmatrix} + b \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 16 \\ 3 \\ 3 \\ 3 \\ 29 \end{bmatrix}$$

The matrix $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$ has $A^T = \begin{bmatrix} 1 & 1 & 0 & 1 & 4 \\ -1 & -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$

$$A^T\vec{y} = \begin{bmatrix} 1 & 1 & 0 & 1 & 4 \\ -1 & -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 16 \\ 3 \\ 3 \\ 3 \\ 29 \end{bmatrix} = \begin{bmatrix} 138 \\ 42 \\ 54 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 1 & 1 & 0 & 1 & 4 \\ -1 & -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 19 & 7 & 7 \\ 7 & 7 & 1 \\ 7 & 1 & 5 \end{bmatrix}$$

Solving $\left[\begin{array}{ccc|c} 19 & 7 & 7 & 138 \\ 7 & 7 & 1 & 42 \\ 7 & 1 & 5 & 54 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \end{array}\right]$ tells us that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 8 \\ -2 \\ 0 \end{bmatrix}$

Therefore the parabola of best fit is $y = 8x^2 - 2x + 0$

Exercise 3.6.6a

Find the circle of best fit for the points $(1,1),(1,-5),(3,-1)$ and $(-1,1)$

Answer

We want $x^2+y^2=ax+by+c$, so we set $\begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \\ x_4^2 \end{bmatrix} + \begin{bmatrix} y_1^2 \\ y_2^2 \\ y_3^2 \\ y_4^2 \end{bmatrix} = a\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + b\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} + c\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 1^2+1^2 \\ 1^2+(-5)^2 \\ 3^2+(-1)^2 \\ (-1)^2+1^2 \end{bmatrix} = a\begin{bmatrix} 1 \\ 1 \\ 3 \\ -1 \end{bmatrix} + b\begin{bmatrix} 1 \\ -5 \\ -1 \\ 1 \end{bmatrix} + c\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 1 & 1 & 1 \\ 1 & -5 & 1 \\ 3 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2 \\ 26 \\ 10 \\ 2 \end{bmatrix}$$

The matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -5 & 1 \\ 3 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}$ has $A^T = \begin{bmatrix} 1 & 1 & 3 & -1 \\ 1 & -5 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

$$A^T\vec{z} = \begin{bmatrix} 1 & 1 & 3 & -1 \\ 1 & -5 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 26 \\ 10 \\ 2 \end{bmatrix} = \begin{bmatrix} 56 \\ -136 \\ 40 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 1 & 1 & 3 & -1 \\ 1 & -5 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 1 & -5 & 1 \\ 3 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 12 & -8 & 4 \\ -8 & 28 & -4 \\ 4 & -4 & 4 \end{bmatrix}$$

Solving $\left[\begin{array}{ccc|c} 12 & -8 & 4 & 56 \\ -8 & 28 & -4 & -136 \\ 4 & -4 & 4 & 40 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 6 \end{array}\right]$ tells us that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ -4 \\ 6 \end{bmatrix}$

Therefore the circle of best fit is $x^2+y^2=0x-4y+6$, which we can rewrite as $(x-0)^2+(y+2)^2=10$

Exercise 3.6.6b

Find the circle of best fit for the points $(5,1),(3,1),(3,-3)$ and $(1,-3)$

Answer

We want $x^2+y^2=ax+by+c$, so we set $\begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \\ x_4^2 \end{bmatrix} + \begin{bmatrix} y_1^2 \\ y_2^2 \\ y_3^2 \\ y_4^2 \end{bmatrix} = a\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + b\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} + c\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

$$\begin{bmatrix} 5^2+1^2 \\ 3^2+1^2 \\ 3^2+(-3)^2 \\ 1^2+(-3)^2 \end{bmatrix} = a\begin{bmatrix} 5 \\ 3 \\ 3 \\ 1 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ -3 \\ -3 \end{bmatrix} + c\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 5 & 1 & 1 \\ 3 & 1 & 1 \\ 3 & -3 & 1 \\ 1 & -3 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 26 \\ 10 \\ 18 \\ 10 \end{bmatrix}$$

The matrix $A = \begin{bmatrix} 5 & 1 & 1 \\ 3 & 1 & 1 \\ 3 & -3 & 1 \\ 1 & -3 & 1 \end{bmatrix}$ has $A^T = \begin{bmatrix} 5 & 3 & 3 & 1 \\ 1 & 1 & -3 & -3 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

$$A^T\vec{z} = \begin{bmatrix} 5 & 3 & 3 & 1 \\ 1 & 1 & -3 & -3 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 26 \\ 10 \\ 18 \\ 10 \end{bmatrix} = \begin{bmatrix} 224 \\ -48 \\ 64 \end{bmatrix} \text{ and } A^TA = \begin{bmatrix} 5 & 3 & 3 & 1 \\ 1 & 1 & -3 & -3 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 5 & 1 & 1 \\ 3 & 1 & 1 \\ 3 & -3 & 1 \\ 1 & -3 & 1 \end{bmatrix} = \begin{bmatrix} 44 & -4 & 12 \\ -4 & 20 & -4 \\ 12 & -4 & 4 \end{bmatrix}$$

Solving $\left[\begin{array}{ccc|c} 44 & -4 & 12 & 224 \\ -4 & 20 & -4 & -48 \\ 12 & -4 & 4 & 64 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -4 \end{array}\right]$ tells us that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 6 \\ -2 \\ -4 \end{bmatrix}$

Therefore the circle of best fit is $x^2+y^2=6x-2y-4$, which we can rewrite as $(x-3)^2+(y+1)^2=6$

3.6: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

3.6: Orthogonal Sets

Learning Objectives

1. Understand which is the best method to use to compute an orthogonal projection in a given situation.
2. *Recipes:* an orthonormal set from an orthogonal set, Projection Formula, B-coordinates when B is an orthogonal set, Gram–Schmidt process.
3. *Vocabulary words:* ***orthogonal set***, ***orthonormal set***.

In this section, we give a formula for orthogonal projection that is considerably simpler than the previous one we used, in that it does not require row reduction or matrix inversion. However, this formula, called the Projection Formula, only works in the presence of an *orthogonal* basis. We will also present the Gram–Schmidt process for turning an arbitrary basis into an orthogonal one.

This page titled 3.6: Orthogonal Sets is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **6.4: Orthogonal Sets** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.7: Orthogonal Sets and Projection

Computations involving projections tend to be much easier in the presence of an *orthogonal* set of vectors.

Definition: Orthogonal and Orthonormal

A set of *nonzero* vectors $\{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_m\}$ is called ***orthogonal*** if $\vec{u}_i \cdot \vec{u}_j = 0$ whenever $i \neq j$.

It is ***orthonormal*** if it is orthogonal, and in addition $\vec{u}_i \cdot \vec{u}_i = 1$ for all $i = 1,\ 2, \cdots, m$.

In other words, a set of vectors is orthogonal if different vectors in the set are perpendicular to each other. An orthonormal set is an orthogonal set of unit vectors.

Illustration 3.7.1

The standard coordinate vectors in $\mathbb{R}^n$ always form an orthonormal set. For instance, in $\mathbb{R}^3$ we check that

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 0 \quad \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0 \quad \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$

Since $\vec{e}_i \cdot \vec{e}_i = 1$ for all $i = 1,\ 2,\ 3$, this shows that $\{\vec{e}_1,\ \vec{e}_2,\ \vec{e}_3\}$ is orthonormal.

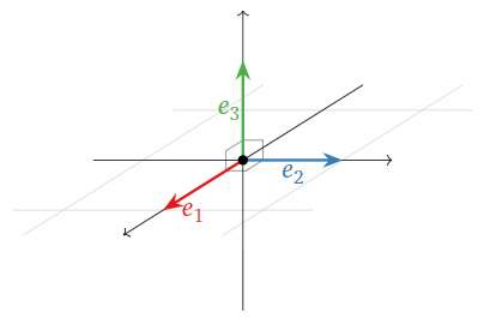


Figure 3.7.1

Example 3.7.1

Is this set orthogonal? Is it orthonormal?

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$$

Solution

We check that

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = 0 \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 0 \quad \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = 0$$

Therefore, $\boxed{\mathcal{B} \text{ is orthogonal}}$

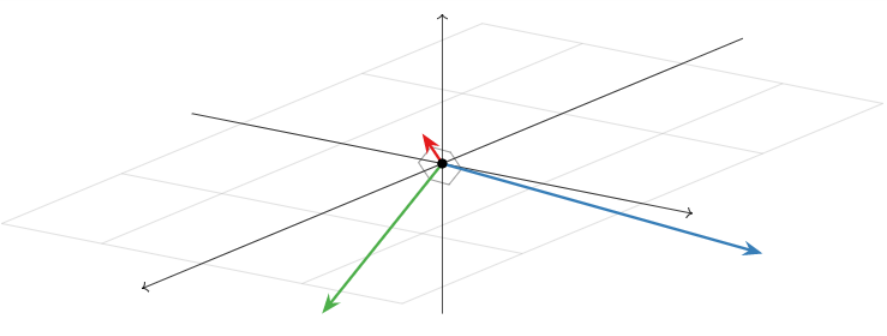


Figure 3.7.2

The set $\boxed{\mathcal{B} \text{ is not orthonormal}}$, for instance we see that:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 3 \neq 1$$

However, we can make it orthonormal by replacing each vector by the unit vector in the direction of each vector:

$$\left\{ \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \right\}$$

is orthonormal.

We saw in the previous example that it is easy to produce an orthonormal set of vectors from an orthogonal one by replacing each vector with the unit vector in the same direction.

Recipe: An Orthonormal Set from an Orthogonal Set

If $\{\vec{v}_1, \vec{v}_2, \cdots, \vec{v}_m\}$ is an orthogonal set of vectors, then $\left\{ \frac{\vec{v}_1}{||\vec{v}_1||}, \frac{\vec{v}_2}{||\vec{v}_2||}, \cdots, \frac{\vec{v}_m}{||\vec{v}_m||} \right\}$ is an orthonormal set.

Example 3.7.2: Non-Example

Is this set orthogonal?

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \right\}$$

Solution

This set is $\boxed{\text{not orthogonal}}$ because

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} = 1 - 1 - 1 = -1 \neq 0$$

We will see how to produce an orthogonal set from $\mathcal{B}$ in this section.

A nice property enjoyed by orthogonal sets is that they are automatically linearly independent.

Fact 3.7.1

An orthogonal set is linearly independent. Therefore, it is a basis for its span.

Proof

Suppose that $\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_m\}$ is orthogonal. We need to show that the equation

$$c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_m\vec{u}_m = \vec{0}$$

has only the trivial solution $c_1 = c_2 = \cdots = c_m = 0$. Taking the dot product of both sides of this equation with $\vec{u}_1$ gives

$$\begin{aligned} 0 = \vec{u}_1 \cdot \vec{0} &= \vec{u}_1 \cdot (c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_m\vec{u}_m) \\ &= c_1(\vec{u}_1 \cdot \vec{u}_1) + c_2(\vec{u}_1 \cdot \vec{u}_2) + \cdots + c_m(\vec{u}_1 \cdot \vec{u}_m) \\ &= c_1(\vec{u}_1 \cdot \vec{u}_1) \end{aligned}$$

because $\vec{u}_1 \cdot \vec{u}_i = 0$ for $i > 1$. Since $\vec{u}_1 \neq \vec{0}$ we have $\vec{u}_1 \cdot \vec{u}_1 \neq 0$, so $c_1 = 0$. Similarly, taking the dot product with $\vec{u}_i$ shows that each $c_i = 0$, as desired.

One advantage of working with orthogonal sets is that it gives a simple formula for the orthogonal projection of a vector.

Theorem 3.7.1: Projection Formula

Let W be a subspace of $\mathbb{R}^n$, and let $\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_m\}$ be an orthogonal basis for W. Then for any vector $\vec{x}$ in $\mathbb{R}^n$, the orthogonal projection of $\vec{x}$ onto W is given by the formula

$$\vec{x}_W = \frac{\vec{x} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 + \frac{\vec{x} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2 + \cdots + \frac{\vec{x} \cdot \vec{u}_m}{\vec{u}_m \cdot \vec{u}_m}\vec{u}_m$$

Proof

Let

$$\vec{y} = \frac{\vec{x} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 + \frac{\vec{x} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2 + \cdots + \frac{\vec{x} \cdot \vec{u}_m}{\vec{u}_m \cdot \vec{u}_m}\vec{u}_m$$

This vector is contained in W because it is a linear combination of $\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_m$. Hence we just need to show that $\vec{x} - \vec{y}$ is in $W^\perp$, i.e., that $\vec{u}_i \cdot (\vec{x} - \vec{y}) = 0$ for each $i = 1,\ 2, \cdots, m$. For $\vec{u}_1$, we have

$$\begin{aligned} \vec{u}_1 \cdot (\vec{x} - \vec{y}) &= \vec{u}_1 \cdot \left(\vec{x} - \frac{\vec{x} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 - \frac{\vec{x} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2 - \cdots - \frac{\vec{x} \cdot \vec{u}_m}{\vec{u}_m \cdot \vec{u}_m}\vec{u}_m\right) \\ &= \vec{u}_1 \cdot \vec{x} - \frac{\vec{x} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}(\vec{u}_1 \cdot \vec{u}_1) - 0 - \cdots - 0 \\ &= 0 \end{aligned}$$

A similar calculation shows that $\vec{u}_i \cdot (\vec{x} - \vec{y}) = 0$ for each i, so $\vec{x} - \vec{y}$ is in $W^\perp$, as desired.

If $\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_m\}$ is an ortho*normal* basis for W, then the denominators $\vec{u}_i \cdot \vec{u}_i = 1$ go away, so the projection formula becomes even simpler:

$$\vec{x}_W = (\vec{x} \cdot \vec{u}_1)\vec{u}_1 + (\vec{x} \cdot \vec{u}_2)\vec{u}_2 + \cdots + (\vec{x} \cdot \vec{u}_m)\vec{u}_m$$

Illustration 3.7.2

Suppose that $L = \text{Span}\{\vec{u}\}$ is a line. The set $\{\vec{u}\}$ is an orthogonal basis for L, so the Projection Formula says that for any vector $\vec{x}$, we have

$$\vec{x}_L = \frac{\vec{x} \cdot \vec{u}}{\vec{u} \cdot \vec{u}}\vec{u}$$

which is exactly the formula we got when we did this problem in the previous section!

Suppose that $\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_m\}$ is an orthogonal basis for a subspace W, and let $L_i = \text{Span}\{\vec{u}_i\}$ for each $i = 1,\ 2, \cdots, m$. Then we see that for any vector $\vec{x}$, we have

$$\begin{aligned}\vec{x}_W &= \frac{\vec{x}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\vec{u}_1 + \frac{\vec{x}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2}\vec{u}_2 + \cdots + \frac{\vec{x}\cdot\vec{u}_m}{\vec{u}_m\cdot\vec{u}_m}\vec{u}_m \\ &= \vec{x}_{L_1} + \vec{x}_{L_2} + \cdots + \vec{x}_{L_m}\end{aligned}$$

In other words, for an *orthogonal* basis, the projection of $\vec{x}$ onto W is the sum of the projections onto the lines spanned by the basis vectors. In this sense, projection onto a line is the most important example of an orthogonal projection.

Illustration 3.7.3: Projection onto the xy-Plane

We use the projection formula, Theorem 3.7.1 to compute the orthogonal projection of a vector onto the xy-plane in $\mathbb{R}^3$.

A basis for the xy-plane is given by the two standard coordinate vectors

$$\vec{e}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix} \quad \vec{e}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$$

The set $\{\vec{e}_1,\ \vec{e}_2\}$ is orthonormal, so for any vector $\vec{x} = (x_1,\ x_2,\ x_3)$, we have

$$\vec{x}_W = (\vec{x}\cdot\vec{e}_1)\vec{e}_1 + (\vec{x}\cdot\vec{e}_2)\vec{e}_2 = x_1\vec{e}_1 + x_2\vec{e}_2 = \begin{bmatrix}x_1\\x_2\\0\end{bmatrix}$$

Example 3.7.3: Projection onto a 3-Space in $\mathbb{R}^4$

Let

$$W = \text{Span}\left\{\begin{bmatrix}1\\0\\-1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\1\\1\end{bmatrix}\right\} \quad \vec{x} = \begin{bmatrix}0\\1\\3\\4\end{bmatrix}$$

Compute $\vec{x}_W$, and find the distance from $\vec{x}$ to W.

Solution

The vectors

$$\vec{u}_1 = \begin{bmatrix}1\\0\\-1\\0\end{bmatrix} \quad \vec{u}_2 = \begin{bmatrix}0\\1\\0\\-1\end{bmatrix} \quad \vec{u}_3 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

are orthogonal, so we can use the Projection Formula:

$$\vec{x}_W = \frac{\vec{x}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1}\vec{u}_1 + \frac{\vec{x}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2}\vec{u}_2 + \frac{\vec{x}\cdot\vec{u}_3}{\vec{u}_3\cdot\vec{u}_3}\vec{u}_3 = \frac{-3}{2}\begin{bmatrix}1\\0\\-1\\0\end{bmatrix} + \frac{-3}{2}\begin{bmatrix}0\\1\\0\\-1\end{bmatrix} + \frac{8}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \boxed{\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{7}{2}\\\frac{7}{2}\end{bmatrix}}$$

$$\vec{x}_{W^\perp} = \vec{x} - \vec{x}_W = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

The distance from $\vec{x}$ to W is

$$\|\vec{x}_{W^\perp}\| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2} = \boxed{1}$$

This gives us a way of expressing $\vec{x}$ as a linear combination of the basis vectors in $\mathcal{B}$: we have computed the $\mathcal{B}$-coordinates of $\vec{x}$ without row reducing!

Recipe: $\mathcal{B}$-Coordinates when $\mathcal{B}$ is an Orthogonal Set

Let W be a subspace of $\mathbb{R}^n$ with orthogonal basis $\mathcal{B} = \{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_m\}$ and let $\vec{x}$ be a vector in W. Then

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} \frac{\vec{x} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \\ \frac{\vec{x} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \\ \vdots \\ \frac{\vec{x} \cdot \vec{u}_m}{\vec{u}_m \cdot \vec{u}_m} \end{bmatrix}$$

As with orthogonal projections, if $\{\vec{u}_1, \vec{u}_2, \cdots, \vec{u}_m\}$ is an ortho*normal* basis of W, then the formula is even simpler:

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} \vec{x} \cdot \vec{u}_1 \\ \vec{x} \cdot \vec{u}_2 \\ \vdots \\ \vec{x} \cdot \vec{u}_m \end{bmatrix}$$

Example 3.7.4: Computing Coordinates with Respect to an Orthogonal Basis

Find the $\mathcal{B}$-coordinates of $\vec{x}$, where

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -4 \\ 2 \end{bmatrix} \right\} \quad \vec{x} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Solution

Since

$$\vec{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \vec{u}_2 = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

form an *orthogonal* basis of $\mathbb{R}^2$, we have

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} \frac{\vec{x} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} \\ \frac{\vec{x} \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2} \end{bmatrix} = \begin{bmatrix} \frac{3 \cdot 2}{1^2 + 2^2} \\ \frac{3 \cdot 2}{(-4)^2 + 2^2} \end{bmatrix} = \boxed{\begin{bmatrix} \frac{6}{5} \\ \frac{3}{10} \end{bmatrix}}$$

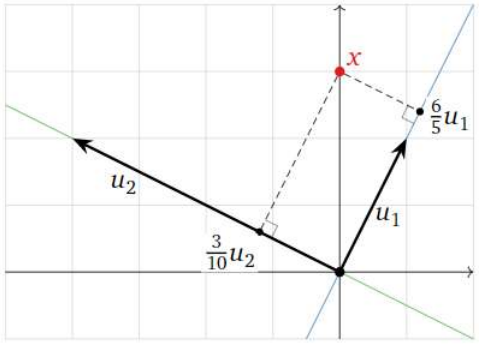


Figure 3.7.3

The following example shows that the Projection Formula does in fact require an orthogonal basis.

Illustration 3.7.4: Non-Example: A Non-Orthogonal Basis

Consider the basis $\mathcal{B} = \{\vec{v}_1, \vec{v}_2\}$ of $\mathbb{R}^2$, where

$$\vec{v}_1 = \begin{bmatrix} 2 \\ -\frac{1}{2} \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

This is not orthogonal because $\vec{v}_1 \cdot \vec{v}_2 = 2 - 1 = 1 \neq 0$. Let $\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Let us try to compute $\vec{x} = \vec{x}_{\mathbb{R}^2}$ using the Projection Formula with respect to the basis $\mathcal{B}$:

$$\vec{x}_{\mathbb{R}^2} = \frac{\vec{x} \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 + \frac{\vec{x} \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 = \frac{\frac{3}{2}}{\frac{17}{4}} \begin{bmatrix} 2 \\ -\frac{1}{2} \end{bmatrix} + \frac{3}{5} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{111}{85} \\ \frac{87}{85} \end{bmatrix} \neq \vec{x}$$

Since $\vec{x} = \vec{x}_{\mathbb{R}^2}$, we see that the Projection Formula does not compute the orthogonal projection in this case. Geometrically, the projections of $\vec{x}$ onto the lines spanned by $\vec{v}_1$ and $\vec{v}_2$ do not sum to $\vec{x}$, as we can see from the picture.

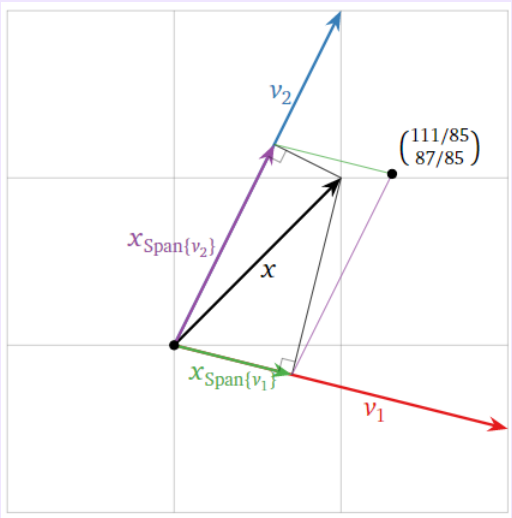


Figure 3.7.4

Note 3.7.1

You need an orthogonal basis to use the Projection Formula.

3.7: Orthogonal Sets and Projection is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.4: Orthogonal Sets** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.7: Gram-Schmidt Process

We saw in the previous subsection that orthogonal projections and $\mathcal{B}$-coordinates are much easier to compute in the presence of an *orthogonal* basis for a subspace. In this subsection, we give a method, called the *Gram–Schmidt Process*, for computing an orthogonal basis of a subspace.

Theorem 3.7.2: The Gram-Schmidt Process

Let $\vec{v}_1,\ \vec{v}_2, \cdots, \vec{v}_m$ be a basis for a subspace W of $\mathbb{R}^n$. Define:

1. $\vec{u}_1 = \vec{v}_1$
2. $\vec{u}_2 = (\vec{v}_2)_{\text{Span}\{\vec{u}_1\}^\perp} = \vec{v}_2 - \dfrac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1$
3. $\vec{u}_3 = (\vec{v}_3)_{\text{Span}\{\vec{u}_1,\vec{u}_2\}^\perp} = \vec{v}_3 - \dfrac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 - \dfrac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2$

$\vdots$

M. $\vec{u}_m = (\vec{v}_m)_{\text{Span}\{\vec{u}_1,\vec{u}_2,\cdots,\vec{u}_{m-1}\}^\perp} = \vec{v}_m - \sum_{i=1}^{m-1} \dfrac{\vec{v}_m \cdot \vec{u}_i}{\vec{u}_i \cdot \vec{u}_i}\vec{u}_i$

Then $\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_m\}$ is an orthogonal basis for the same subspace W.

Proof

First we claim that each $\vec{u}_i$ is in W, and in fact that $\vec{u}_i$ is in $\text{Span}\{\vec{v}_1,\ \vec{v}_2, \cdots, \vec{v}_i\}$. Clearly $\vec{u}_1 = \vec{v}_1$ is in $\text{Span}\{\vec{v}_1\}$. Then $\vec{u}_2$ is a linear combination of $\vec{u}_1$ and $\vec{v}_2$, which are both in $\text{Span}\{\vec{v}_1,\ \vec{v}_2\}$, so $\vec{u}_2$ is in $\text{Span}\{\vec{v}_1,\ \vec{v}_2\}$ as well. Similarly, $\vec{u}_3$ is a linear combination of $\vec{u}_1,\ \vec{u}_2$, and $\vec{v}_3$, which are all in $\text{Span}\{\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3\}$, so $\vec{u}_3$ is in $\text{Span}\{\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3\}$. Continuing in this way, we see that each $\vec{u}_i$ is in $\text{Span}\{\vec{v}_1,\ \vec{v}_2, \cdots, \vec{v}_i\}$.

Now we claim that $\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_m\}$ is an orthogonal set. Let $1 \le i < j \le m$. Then $\vec{u}_j = (\vec{v}_j)_{\text{Span}\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_{j-1}\}^\perp}$, so by definition $\vec{u}_j$ is orthogonal to every vector in $\text{Span}\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_{j-1}\}$. In particular, $\vec{u}_j$ is orthogonal to $\vec{u}_i$.

We still have to prove that each $\vec{u}_i$ is nonzero. Clearly $\vec{u}_1 = \vec{v}_1 \neq \vec{0}$. Suppose that $\vec{u}_i = \vec{0}$. Then $(\vec{v}_i)_{\text{Span}\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_{i-1}\}^\perp} = \vec{0}$, which means that $\vec{v}_i$ is in $\text{Span}\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_{i-1}\}$. But each $\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_{i-1}$ is in $\text{Span}\{\vec{v}_1,\ \vec{v}_2, \cdots, \vec{v}_{i-1}\}$ by the first paragraph, so $\vec{v}_i$ is in $\text{Span}\{\vec{v}_1,\ \vec{v}_2, \cdots, \vec{v}_{i-1}\}$. This contradicts the increasing span criterion since we started with a basis (linearly independent set); therefore, $\vec{u}_i$ must be nonzero.

The previous two paragraphs justify the use of the projection formula, Theorem 3.7.1, in the equalities

$$(\vec{v}_i)_{\text{Span}\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_{i-1}\}^\perp} = \vec{v}_i - (\vec{v}_i)_{\text{Span}\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_{i-1}\}} = \vec{v}_i - \sum_{j=1}^{i-1} \frac{\vec{v}_i \cdot \vec{u}_j}{\vec{u}_j \cdot \vec{u}_j}\vec{u}_j$$

in the statement of the theorem.

Since $\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_m\}$ is an orthogonal set, it is linearly independent. Thus it is a set of m linearly independent vectors in W, so it is a basis for W. Similarly, for every i, we saw that the set $\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_i\}$ is contained in the i-dimensional subspace $\text{Span}\{\vec{v}_1,\ \vec{v}_2, \cdots, \vec{v}_i\}$, so $\{\vec{u}_1,\ \vec{u}_2, \cdots, \vec{u}_i\}$ is an orthogonal basis for $\text{Span}\{\vec{v}_1,\ \vec{v}_2, \cdots, \vec{v}_i\}$

Illustration 3.7.5: Two Vectors

In order to find an orthogonal basis $\{\vec{u}_1,\ \vec{u}_2\}$ for $W = \text{Span}\{\vec{v}_1,\ \vec{v}_2\}$, where

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

We will use Gram–Schmidt: first take $\vec{u}_1 = \vec{v}_1$, then

$$\vec{u}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Then $\{\vec{u}_1,\ \vec{u}_2\}$ is an orthogonal basis for W: indeed, it is clear that $\vec{u}_1 \cdot \vec{u}_2 = 0$.

Geometrically, we are simply replacing $\vec{v}_2$ with the part of $\vec{v}_2$ that is perpendicular to line $L_1 = \text{Span}\{\vec{v}_1\}$:

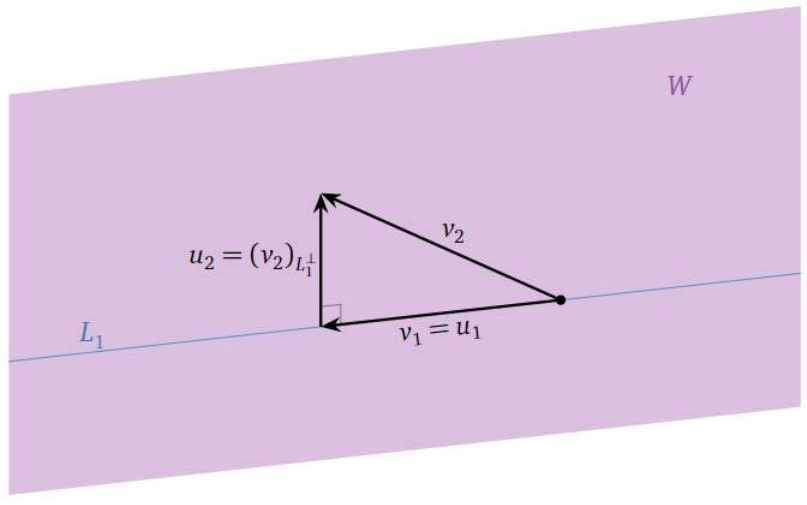


Figure 3.7.5

Example 3.7.5: Three Vectors in $\mathbb{R}^4$

Find an orthonormal basis $\{\vec{w}_1,\ \vec{w}_2,\ \vec{w}_3\}$ for $W = \text{Span}\{\vec{v}_1,\ \vec{v}_2,\ \vec{v}_3\}$, where

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \vec{v}_2 = \begin{bmatrix} -1 \\ 4 \\ 4 \\ -1 \end{bmatrix} \quad \vec{v}_3 = \begin{bmatrix} 4 \\ -2 \\ -2 \\ 0 \end{bmatrix}$$

Solution

We run Gram–Schmidt:

$$\vec{u}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\vec{u}_2 = \vec{v}_2 - \frac{\vec{v}_2 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 = \begin{bmatrix} -1 \\ 4 \\ 4 \\ -1 \end{bmatrix} - \frac{6}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{5}{2} \\ \frac{5}{2} \\ \frac{5}{2} \\ -\frac{5}{2} \end{bmatrix}$$

$$\vec{u}_3 = \vec{v}_3 - \frac{\vec{v}_3 \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 - \frac{\vec{v}_3 \cdot \vec{u}_2}{\vec{u}_2 \cdot \vec{u}_2}\vec{u}_2$$

$$= \begin{bmatrix} 4 \\ -2 \\ -2 \\ 0 \end{bmatrix} - \frac{0}{24}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{-20}{25}\begin{bmatrix} -\frac{5}{2} \\ \frac{5}{2} \\ \frac{5}{2} \\ -\frac{5}{2} \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \\ -2 \end{bmatrix}$$

Then $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthogonal basis for W.

To make it orthonormal, we divide each $\vec{u}_i$ by its length $\|\vec{u}_i\|$:

$$\boxed{\vec{w}_1 = \frac{1}{2}\vec{u}_1 = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \quad \vec{w}_2 = \frac{1}{5}\vec{u}_2 = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \quad \vec{w}_3 = \frac{1}{2\sqrt{2}}\vec{u}_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}}$$

Note 3.7.2

You can use the Gram–Schmidt Process to produce an orthogonal basis from any spanning set: if some $\vec{u}_i = 0$, just throw away $\vec{u}_i$ and $\vec{v}_i$, and continue.

Sometimes one can use the properties of orthogonal sets to solve least-squares questions more easily.

Example 3.7.6: Best-fit linear function

Find the linear function $f(x, y)$ that best approximates the following data:

x	y	$f(x, y)$
1	0	0
0	1	1
−1	0	3
0	−1	4

What quantity is being minimized?

Solution

The general equation for a linear function in two variables is

$$f(x, y) = Bx + Cy + D$$

We want to solve the following system of equations in the unknowns B, C, D:

$$\begin{aligned} B(1) + C(0) + D &= 0 \\ B(0) + C(1) + D &= 1 \\ B(-1) + C(0) + D &= 3 \\ B(0) + C(-1) + D &= 4 \end{aligned}$$

In matrix form, we can write this as $A\vec{x} = \vec{b}$ for

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \qquad \vec{x} = \begin{bmatrix} B \\ C \\ D \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ 4 \end{bmatrix}$$

We observe that the columns $\vec{u}_1, \vec{u}_2, \vec{u}_3$ of A are *orthogonal*, so:

$$\hat{x} = \begin{bmatrix} \dfrac{\vec{b}\cdot\vec{u}_1}{\vec{u}_1\cdot\vec{u}_1} \\ \dfrac{\vec{b}\cdot\vec{u}_2}{\vec{u}_2\cdot\vec{u}_2} \\ \dfrac{\vec{b}\cdot\vec{u}_3}{\vec{u}_3\cdot\vec{u}_3} \end{bmatrix} = \begin{bmatrix} \dfrac{-3}{2} \\ \dfrac{-3}{2} \\ \dfrac{8}{4} \end{bmatrix} = \begin{bmatrix} -\dfrac{3}{2} \\ -\dfrac{3}{2} \\ 2 \end{bmatrix}$$

We find a least-squares solution by multiplying both sides by the transpose:

$$A^TA = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \qquad A^T\vec{b} = \begin{bmatrix} -3 \\ -3 \\ 8 \end{bmatrix}$$

The matrix A^TA is diagonal (do you see why that happened?), so it is easy to solve the equation $A^TA\vec{x} = A^T\vec{b}$:

$$\left[\begin{array}{ccc|c} 2 & 0 & 0 & -3 \\ 0 & 2 & 0 & -3 \\ 0 & 0 & 4 & 8 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -\dfrac{3}{2} \\ 0 & 1 & 0 & -\dfrac{3}{2} \\ 0 & 0 & 1 & 2 \end{array}\right] \implies \hat{x} = \begin{bmatrix} -\dfrac{3}{2} \\ -\dfrac{3}{2} \\ 2 \end{bmatrix}$$

Therefore, the best-fit linear equation is

$$\boxed{f(x,y) = -\frac{3}{2}x - \frac{3}{2}y + 2}$$

Here is a picture of the graph of $f(x,y)$:

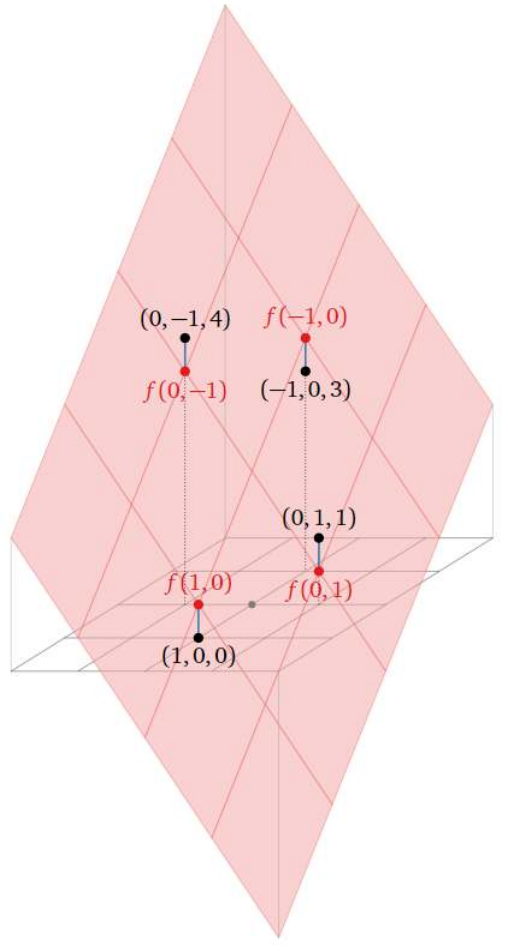


Figure 3.7.6

Now we consider what quantity is being minimized by the function $f(x, y)$. The least-squares solution $\hat{x}$ minimizes the sum of the squares of the entries of the vector $\vec{b} - A\hat{x}$.

$$A\hat{x} = \begin{bmatrix} -\frac{3}{2}(1) & - & \frac{3}{2}(0) & + & 2 \\ -\frac{3}{2}(0) & - & \frac{3}{2}(1) & + & 2 \\ -\frac{3}{2}(-1) & - & \frac{3}{2}(0) & + & 2 \\ -\frac{3}{2}(0) & - & \frac{3}{2}(-1) & + & 2 \end{bmatrix} = \begin{bmatrix} f(1,0) \\ f(0,1) \\ f(-1,0) \\ f(0,-1) \end{bmatrix}$$

In other words, $A\hat{x}$ is the vector whose entries are the values of f evaluated on the points (x, y) we specified in our data table, and $\vec{b}$ is the vector whose entries are the desired values of f evaluated at those points. The difference $\vec{b} - A\hat{x}$ is the vertical distance of the graph from the data points, as indicated in the above picture. The best-fit linear function minimizes the sum of these vertical distances.

3.7: Gram-Schmidt Process is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **6.4: Orthogonal Sets** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **6.5: The Method of Least Squares** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

3.7: Practice Set

Exercise 3.7.1

Determine if each set below is orthogonal, orthonormal, or neither.

(a) $S_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\}$

(b) $S_2 = \left\{ \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix} \right\}$

(c) $S_3 = \left\{ \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 16 \\ -7 \\ -5 \end{bmatrix} \right\}$

(d) $S_4 = \left\{ \begin{bmatrix} \sqrt{2} \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} \sqrt{2} \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ \sqrt{2} \\ \sqrt{2} \end{bmatrix} \right\}$

(e) $S_5 = \left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

(f) $S_6 = \left\{ \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \right\}$

Answer

(a) Neither

(b) Orthogonal

(c) Orthogonal

(d) Orthogonal

(e) Neither

(f) Orthonormal

Exercise 3.7.2a

Find an orthonormal basis for the subspace $W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = 0$ this set is already orthogonal, so we divide each vector by its length.

$$\|\vec{v}_1\| = \sqrt{2^2 + 4^2} = 2\sqrt{5} \longrightarrow \vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$$

$$\|\vec{v}_2\| = \sqrt{2^2 + (-1)^2} = \sqrt{5} \longrightarrow \vec{u}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{bmatrix}$$

Therefore an orthonormal basis for W is the set $\left\{\begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}, \begin{bmatrix} \frac{2}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} \end{bmatrix}\right\}$

? Exercise 3.7.2b

Find an orthonormal basis for the subspace $W = \text{Span}\left\{\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix}\right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 4 \\ -1 \\ 1 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = 0$ this set is already orthogonal, so we divide each vector by its length.

$$\|\vec{v}_1\| = \sqrt{1^2 + 2^2 + (-2)^2} = 3 \longrightarrow \vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{bmatrix}$$

$$\|\vec{v}_2\| = \sqrt{4^2 + (-1)^2 + 1^2} = 3\sqrt{2} \longrightarrow \vec{u}_2 = \frac{1}{\|\vec{v}_2\|}\vec{v}_2 = \begin{bmatrix} \frac{4}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \end{bmatrix}$$

Therefore an orthonormal basis for W is the set $\left\{\begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{bmatrix}, \begin{bmatrix} \frac{4}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \\ \frac{1}{3\sqrt{2}} \end{bmatrix}\right\}$

? Exercise 3.7.2c

Find an orthonormal basis for the subspace $W = \text{Span}\left\{\begin{bmatrix} 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 10 \end{bmatrix}\right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = 50$ this set is not orthogonal, so we orthogonalize it:

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

$$\vec{w}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 = \vec{v}_2 - 2\vec{w}_1 = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$

Now we divide each orthogonalized vector by its length.

$$\|\vec{w}_1\| = \sqrt{4^2 + 3^2} = 5 \longrightarrow \vec{u}_1 = \frac{1}{\|\vec{w}_1\|}\vec{w}_1 = \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$$

$$\|\vec{w}_2\| = \sqrt{(-3)^2 + 4^2} = 5 \longrightarrow \vec{u}_2 = \frac{1}{\|\vec{w}_2\|}\vec{w}_2 = \begin{bmatrix} -\frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$$

Therefore an orthonormal basis for W is the set $\left\{\begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix}, \begin{bmatrix} -\frac{3}{5} \\ \frac{4}{5} \end{bmatrix}\right\}$

Exercise 3.7.2d

Find an orthonormal basis for the subspace $W = \text{Span}\left\{\begin{bmatrix}1\\2\\0\\-1\end{bmatrix}, \begin{bmatrix}0\\1\\1\\2\end{bmatrix}, \begin{bmatrix}1\\8\\0\\5\end{bmatrix}\right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix}1\\2\\0\\-1\end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix}0\\1\\1\\2\end{bmatrix}$, and $\vec{v}_3 = \begin{bmatrix}1\\8\\0\\5\end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_3 = 12$ this set is not orthogonal, so we orthogonalize it:

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix}1\\2\\0\\-1\end{bmatrix}$$

$$\vec{w}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 = \vec{v}_2 - 0\vec{w}_1 = \begin{bmatrix}0\\1\\1\\2\end{bmatrix}$$

$$\vec{w}_3 = \vec{v}_3 - \left(\frac{\vec{v}_3 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 - \left(\frac{\vec{v}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 = \vec{v}_3 - 2\vec{w}_1 - 3\vec{w}_2 = \begin{bmatrix}-1\\1\\-3\\1\end{bmatrix}$$

Now we divide each orthogonalized vector by its length.

$$\|\vec{w}_1\| = \sqrt{1^2+2^2+0^2+(-1)^2} = \sqrt{6} \longrightarrow \vec{u}_1 = \frac{1}{\|\vec{w}_1\|}\vec{w}_1 = \begin{bmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\0\\-\frac{1}{\sqrt{6}}\end{bmatrix}$$

$$\|\vec{w}_2\| = \sqrt{0^2+1^2+1^2+2^2} = \sqrt{6} \longrightarrow \vec{u}_2 = \frac{1}{\|\vec{w}_2\|}\vec{w}_2 = \begin{bmatrix}0\\\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\end{bmatrix}$$

$$\|\vec{w}_3\| = \sqrt{(-1)^2+1^2+(-3)^2+1^2} = 2\sqrt{3} \longrightarrow \vec{u}_3 = \frac{1}{\|\vec{w}_3\|}\vec{w}_3 = \begin{bmatrix}-\frac{1}{2\sqrt{3}}\\\frac{1}{2\sqrt{3}}\\-\frac{3}{2\sqrt{3}}\\\frac{1}{2\sqrt{3}}\end{bmatrix}$$

Therefore an orthonormal basis for W is the set $\left\{\begin{bmatrix}\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\\0\\-\frac{1}{\sqrt{6}}\end{bmatrix}, \begin{bmatrix}0\\\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\\frac{2}{\sqrt{6}}\end{bmatrix}, \begin{bmatrix}-\frac{1}{2\sqrt{3}}\\\frac{1}{2\sqrt{3}}\\-\frac{3}{2\sqrt{3}}\\\frac{1}{2\sqrt{3}}\end{bmatrix}\right\}$

Exercise 3.7.2e

Find an orthonormal basis for the subspace $W = \text{Span}\left\{\begin{bmatrix}1\\2\\-2\\0\\4\end{bmatrix}, \begin{bmatrix}1\\4\\0\\0\\4\end{bmatrix}, \begin{bmatrix}-1\\7\\-7\\1\\12\end{bmatrix}, \begin{bmatrix}0\\1\\1\\19\\0\end{bmatrix}\right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix}1\\2\\-2\\0\\4\end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix}1\\4\\0\\0\\4\end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix}-1\\7\\-7\\1\\12\end{bmatrix}$, and $\vec{v}_4 = \begin{bmatrix}0\\1\\1\\19\\0\end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = 25$ this set is not orthogonal, so we orthogonalize it:

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix}1\\2\\-2\\0\\4\end{bmatrix}$$

$$\vec{w}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 = \vec{v}_2 - 1\vec{w}_1 = \begin{bmatrix}0\\2\\2\\0\\0\end{bmatrix}$$

$$\vec{w}_3 = \vec{v}_3 - \left(\frac{\vec{v}_3 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 - \left(\frac{\vec{v}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 = \vec{v}_3 - 3\vec{w}_1 - 0\vec{w}_2 = \begin{bmatrix}-4\\1\\-1\\1\\0\end{bmatrix}$$

$$\vec{w}_4 = \vec{v}_4 - \left(\frac{\vec{v}_4 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 - \left(\frac{\vec{v}_4 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 - \left(\frac{\vec{v}_4 \cdot \vec{w}_3}{\vec{w}_3 \cdot \vec{w}_3}\right)\vec{w}_3 = \vec{v}_4 - 0\vec{w}_1 - \tfrac{1}{2}\vec{w}_2 - 1\vec{w}_3 = \begin{bmatrix}4\\-1\\1\\18\\0\end{bmatrix}$$

Now we divide each orthogonalized vector by its length.

$$\|\vec{w}_1\| = \sqrt{1^2+2^2+(-2)^2+0^2+4^2} = 5 \longrightarrow \vec{u}_1 = \frac{1}{\|\vec{w}_1\|}\vec{w}_1 = \begin{bmatrix}\frac{1}{5}\\\frac{2}{5}\\-\frac{2}{5}\\0\\\frac{4}{5}\end{bmatrix}$$

$$\|\vec{w}_2\| = \sqrt{0^2+2^2+2^2+0^2+0^2} = 2\sqrt{2} \longrightarrow \vec{u}_2 = \frac{1}{\|\vec{w}_2\|}\vec{w}_2 = \begin{bmatrix}0\\\frac{1}{\sqrt{2}}\\\frac{1}{\sqrt{2}}\\0\\0\end{bmatrix}$$

$$\|\vec{w}_3\| = \sqrt{(-4)^2+1^2+(-1)^2+1^2+0^2} = \sqrt{19} \longrightarrow \vec{u}_3 = \frac{1}{\|\vec{w}_3\|}\vec{w}_3 = \begin{bmatrix}-\frac{4}{\sqrt{19}}\\\frac{1}{\sqrt{19}}\\-\frac{1}{\sqrt{19}}\\\frac{1}{\sqrt{19}}\\0\end{bmatrix}$$

$$\|\vec{w}_4\| = \sqrt{4^2 + (-1)^2 + 1^2 + 18^2 + 0^2} = 3\sqrt{38} \longrightarrow \vec{u}_4 = \frac{1}{\|\vec{w}_4\|}\vec{w}_4 = \begin{bmatrix} \frac{4}{3\sqrt{38}} \\ -\frac{1}{3\sqrt{38}} \\ \frac{1}{3\sqrt{38}} \\ \frac{6}{\sqrt{38}} \\ 0 \end{bmatrix}$$

Therefore an orthonormal basis for W is the set $\left\{ \begin{bmatrix} \frac{1}{5} \\ \frac{2}{5} \\ -\frac{2}{5} \\ 0 \\ \frac{4}{5} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{4}{\sqrt{19}} \\ \frac{1}{\sqrt{19}} \\ -\frac{1}{\sqrt{19}} \\ \frac{1}{\sqrt{19}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{4}{3\sqrt{38}} \\ -\frac{1}{3\sqrt{38}} \\ \frac{1}{3\sqrt{38}} \\ \frac{6}{\sqrt{38}} \\ 0 \end{bmatrix} \right\}$

Exercise $3.7.2f$

Find an orthonormal basis for the subspace $W = \text{Span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 3 \\ 0 \\ 0 \\ 3 \end{bmatrix}$, and $\vec{v}_4 = \begin{bmatrix} 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = -2$ this set is not orthogonal, so we orthogonalize it:

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 = \vec{v}_2 + 1\vec{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\vec{w}_3 = \vec{v}_3 - \left(\frac{\vec{v}_3 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 - \left(\frac{\vec{v}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 = \vec{v}_3 - 0\vec{w}_1 - 3\vec{w}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

(this means that $\vec{v}_3$ is in the span of $\vec{v}_1, \vec{v}_2$ and can be omitted from our basis.)

$$\vec{w}_4 = \vec{v}_4 - \left(\frac{\vec{v}_4 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 - \left(\frac{\vec{v}_4 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 = \vec{v}_4 - 0\vec{w}_1 - 1\vec{w}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

Now we divide each orthogonalized vector by its length.

$$\|\vec{w}_1\| = \sqrt{0^2 + 1^2 + 1^2 + 0^2} = \sqrt{2} \longrightarrow \vec{u}_1 = \frac{1}{\|\vec{w}_1\|}\vec{w}_1 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$\|\vec{w}_2\| = \sqrt{1^2+0^2+0^2+1^2} = \sqrt{2} \longrightarrow \vec{u}_2 = \frac{1}{\|\vec{w}_2\|}\vec{w}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\|\vec{w}_4\| = \sqrt{1^2+(-1)^2+1^2+(-1)^2} = 2 \longrightarrow \vec{u}_4 = \frac{1}{\|\vec{w}_4\|}\vec{w}_4 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$$

Therefore an orthonormal basis for W is the set $\left\{ \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix} \right\}$

Exercise 3.7.3a

Find the projection of $\vec{x} = \begin{bmatrix} -2 \\ 9 \\ -2 \end{bmatrix}$ onto $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 17 \end{bmatrix} \right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ 4 \\ 17 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = 0$ this set is orthogonal and we can use the projection formula:

$$\vec{x}_W = \left(\frac{\vec{x}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\right)\vec{v}_1 + \left(\frac{\vec{x}\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2}\right)\vec{v}_2 = 2\vec{v}_1 + 0\vec{v}_2 = \begin{bmatrix} 2 \\ 8 \\ -2 \end{bmatrix}$$

Exercise 3.7.3b

Find the projection of $\vec{x} = \begin{bmatrix} -2 \\ 3 \\ 10 \\ 16 \end{bmatrix}$ onto $W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 0 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 0 \\ 2 \end{bmatrix} \right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 1 \\ -3 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and $\vec{v}_3 = \begin{bmatrix} 3 \\ 0 \\ 0 \\ 2 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = \vec{v}_1 \cdot \vec{v}_3 = \vec{v}_2 \cdot \vec{v}_3 = 0$ this set is orthogonal and we can use the projection formula:

$$\vec{x}_W = \left(\frac{\vec{x}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\right)\vec{v}_1 + \left(\frac{\vec{x}\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2}\right)\vec{v}_2 + \left(\frac{\vec{x}\cdot\vec{v}_3}{\vec{v}_3\cdot\vec{v}_3}\right)\vec{v}_3 = -3\vec{v}_1 + 3\vec{v}_2 + 2\vec{v}_3 = \begin{bmatrix} 0 \\ 3 \\ -3 \\ 13 \end{bmatrix}$$

Exercise 3.7.3c

Find the projection of $\vec{x} = \begin{bmatrix} 8 \\ -2 \\ 23 \\ -25 \end{bmatrix}$ onto $W = \text{Span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 2 \\ 10 \end{bmatrix} \right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 5 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} -1 \\ 3 \\ 2 \\ 10 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = 56$ this set is not orthogonal, so we begin by orthogonalizing with Gram Schmidt:

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 5 \end{bmatrix}$$

$$\vec{w}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 = \vec{v}_2 - 2\vec{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Now we can use the projection formula:

$$\vec{x}_W = \left(\frac{\vec{x} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 + \left(\frac{\vec{x} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 = -4\vec{w}_1 + 3\vec{w}_2 = \begin{bmatrix} 7 \\ -1 \\ -4 \\ -20 \end{bmatrix}$$

? Exercise 3.7.3*d*

Find the projection of $\vec{x} = \begin{bmatrix} -9 \\ 9 \\ 12 \\ 12 \end{bmatrix}$ onto $W = \text{Span}\left\{\begin{bmatrix} 2 \\ 1 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 11 \\ -2 \\ -6 \\ 2 \end{bmatrix}, \begin{bmatrix} 15 \\ 4 \\ -1 \\ 1 \end{bmatrix}\right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 3 \\ -1 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 11 \\ -2 \\ -6 \\ 2 \end{bmatrix}$ and $\vec{v}_3 = \begin{bmatrix} 15 \\ 4 \\ -1 \\ 1 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_3 \neq 0$ this set is not orthogonal, so we begin by orthogonalizing with Gram Schmidt:

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 3 \\ -1 \end{bmatrix}$$

$$\vec{w}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 = \vec{v}_2 - 0\vec{w}_1 = \begin{bmatrix} 11 \\ -2 \\ -6 \\ 2 \end{bmatrix}$$

$$\vec{w}_3 = \vec{v}_3 - \left(\frac{\vec{v}_3 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 - \left(\frac{\vec{v}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 = \vec{v}_3 - 2\vec{w}_1 - 1\vec{w}_2 = \begin{bmatrix} 0 \\ 4 \\ -1 \\ 1 \end{bmatrix}$$

Now we can use the projection formula:

$$\vec{x}_W = \left(\frac{\vec{x} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 + \left(\frac{\vec{x} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 + \left(\frac{\vec{x} \cdot \vec{w}_3}{\vec{w}_3 \cdot \vec{w}_3}\right)\vec{w}_3 = 1\vec{w}_1 - 1\vec{w}_2 + 2\vec{w}_3 = \begin{bmatrix} -9 \\ 11 \\ 7 \\ -1 \end{bmatrix}$$

Exercise 3.7.3e

Find the projection of $\vec{x} = \begin{bmatrix} -11 \\ 3 \\ 10 \\ 1 \end{bmatrix}$ onto $W = \text{Span}\left\{ \begin{bmatrix} -1 \\ -1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ -3 \\ -2 \\ 7 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 1 \\ 3 \end{bmatrix} \right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 2 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} -2 \\ -3 \\ -2 \\ 7 \end{bmatrix}$ and $\vec{v}_3 = \begin{bmatrix} 2 \\ 3 \\ 1 \\ 3 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = 21$ this set is not orthogonal, so we begin by orthogonalizing with Gram Schmidt:

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 2 \end{bmatrix}$$

$$\vec{w}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1} \right) \vec{w}_1 = \vec{v}_2 - 3\vec{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\vec{w}_3 = \vec{v}_3 - \left(\frac{\vec{v}_3 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1} \right) \vec{w}_1 - \left(\frac{\vec{v}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2} \right) \vec{w}_2 = \vec{v}_3 - 0\vec{w}_1 - 2\vec{w}_2 = \begin{bmatrix} 0 \\ 3 \\ -1 \\ 1 \end{bmatrix}$$

Now we can use the projection formula:

$$\vec{x}_W = \left(\frac{\vec{x} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1} \right) \vec{w}_1 + \left(\frac{\vec{x} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2} \right) \vec{w}_2 + \left(\frac{\vec{x} \cdot \vec{w}_3}{\vec{w}_3 \cdot \vec{w}_3} \right) \vec{w}_3 = 0\vec{w}_1 + 0\vec{w}_2 + 0\vec{w}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

(The vector $\vec{x}$ is entirely contained in $W^{\perp}$ in this problem.)

Exercise 3.7.3f

Find the projection of $\vec{x} = \begin{bmatrix} 11 \\ 16 \\ 9 \\ -11 \\ -14 \end{bmatrix}$ onto $W = \text{Span}\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ -4 \\ 6 \\ 2 \end{bmatrix}, \begin{bmatrix} 8 \\ 11 \\ 3 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 9 \\ -8 \\ -9 \\ 0 \\ -9 \end{bmatrix} \right\}$

Answer

We denote $\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$, $\vec{v}_2 = \begin{bmatrix} 5 \\ 1 \\ -4 \\ 6 \\ 2 \end{bmatrix}$, $\vec{v}_3 = \begin{bmatrix} 8 \\ 11 \\ 3 \\ -2 \\ 1 \end{bmatrix}$ and $\vec{v}_4 = \begin{bmatrix} 9 \\ -8 \\ -9 \\ 0 \\ -9 \end{bmatrix}$

Since $\vec{v}_1 \cdot \vec{v}_2 = 15$ this set is not orthogonal, so we begin by orthogonalizing with Gram Schmidt:

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 = \vec{v}_2 - 5\vec{w}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

$$\vec{w}_3 = \vec{v}_3 - \left(\frac{\vec{v}_3 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 - \left(\frac{\vec{v}_3 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 = \vec{v}_3 - 1\vec{w}_1 - 2\vec{w}_2 = \begin{bmatrix} 7 \\ 9 \\ 2 \\ -5 \\ -3 \end{bmatrix}$$

$$\vec{w}_4 = \vec{v}_4 - \left(\frac{\vec{v}_4 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 - \left(\frac{\vec{v}_4 \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 - \left(\frac{\vec{v}_4 \cdot \vec{w}_3}{\vec{w}_3 \cdot \vec{w}_3}\right)\vec{w}_3 = \vec{v}_4 - 6\vec{w}_1 + 5\vec{w}_2 = \begin{bmatrix} 3 \\ -3 \\ 2 \\ -1 \\ 1 \end{bmatrix}$$

Now we can use the projection formula:

$$\vec{x}_W = \left(\frac{\vec{x} \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1}\right)\vec{w}_1 + \left(\frac{\vec{x} \cdot \vec{w}_2}{\vec{w}_2 \cdot \vec{w}_2}\right)\vec{w}_2 + \left(\frac{\vec{x} \cdot \vec{w}_3}{\vec{w}_3 \cdot \vec{w}_3}\right)\vec{w}_3 + \left(\frac{\vec{x} \cdot \vec{w}_4}{\vec{w}_4 \cdot \vec{w}_4}\right)\vec{w}_4 = -3\vec{w}_1 - 2\vec{w}_2 + 2\vec{w}_3 + 0\vec{w}_4 = \begin{bmatrix} 11 \\ 16 \\ 5 \\ -15 \\ -10 \end{bmatrix}$$

Exercise 3.7.4a

Find the line of best fit for the points $(1,-2), (2,8), (-4,-10), (4,18)$, and $(-3,-4)$

Answer

We set up $\vec{y} = m\vec{x} + b\vec{u}$ for $\vec{x} = \begin{bmatrix} 1 \\ 2 \\ -4 \\ 4 \\ -3 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} -2 \\ 8 \\ -10 \\ 18 \\ -4 \end{bmatrix}$

Since $\vec{x} \cdot \vec{u} = 0$ they form an orthogonal set, so we must have

$m = \dfrac{\vec{y} \cdot \vec{x}}{\vec{x} \cdot \vec{x}} = \dfrac{138}{46} = 3$ and $b = \dfrac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} = \dfrac{10}{5} = 2$

So the line of best fit is $y = 3x + 2$

Exercise 3.7.4b

Find the line of best fit for the points $(6,5), (-1,3), (5,4), (-3,-6), (-2,-9), (-4,-4)$, and $(-1,7)$

Answer

We set up $\vec{y} = m\vec{x} + b\vec{u}$ for $\vec{x} = \begin{bmatrix} 6 \\ -1 \\ 5 \\ -3 \\ -2 \\ -4 \\ -1 \end{bmatrix}$, $\vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} 5 \\ 3 \\ 4 \\ -6 \\ -9 \\ -4 \\ 7 \end{bmatrix}$

Since $\vec{x} \cdot \vec{u} = 0$ they form an orthogonal set, so we must have

$m = \dfrac{\vec{y} \cdot \vec{x}}{\vec{x} \cdot \vec{x}} = \dfrac{92}{92} = 1$ and $b = \dfrac{\vec{y} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} = \dfrac{0}{7} = 0$

So the line of best fit is $y = x$

Exercise 3.7.4c

Find the plane of best fit for the points $(2, 9, 40), (-1, 0, -58), (4, -7, 22)$, and $(-5, -2, 36)$

Answer

We set up $\vec{z} = a\vec{x} + b\vec{y} + c\vec{u}$ for $\vec{x} = \begin{bmatrix} 2 \\ -1 \\ 4 \\ -5 \end{bmatrix}, \vec{y} = \begin{bmatrix} 9 \\ 0 \\ -7 \\ -2 \end{bmatrix}, \vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $\vec{z} = \begin{bmatrix} 40 \\ -58 \\ 22 \\ 36 \end{bmatrix}$

Since $\vec{x} \cdot \vec{u} = \vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{u} = 0$ they form an orthogonal set, so we must have

$a = \dfrac{\vec{z} \cdot \vec{x}}{\vec{x} \cdot \vec{x}} = \dfrac{46}{46} = 1$ and $b = \dfrac{\vec{z} \cdot \vec{y}}{\vec{y} \cdot \vec{y}} = \dfrac{134}{134} = 1$ and $c = \dfrac{\vec{z} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} = \dfrac{40}{4} = 10$

So the plane of best fit is $z = x + y + 10$

Exercise 3.7.4d

Find the plane of best fit for the points $(4, 5, -15), (2, -7, 7), (-3, 2, -55), (-2, -1, -4), (-1, 2, -9)$, and $(0, -1, 4)$

Answer

We set up $\vec{z} = a\vec{x} + b\vec{y} + c\vec{u}$ for $\vec{x} = \begin{bmatrix} 4 \\ 2 \\ -3 \\ -2 \\ -1 \\ 0 \end{bmatrix}, \vec{y} = \begin{bmatrix} 5 \\ -7 \\ 2 \\ -1 \\ 2 \\ -1 \end{bmatrix}, \vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $\vec{z} = \begin{bmatrix} -15 \\ 7 \\ -55 \\ -4 \\ -9 \\ 4 \end{bmatrix}$

Since $\vec{x} \cdot \vec{u} = \vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{u} = 0$ they form an orthogonal set, so we must have

$a = \dfrac{\vec{z} \cdot \vec{x}}{\vec{x} \cdot \vec{x}} = \dfrac{136}{34} = 4$ and $b = \dfrac{\vec{z} \cdot \vec{y}}{\vec{y} \cdot \vec{y}} = \dfrac{-252}{84} = -3$ and $c = \dfrac{\vec{z} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} = \dfrac{-72}{6} = -12$

So the plane of best fit is $z = 4x - 3y - 12$

Exercise 3.7.4e

Find the circle of best fit for the points $(4, 4), (4, -4), (-4, 8)$, and $(-4, -8)$

Answer

We set up $\vec{s} = a\vec{x} + b\vec{y} + c\vec{u}$ for $\vec{x} = \begin{bmatrix} 4 \\ 4 \\ -4 \\ -4 \end{bmatrix}, \vec{y} = \begin{bmatrix} 4 \\ -4 \\ 8 \\ -8 \end{bmatrix}, \vec{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ and $\vec{s} = \begin{bmatrix} 4^2 + 4^2 \\ 4^2 + (-4)^2 \\ (-4)^2 + 8^2 \\ (-4)^2 + (-8)^2 \end{bmatrix} = \begin{bmatrix} 32 \\ 32 \\ 80 \\ 80 \end{bmatrix}$

Since $\vec{x} \cdot \vec{u} = \vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{u} = 0$ they form an orthogonal set, so we must have

$a = \dfrac{\vec{s} \cdot \vec{x}}{\vec{x} \cdot \vec{x}} = \dfrac{-384}{64} = -6$ and $b = \dfrac{\vec{s} \cdot \vec{y}}{\vec{y} \cdot \vec{y}} = \dfrac{0}{160} = 0$ and $c = \dfrac{\vec{s} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} = \dfrac{224}{4} = 56$

So the circle of best fit is $x^2 + y^2 = -6x + 0y + 56$ which can be written as $(x + 3)^2 + y^2 = 65$

Exercise 3.7.4f

Find the circle of best fit for the points $(0,-2),(0,0),(3,1)$, and $(-3,1)$

Answer

We set up $\vec{s}=a\vec{x}+b\vec{y}+c\vec{u}$ for $\vec{x}=\begin{bmatrix}0\\0\\3\\-3\end{bmatrix}, \vec{y}=\begin{bmatrix}-2\\0\\1\\1\end{bmatrix}, \vec{u}=\begin{bmatrix}1\\1\\1\\1\end{bmatrix}$ and $\vec{s}=\begin{bmatrix}0^2+(-2)^2\\0^2+0^2\\3^2+1^2\\(-3)^2+1^2\end{bmatrix}=\begin{bmatrix}4\\0\\10\\10\end{bmatrix}$

Since $\vec{x}\cdot\vec{u}=\vec{x}\cdot\vec{y}=\vec{y}\cdot\vec{u}=0$ they form an orthogonal set, so we must have

$a=\dfrac{\vec{s}\cdot\vec{x}}{\vec{x}\cdot\vec{x}}=\dfrac{0}{18}=0$ and $b=\dfrac{\vec{s}\cdot\vec{y}}{\vec{y}\cdot\vec{y}}=\dfrac{12}{6}=2$ and $c=\dfrac{\vec{s}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}=\dfrac{24}{4}=6$

So the circle of best fit is $x^2+y^2=0x+2y+6$ which can be written as $x^2+(y-1)^2=7$

Exercise 3.7.5

Determine if the following statements are Always True, Always False, or Sometimes True/Sometimes False.

(a) If a set of vectors is orthonormal, then it is linearly dependent.

(b) If the columns of an $n\times n$ matrix A form an orthogonal set, then A^TA is a diagonal matrix.

(c) If $\vec{x}$ is a unit vector and $\vec{x}\cdot\vec{v}=-\|\vec{v}\|$, then $\vec{v}=-\|\vec{v}\|\vec{x}$.

(d) If $\{\vec{u},\vec{v}\}$ is orthogonal and $\{\vec{u},\vec{w}\}$ is orthogonal, then $\{\vec{u},\vec{v},\vec{w}\}$ is orthogonal.

(e) If $W=\text{Span}\{\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_n\}$, then $W=\text{Span}\{\vec{w}_1,\vec{w}_2,\ldots,\vec{w}_n\}$ for some orthonormal set $\{\vec{w}_1,\vec{w}_2,\ldots,\vec{w}_n\}$

(f) If every $\vec{x}\in\mathbb{R}^2$ satisfies $\vec{x}=\left(\dfrac{\vec{x}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\right)\vec{u}+\left(\dfrac{\vec{x}\cdot\vec{v}}{\vec{v}\cdot\vec{v}}\right)\vec{v}$, then $\{\vec{u},\vec{v}\}$ is an orthogonal set.

Answer

(a) **Always False**

Orthogonal sets are always linearly independent.

(b) **Always True**

If the columns of A are $\vec{v}_1,\ldots,\vec{v}_n$, the entry in row i, column j of A^TA is $\vec{v}_i\cdot\vec{v}_j$, which is zero when $i\neq j$.

(c) **Always True**

If $\vec{v}\neq 0$, then the angle between $\vec{x},\vec{v}$ is $\cos^{-1}\left(\dfrac{\vec{x}\cdot\vec{v}}{\|\vec{x}\|\|\vec{v}\|}\right)=\cos^{-1}(-1)=180^\circ$, and when $\vec{v}=\vec{0}$ this is already true.

(d) **Sometimes True Sometimes False**

This is true using $\vec{u}=\vec{e}_1,\vec{v}=\vec{e}_2,\vec{w}=\vec{e}_3$ for one example.
This is false using $\vec{u}=\vec{e}_1,\vec{v}=\vec{e}_2,\vec{w}=\vec{e}_2$ for another example.

(e) **Sometimes True Sometimes False**

This is true when $\{\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_n\}$ is linearly independent, you can use Gram-Schmidt and normalize each vector to a unit vector. But in the case where $\{\vec{v}_1,\vec{v}_2,\ldots,\vec{v}_n\}$ is dependent, Gram-Schmidt will turn at least one vector into the zero vector and therefore the set will not have n linearly independent vectors.

(f) **Always True**

Plug in $\vec{x}=\vec{u}$ to prove $\vec{u}\cdot\vec{v}=0$ must hold.

3.7: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

4: Vector Spaces

4.1: General Vector Spaces and Subspaces

4.1: General Vector Spaces

4.1: Calculations in Vector Spaces

4.1: Practice Set

4.2: Linear Transformations in General Vector Spaces

4.2: Linear Transformations

4.2: Coordinatization

4.2: Practice Set

4.3: Examples and Applications

4.3: Differential Equations

4.3: Famous Transformations

4.3: Inner Product Spaces

4.3: Markov Chains

4.3: Page Rank

4.3: Practice Set

This page titled 4: Vector Spaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

4.1: General Vector Spaces and Subspaces

Outcomes

1. Develop the abstract concept of a vector space through axioms.
2. Deduce basic properties of vector spaces.
3. Use the vector space axioms to determine if a set and its operations constitute a vector space.
4. Prove or disprove a subset of a vector space is a subspace.

In this section, we consider the idea of an abstract vector space. A vector space is something which has two operations which mimic the structure of $\mathbb{R}^n$.

This page titled 4.1: General Vector Spaces and Subspaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

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

4.1: General Vector Spaces

We begin by defining exactly what properties of addition we appreciate about the vector space we have become familiar with $V = \mathbb{R}^n$:

Theorem 4.1.1: 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

$$\vec{v} + (-\vec{v}) = \vec{0}$$

And similarly, the properties of scalar multiplication (and how it interacts with addition) that we use repeatedly in $V = \mathbb{R}^n$. (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.)

Theorem 4.1.2: 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

$$\text{If } a \text{ is a real number, and } \vec{v} \text{ is in } V, \text{ then } a\vec{v} \text{ is in } V$$

- Distribution across Vector Addtion

$$a\left(\vec{v} + \vec{w}\right) = a\vec{v} + a\vec{w}$$

- Distribution across Scalar Addition

$$(a + b)\,\vec{v} = a\vec{v} + b\vec{v}$$

- Associative Law of Scalar Multiplication

$$a\left(b\vec{v}\right) = (ab)\vec{v}$$

- 1 is the Multiplicative Indentity

$$1\vec{v} = \vec{v}$$

If we can find *any* collection of objects and define addition and scalar multiplication in a way that they satisfy all of the axioms above, we have defined a vector space.

Definition: 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.

Consider the following example, in which we prove that $\mathbb{R}^n$ is in fact an example of a vector space.

Illustration 4.1.1: $\mathbb{R}^n$

$\mathbb{R}^n$, under the usual operations of vector addition and scalar multiplication, is a vector space.

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:

$$\vec{x}+\vec{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} = \begin{bmatrix} y_1+x_1 \\ y_2+x_2 \\ \vdots \\ y_n+x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \vec{y}+\vec{x}$$

  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:

$$(\vec{x}+\vec{y})+\vec{z} = \left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \right) + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} (x_1+y_1)+z_1 \\ (x_2+y_2)+z_2 \\ \vdots \\ (x_n+y_n)+z_n \end{bmatrix} =$$

$$\begin{bmatrix} x_1+(y_1+z_1) \\ x_2+(y_2+z_2) \\ \vdots \\ x_n+(y_n+z_n) \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1+z_1 \\ y_2+z_2 \\ \vdots \\ y_n+z_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \left(\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \right) = \vec{x}+(\vec{y}+\vec{z})$$

  Hence, addition of vectors is associative.
- Next, we show the existence of an additive identity. Let $\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$:

$$\vec{x}+\vec{0}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}+\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}=\begin{bmatrix}x_1+0\\x_2+0\\\vdots\\x_n+0\end{bmatrix}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\vec{x}$$

Hence, the zero vector $\vec{0}$ is an additive identity.

- Next, we prove the existence of an additive inverse. Let $-\vec{x}=\begin{bmatrix}-x_1\\-x_2\\\vdots\\-x_n\end{bmatrix}$:

$$\vec{x}+(-\vec{x})=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}+\begin{bmatrix}-x_1\\-x_2\\\vdots\\-x_n\end{bmatrix}=\begin{bmatrix}x_1-x_1\\x_2-x_2\\\vdots\\x_n-x_n\end{bmatrix}=\begin{bmatrix}0\\0\\\vdots\\0\end{bmatrix}=\vec{0}$$

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 $a(\vec{x}+\vec{y})=a\vec{x}+a\vec{y}$:

$$a(\vec{x}+\vec{y})=a\left(\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}+\begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}\right)=a\begin{bmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{bmatrix}=\begin{bmatrix}a(x_1+y_1)\\a(x_2+y_2)\\\vdots\\a(x_n+y_n)\end{bmatrix}=\begin{bmatrix}ax_1+ay_1\\ax_2+ay_2\\\vdots\\ax_n+ay_n\end{bmatrix}$$
$$=\begin{bmatrix}ax_1\\ax_2\\\vdots\\ax_n\end{bmatrix}+\begin{bmatrix}ay_1\\ay_2\\\vdots\\ay_n\end{bmatrix}=a\vec{x}+a\vec{y}$$

- Next, we wish to show that $(a+b)\vec{x}=a\vec{x}+b\vec{x}$:

$$(a+b)\vec{x}=(a+b)\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=\begin{bmatrix}(a+b)x_1\\(a+b)x_2\\\vdots\\(a+b)x_n\end{bmatrix}=\begin{bmatrix}ax_1+bx_1\\ax_2+bx_2\\\vdots\\ax_n+bx_n\end{bmatrix}$$
$$=\begin{bmatrix}ax_1\\ax_2\\\vdots\\ax_n\end{bmatrix}+\begin{bmatrix}bx_1\\bx_2\\\vdots\\bx_n\end{bmatrix}=a\vec{x}+b\vec{x}$$

- We wish to show that $a(b\vec{x})=(ab)\vec{x}$:

$$a(b\vec{x})=a\left(b\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\right)=a\left(\begin{bmatrix}bx_1\\bx_2\\\vdots\\bx_n\end{bmatrix}\right)=\begin{bmatrix}a(bx_1)\\a(bx_2)\\\vdots\\a(bx_n)\end{bmatrix}$$
$$=\begin{bmatrix}(ab)x_1\\(ab)x_2\\\vdots\\(ab)x_n\end{bmatrix}=(ab)\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=(ab)\vec{x}$$

- Finally, we need to show that $1\vec{x} = \vec{x}$:

$$1\vec{x} = 1\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 1x_1 \\ 1x_2 \\ \vdots \\ 1x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \vec{x}$$

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.

We now consider some examples of vector spaces.

Illustration 4.1.2: Vector Space of Polynomials

Let $\mathbb{P}_n(\mathbb{R})$ be the set of all polynomials of at most degree n 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 $\mathbb{P}_n(\mathbb{R})$ is a vector space.

We can write $\mathbb{P}_n(\mathbb{R})$ explicitly as

$$\mathbb{P}_n(\mathbb{R}) = \{a_n x^n + \cdots + a_1 x + a_0 \, : \, a_i \in \mathbb{R} \text{ for all } i\}$$

To show that $\mathbb{P}_n$ is a vector space, we verify the axioms. Let $p(x), q(x), r(x)$ be polynomials in $\mathbb{P}_n(\mathbb{R})$ and let a, b, c be real numbers. Write $p(x) = p_n x^n + \cdots + p_1 x + p_0$, $q(x) = q_n x^n + \cdots + q_1 x + q_0$, and $r(x) = r_n x^n + \cdots + r_1 x + r_0$.

- We first prove that addition of polynomials in $\mathbb{P}_n(\mathbb{R})$ is closed. For two polynomials in $\mathbb{P}_n(\mathbb{R})$ we need to show that their sum is also a polynomial in $\mathbb{P}_n(\mathbb{R})$. From the definition of $\mathbb{P}_n(\mathbb{R})$, a polynomial is contained in $\mathbb{P}_n(\mathbb{R})$ if it is of degree at most n or the zero polynomial.

$$\begin{aligned} p(x) + q(x) &= p_n x^n + \cdots + p_1 x + p_0 + q_n x^n + \cdots + q_1 x + q_0 \\ &= (p_n + q_n)x^n + \cdots + (p_1 + q_1)x + (p_0 + q_0) \end{aligned}$$

  The sum is a polynomial of degree n and therefore is in $\mathbb{P}_n(\mathbb{R})$. It follows that $\mathbb{P}_n(\mathbb{R})$ is closed under addition.
- We need to show that addition is commutative, that is $p(x) + q(x) = q(x) + p(x)$:

$$\begin{aligned} p(x) + q(x) &= p_n x^n + \cdots + p_1 x + p_0 + q_n x^n + \cdots + q_1 x + q_0 \\ &= (p_n + q_n)x^n + \cdots + (p_1 + q_1)x + (p_0 + q_0) \\ &= (q_n + p_n)x^n + \cdots + (q_1 + p_1)x + (q_0 + p_0) \\ &= q_n x^n + \cdots + q_1 x + q_0 + p_n x^n + \cdots + p_1 x + p_0 \\ &= q(x) + p(x) \end{aligned}$$

- Next, we need to show that addition is associative. That is, that $(p(x) + q(x)) + r(x) = p(x) + (q(x) + r(x))$:

$$\begin{aligned} (p(x) + q(x)) + r(x) &= (p_n x^n + \cdots + p_1 x + p_0 + q_n x^n + \cdots + q_1 x + q_0) + r_n x^n + \cdots + r_1 x + r_0 \\ &= (p_n + q_n)x^n + \cdots + (p_1 + q_1)x + (p_0 + q_0) + r_n x^n + \cdots + r_1 x + r_0 \\ &= (p_n + q_n + r_n)x^n + \cdots + (p_1 + q_1 + r_1)x + (p_0 + q_0 + r_0) \\ &= p_n x^n + \cdots + p_1 x + p_0 + (q_n + r_n)x^n + \cdots + (q_1 + r_1)x + (q_0 + r_0) \\ &= p_n x^n + \cdots + p_1 x + p_0 + (q_n x^n + \cdots + q_1 x + q_0 + r_n x^n + \cdots + r_1 x + r_0) \\ &= p(x) + (q(x) + r(x)) \end{aligned}$$

- Next, we must prove that there exists an additive identity. Let $0(x) = 0x^n + \cdots + 0x + 0$:

$$\begin{aligned} p(x) + 0(x) &= p_n x^n + \cdots + p_1 x + p_0 + 0x^n + \cdots + 0x + 0 \\ &= (p_n + 0)x^n + \cdots + (p_1 + 0)x + (p_0 + 0) \\ &= p_n x^n + \cdots + p_1 x + p_0 \\ &= p(x) \end{aligned}$$

  Hence, an additive identity exists, specifically the zero polynomial.
- Next, we must prove that there exists an additive inverse. Let $-p(x) = -p_n x^n - \cdots - p_1 x - p_0$ and consider the following:

$$\begin{aligned} p(x)+(-p(x)) &= p_n x^n + \cdots + p_1 x + p_0 + (-p_n x^n - \cdots - p_1 x - p_0) \\ &= (p_n - p_n)x^n + \cdots + (p_1 - p_1)x + (p_0 - p_0) \\ &= 0x^n + \cdots + 0x + 0 \\ &= 0(x) \end{aligned}$$

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 $\mathbb{P}_n(\mathbb{R})$ is closed under scalar multiplication. That is, we show that $ap(x)$ is also a polynomial of degree at most n.

$$ap(x) = a\left(p_n x^n + \cdots + p_1 x + p_0\right) = ap_n x^n + \cdots + ap_1 x + ap_0$$

  Therefore, $\mathbb{P}_n(\mathbb{R})$ is closed under scalar multiplication.
- We need to show that $a(p(x)+q(x)) = ap(x) + aq(x)$:

$$\begin{aligned} a(p(x)+q(x)) &= a\left(p_n x^n + \cdots + p_1 x + p_0 + q_n x^n + \cdots + q_1 x + q_0\right) \\ &= a\left((p_n + q_n)x^n + \cdots + (p_1 + q_1)x + (p_0 + q_0)\right) \\ &= a(p_n + q_n)x^n + \cdots + a(p_1 + q_1)x + a(p_0 + q_0) \\ &= (ap_n + aq_n)x^n + \cdots + (ap_1 + aq_1)x + (ap_0 + aq_0) \\ &= ap_n x^n + \cdots + ap_1 x + ap_0 + aq_n x^n + \cdots + aq_1 x + aq_0 \\ &= ap(x) + aq(x) \end{aligned}$$

- Next, we show that $(a+b)p(x) = ap(x) + bp(x)$:

$$\begin{aligned} (a+b)p(x) &= (a+b)(p_n x^n + \cdots + p_1 x + p_0) \\ &= (a+b)p_n x^n + \cdots + (a+b)p_1 x + (a+b)p_0 \\ &= ap_n x^n + \cdots + ap_1 x + ap_0 + bp_n x^n + \cdots + bp_1 x + bp_0 \\ &= ap(x) + bp(x) \end{aligned}$$

- The next axiom which needs to be verified is $a(bp(x)) = (ab)p(x)$:

$$\begin{aligned} a(bp(x)) &= a\left(b\left(p_n x^n + \cdots + p_1 x + p_0\right)\right) \\ &= a\left(bp_n x^n + \cdots + bp_1 x + bp_0\right) \\ &= abp_n x^n + \cdots + abp_1 x + abp_0 \\ &= (ab)(p_n x^n + \cdots + p_1 x + p_0) \\ &= (ab)p(x) \end{aligned}$$

- Finally, we show that $1p(x) = p(x)$:

$$\begin{aligned} 1p(x) &= 1\left(p_n x^n + \cdots + p_1 x + p_0\right) \\ &= 1p_n x^n + \cdots + 1p_1 x + 1p_0 \\ &= p_n x^n + \cdots + p_1 x + p_0 \\ &= p(x) \end{aligned}$$

  Since the above axioms hold, we know that $\mathbb{P}_n(\mathbb{R})$ as described above is a vector space.

Another important example of a vector space is the set of all matrices of the same size.

Illustration 4.1.3: Vector Space of Matrices

Let $\mathbb{M}_{m\times n}(\mathbb{R})$ be the set of all $m \times n$ matrices. Using the usual operations of matrix addition and scalar multiplication, show that $\mathbb{M}_{m\times n}(\mathbb{R})$ is a vector space.

Let A, B be $m \times n$ matrices in $\mathbb{M}_{m\times n}(\mathbb{R})$. We first prove the axioms for addition.

- In order to prove that $\mathbb{M}_{m\times n}(\mathbb{R})$ is closed under matrix addition, we show that the sum $A+B$ is in $\mathbb{M}_{m\times n}(\mathbb{R})$. This means showing that $A+B$ is an $m \times n$ matrix.

$$A+B=\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}+\begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{bmatrix}$$
$$=\begin{bmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{bmatrix}$$

You can see that the sum is a $m \times n$ matrix, so it is in $\mathbb{M}_{m\times n}(\mathbb{R})$. It follows that $\mathbb{M}_{m\times n}(\mathbb{R})$ is closed under matrix addition.

- The remaining axioms regarding matrix addition follow from properties of matrix addition. Therefore, $\mathbb{M}_{m\times n}(\mathbb{R})$ satisfies the axioms of vector addition.

We now turn our attention to the axioms regarding scalar multiplication. Let A be a matrix in $\mathbb{M}_{m\times n}(\mathbb{R})$ and let c be a real number.

- We first show that $\mathbb{M}_{m\times n}(\mathbb{R})$ is closed under scalar multiplication. That is, we show that cA is an $m \times n$ matrix.

$$cA=c\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$
$$=\begin{bmatrix} ca_{11} & \cdots & ca_{1n} \\ \vdots & \ddots & \vdots \\ ca_{m1} & \cdots & ca_{mn} \end{bmatrix}$$

This is a $m \times n$ matrix in $\mathbb{M}_{m\times n}(\mathbb{R})$ 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, $\mathbb{M}_{m\times n}$ satisfies the axioms of scalar multiplication.

In conclusion, $\mathbb{M}_{m\times n}(\mathbb{R})$ satisfies the required axioms and is a vector space.

We now examine an example of a set that does not satisfy all of the above axioms, and is therefore *not* a vector space. Here we stress that the vector space is not just the objects themselves, but how we define addition and scalar multiplication.

Illustration 4.1.4: Not a Vector Space

Let V denote the set of 2×2 matrices. Let addition in V be defined by $\begin{bmatrix} a & b \\ c & d \end{bmatrix}+\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}=\begin{bmatrix} a+\delta & b+\gamma \\ 0 & 0 \end{bmatrix}$ for matrices in V. Let scalar multiplication in V be the usual scalar multiplication of matrices.

This V is not a vector space. 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 $A+B$ is a 2×2 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 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}+\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}=\begin{bmatrix} a+\delta & b+\gamma \\ 0 & 0 \end{bmatrix}$ and $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}+\begin{bmatrix} a & b \\ c & d \end{bmatrix}=\begin{bmatrix} \alpha+d & \beta+c \\ 0 & 0 \end{bmatrix}$. Therefore, we can find A, B in V such that these sums are not equal. One example is

$$A=\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B=\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

It follows that $A+B=\begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 5 & 4 \\ 0 & 0 \end{bmatrix}=B+A$.

Therefore, addition as defined for V is not commutative and V fails this axiom. Hence, V is not a vector space.

Consider another example of a vector space.

Illustration 4.1.5: Vector Space of Functions

Let S be a nonempty set and define $\mathbb{F}(S,\mathbb{R})$ to be the set of *real* functions defined on S. In other words, we write $\mathbb{F}(S,\mathbb{R}) : S \mapsto \mathbb{R}$. Letting a,b,c be scalars and f,g,h functions, the vector operations are defined as:

$$\begin{aligned}(f+g)(x) &= f(x)+g(x)\\ (af)(x) &= a(f(x))\end{aligned}$$

To verify that $\mathbb{F}(S,\mathbb{R})$ is a vector space, we must prove the axioms beginning with those for addition. Let f,g,h be functions in $\mathbb{F}(S,\mathbb{R})$.

- 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 $\mathbb{F}(S,\mathbb{R})$ and $\mathbb{F}(S,\mathbb{R})$ 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:

  $$((f+g)+h)(x) = (f+g)(x)+h(x) = (f(x)+g(x))+h(x)$$

  $$= f(x)+(g(x)+h(x)) = (f(x)+(g+h)(x)) = (f+(g+h))(x)$$

  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 $\mathbb{F}(S,\mathbb{R})$ 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 $\mathbb{F}(S,\mathbb{R})$ and $\mathbb{F}(S,\mathbb{R})$ is closed under scalar multiplication.
- $$((a+b)f)(x) = (a+b)f(x) = af(x)+bf(x) = (af+bf)(x)$$

  and so $(a+b)f = af+bf$.
- $$(a(f+g))(x) = a(f+g)(x) = a(f(x)+g(x))$$

  $$= af(x)+ag(x) = (af+ag)(x)$$

  and so $a(f+g) = af+ag$.
- $$((ab)f)(x) = (ab)f(x) = a(bf(x)) = (a(bf))(x)$$

  so $(abf) = a(bf)$.
- Finally $(1f)(x) = 1f(x) = f(x)$ so $1f=f$.

It follows that $\mathbb{F}(S,\mathbb{R})$ satisfies all the required axioms and is a vector space.

Consider the following important theorem.

Theorem 4.1.3: 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:

   $$\vec{0} = \vec{0} + \vec{0}' = \vec{0}' + \vec{0} = \vec{0}'$$

   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

   $$\begin{aligned} 0\vec{x} + (-0\vec{x}) &= 0\vec{x} + 0\vec{x} + (-0\vec{x}) \\ \vec{0} + 0 &= 0\vec{x} + 0 \\ \vec{0} &= 0\vec{x} \end{aligned}$$

   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 4.1.4

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 $(-\vec{v})$. The proof is left as an exercise to the reader.

4.1: General Vector Spaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

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

4.1: Calculations in Vector Spaces

Now that we have objects that "act like $\mathbb{R}^n$", we can do similar things to them.

Definition: Span of Vectors

Let $\{\vec{v}_1, \cdots, \vec{v}_n\} \subseteq V$. Then the **span** of these vectors is

$$\text{Span}\{\vec{v}_1, \cdots, \vec{v}_n\} = \left\{ \sum_{i=1}^{n} c_i \vec{v}_i : c_i \in \mathbb{R} \right\}$$

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 $\vec{v}_i$:

$$\vec{w} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$$

We say that a collection of vectors $\{\vec{v}_1, \cdots, \vec{v}_n\}$ is a **spanning set** for V if $V = \text{Span}\{\vec{v}_1, \cdots, \vec{v}_n\}$

Example 4.1.1: Matrix Span

Let $M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, and $M_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

(a) Determine whether or not $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ is in $\text{Span}\{M_1, M_2\}$

(b) Determine whether or not $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is in $\text{Span}\{M_1, M_2\}$

Solution

(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 $\boxed{A \text{ is in } \text{Span}\{M_1, M_2\}}$

(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, $\boxed{B \text{ is not in } \text{Span}\{M_1, M_2\}}$

Example 4.1.2: Polynomial Span

Determine if $p(x) = 7x^2 + 4x - 3$ is in $\text{Span}\{4x^2 + x, x^2 - 2x + 3\}$

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 system of equations must hold:

$$\begin{cases} 4a+b=7 \\ a-2b=4 \\ 3b=-3 \end{cases}$$

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, $\boxed{p(x) \text{ is in the given span}}$

Illustration 4.1.6: Spanning Set

The set $S=\{x^2+1, x-2, 2x^2-x\}$ is a spanning set for $\mathbb{P}_2$, the set of all polynomials of degree at most 2.

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 for all choices of a, b, c:

$$\begin{cases} r+2t=a \\ s-t=b \\ r-2s=c \end{cases}$$

To check that a solution exists, set up the matrix and row reduce to see if it is invertible:

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \xrightarrow{\text{RREF}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Clearly, this matrix has RREF I, so it is invertible and a solution exists for any choice of a, b, c. Hence, S is a spanning set for $\mathbb{P}_2$.

We can also expand the definition of linear independence to these general vector spaces.

Definition 4.1.3: Linear Independence and Bases

Let $\{\vec{v}_1, \cdots, \vec{v}_n\} \subseteq V$. This set of vectors is called **linearly independent** if the only solution to

$$c_1\vec{v}_1+\cdots+c_n\vec{v}_n=\vec{0}$$

is $c_1=c_2=\cdots=c_n=0$, the trivial linear combination.

When $\{\vec{v}_1, \cdots, \vec{v}_n\} \subseteq V$ is both linearly independent and a spanning set of V, we call it a **basis** of V.

If V has basis $\{\vec{v}_1, \cdots, \vec{v}_n\}$, we say $dim(V)=n$, or that V has n-dimensions.

This matches with our definitions in $\mathbb{R}^n$, so all previous theorems regarding linear independence, bases, and dimension still hold.

Example 4.1.3: Basis of a Vector Space

Let $S = \{x^2+1, x-2, 2x^2-x\}$

(a) Show that S is a basis for $\mathbb{P}_2$.

(b) What is the dimension of $\mathbb{P}_2$?

(c) What is the dimension of $\mathbb{P}_n$?

Solution

(a) As we saw in Illustration 4.1.6 above, S is a spanning set of $\mathbb{P}_2$. To prove it is a basis, we need the vectors inside to be linearly independent. For $\vec{0} = 0(x) = 0x^2 + 0x + 0$ in $\mathbb{P}_2$, can we find r, s, t such that:

$$0(x) = r(x^2+1) + s(x-2) + t(2x^2 - x)$$

For this to be true, the following must hold:

$$\begin{cases} r + 2t = 0 \\ s - t = 0 \\ r - 2s = 0 \end{cases}$$

We already showed that the matrix $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix}$ is invertible, so we know the only solution to $A \begin{bmatrix} r \\ s \\ t \end{bmatrix} = \vec{0}$ is the trivial one.

This proves that S is linearly independent, and must therefore be a basis for $\mathbb{P}_2$.

(b) Since there is a basis for $\mathbb{P}_2$ containing exactly 3 vectors, we must have $\dim(\mathbb{P}_2) = 3$.

This means any such basis for $\mathbb{P}_2$ would contain 3 vectors. A simple example would be $S' = \{x^2, x, 1\}$.

Clearly, these vectors are a spanning set of $\mathbb{P}_2$ and must be linearly independent. This proves $\boxed{\dim(\mathbb{P}_2) = 3}$

(c) The easiest linearly independent spanning set of $\mathbb{P}_n$ would be $\{x^n, \ldots, x^2, x, 1\}$.

Since this is a basis for $\mathbb{P}_n$, and it contains exactly $n+1$ vectors, we have $\boxed{\dim(\mathbb{P}_n) = n+1}$

Consider how the plane $W = \left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \right\}$ in $\mathbb{R}^3$ is the span of two linearly independent vectors $\{\vec{e}_1, \vec{e}_2\}$. This collection of vectors from $\mathbb{R}^3$ together act like an entirely different vector space: $\mathbb{R}^2$.

To be clear, this is **not saying** that $W = \mathbb{R}^2$, but that it "acts like" a copy of $\mathbb{R}^2$ inside of $\mathbb{R}^3$.

Definition 4.1.4: Subspaces

Let $W \subseteq V$ be nonempty. If W itself is a vector space (satisfies the addition and scalar multiplication axioms), then we say that W is a **subspace** of V.

Note that the axioms of addition and scalar multiplication are satisfied by all vectors in V, we do not need to check most of the properties again.

Instead, we only need to check that:

1. $\vec{0}$ is in W.

2. $a\vec{x} + b\vec{y} \in W$ for any real numbers a, b and any vectors $\vec{x}, \vec{y}$ in W.

Theorem 4.1.5: 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.

Illustration 4.1.7: 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.

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 $\left\{\vec{0}\right\}$

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

$$a\vec{w}_1 + b\vec{w}_2 = a\vec{0} + b\vec{0} = \vec{0}$$

It follows that the sum is contained in $\left\{\vec{0}\right\}$ and 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**.

Consider another example.

Example 4.1.4: 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 $\mathbb{P}_2$ and find its dimension.

Solution

First, express W as follows:

$$W = \left\{p(x) = ax^2 + bx + c \,:\, p(1) = 0\right\}$$

We need to show that W satisfies the conditions of a subspace:

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 $ap(x) + bq(x)$. Let $r(x)$ represent this sum.

$$r(1) = ap(1) + bq(1) = a0 + b0 = 0$$

Therefore, the sum is also in W and the second condition is satisfied.

It follows that W is a subspace of $\mathbb{P}_2$.

To find its dimension, we need to find a basis for W.

For any polynomial $p(x) = ax^2 + bx + c$ in W, we know that $p(1) = a + b + c = 0$ must hold.

This is a single equation with three unknowns, so we know its solution can be written as $a = -b - c$ and b, c are free variables.

$$\begin{aligned} W &= \left\{ax^2 + bx + c \,:\, a = -b - c\right\} \\ &= \left\{(-b - c)x^2 + bx + c\right\} \\ &= \left\{b(-x^2 + x) + c(-x^2 + 1)\right\} \\ &= \text{Span}\left\{-x^2 + x, -x^2 + 1\right\} \end{aligned}$$

Clearly, these two vectors are not scalar multiples of one another, so they must be linearly independent. Since they are a spanning set for W, we know that $\{-x^2+x, -x^2+1\}$ is a basis of W, and $\boxed{\dim(W)=2}$

Note

Just like in $\mathbb{R}^n$, the dimension of any subspace W of V must satisfy $\dim(W) \leq \dim(V)$.

If W satisfies $\dim(W)=0$, then $W=\{\vec{0}\}$, and if a finite dimensional subspace W of V satisfies $\dim(W)=\dim(V)$, then $W=V$.

4.1: Calculations in Vector Spaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **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** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.
- **9.4: Subspaces and Basis** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

4.1.3: Practice Set

Exercise 4.1.3.1a

Fix some $r \in \mathbb{R}$ and consider the collection of vectors $V = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \right\}$ with addition and scalar multiplication defined as:

$\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x+a+r \\ y+b-r \end{bmatrix}$ and $t \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} tx+tr-r \\ ty-tr+r \end{bmatrix}$

Is V a vector space or not? Justify your answer.

Answer

1. $\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x+a+r \\ y+b-r \end{bmatrix}$ so the sum of two vectors in V is still in V.
2. $\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x+a+r \\ y+b-r \end{bmatrix} = \begin{bmatrix} a+x+r \\ b+y-r \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \oplus \begin{bmatrix} x \\ y \end{bmatrix}$
3. $\left(\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} \right) \oplus \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} x+a+r \\ y+b-r \end{bmatrix} \oplus \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} x+a+c+2r \\ b+y+d-2r \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a+c+r \\ b+d-r \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} \oplus \left(\begin{bmatrix} a \\ b \end{bmatrix} \oplus \begin{bmatrix} c \\ d \end{bmatrix} \right)$
4. We have $\vec{0} = \begin{bmatrix} -r \\ r \end{bmatrix}$ since $\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} -r \\ r \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$
5. We have $-\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x-2r \\ -y+2r \end{bmatrix}$ since $\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} -x-2r \\ -y+2r \end{bmatrix} = \begin{bmatrix} -r \\ r \end{bmatrix} = \vec{0}$
6. $t \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} tx+tr-r \\ ty-tr+r \end{bmatrix}$ is still in V for all $t \in \mathbb{R}$.
7. $t \odot \left(\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} \right) = t \odot \begin{bmatrix} x+a+r \\ y+b-r \end{bmatrix} = \begin{bmatrix} t(x+a+r)+tr-r \\ t(y+b-r)-tr+r \end{bmatrix} = \begin{bmatrix} tx+ta+2tr-r \\ ty+tb-2tr+r \end{bmatrix}$
$\left(t \odot \begin{bmatrix} x \\ y \end{bmatrix} \right) \oplus \left(t \odot \begin{bmatrix} a \\ b \end{bmatrix} \right) = \begin{bmatrix} tx+tr-r \\ ty-tr+r \end{bmatrix} \oplus \begin{bmatrix} ta+tr-r \\ tb-tr+r \end{bmatrix} = \begin{bmatrix} (tx+tr-r)+(ta+tr-r)+r \\ (ty-tr+r)+(tb-tr+r)-r \end{bmatrix} = \begin{bmatrix} tx+ta+2tr-r \\ ty+tb-2tr+r \end{bmatrix}$
8. $(t+s) \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (t+s)x+(t+s)r-r \\ (t+s)y-(t+s)r+r \end{bmatrix} = \begin{bmatrix} (tx+tr-r)+(sx+sr-r)+r \\ (ty-tr+r)+(sy-sr+r)-r \end{bmatrix} = \left(t \odot \begin{bmatrix} x \\ y \end{bmatrix} \right) \oplus \left(s \odot \begin{bmatrix} x \\ y \end{bmatrix} \right)$
9. $(ts) \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} (ts)x+(ts)r-r \\ (ts)y-(ts)r+r \end{bmatrix} = \begin{bmatrix} t(sx+sr-r)+tr-r \\ t(sy-sr+r)-tr+r \end{bmatrix} = t \odot \begin{bmatrix} sx+sr-r \\ sy-sr+r \end{bmatrix} = t \odot \left(s \odot \begin{bmatrix} x \\ y \end{bmatrix} \right)$
10. $1 \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1x+1r-r \\ 1y-1r+r \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$

All ten rules hold, so this **is a vector space**.

Exercise 4.1.3.1b

Consider the collection of vectors $V = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \right\}$ with addition and scalar multiplication defined as:

$\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x+b \\ y+a \end{bmatrix}$ and the usual $\mathbb{R}^2$ scalar multiplication

Is V a vector space or not? Justify your answer.

Answer

1. $\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x+b \\ y+a \end{bmatrix}$ so the sum of two vectors in V is still in V.
2. $\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x+b \\ y+a \end{bmatrix}$ but $\begin{bmatrix} a \\ b \end{bmatrix} \oplus \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a+y \\ b+x \end{bmatrix}$

These are not always equivalent, so this **is not a vector space.**

Exercise 4.1.3.1c

Consider the collection of vectors $V = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} \right\}$ with addition and scalar multiplication defined as:

The usual $\mathbb{R}^2$ addition and $t \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Is V a vector space or not? Justify your answer.

Answer

Since we are using the standard vector addition, the first five rules will work automatically.

6. $t \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is still in V for each $t \in \mathbb{R}$.

7. $t \odot \left(\begin{bmatrix} x \\ y \end{bmatrix} \oplus \begin{bmatrix} a \\ b \end{bmatrix} \right) = t \odot \begin{bmatrix} x+a \\ y+b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$\left(t \odot \begin{bmatrix} x \\ y \end{bmatrix} \right) \oplus \left(t \odot \begin{bmatrix} a \\ b \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0+0 \\ 0+0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

8. $(t+s) \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \left(t \odot \begin{bmatrix} x \\ y \end{bmatrix} \right) \oplus \left(s \odot \begin{bmatrix} x \\ y \end{bmatrix} \right)$

9. $(ts) \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = t \odot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = t \odot \left(s \odot \begin{bmatrix} x \\ y \end{bmatrix} \right)$

10. $1 \odot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \neq \begin{bmatrix} x \\ y \end{bmatrix}$

This fails the tenth rule, so it **is not a vector space.**

Exercise 4.1.3.1d

Consider the collection of vectors $V = \{p(x) \in \mathbb{P}_2(\mathbb{R})\}$ with addition and scalar multiplication defined as:

$p(x) \oplus q(x) = p(x)$ and the usual polynomial scalar multiplication

Is V a vector space or not? Justify your answer.

Answer

1. $p(x) \oplus q(x) = p(x)$ is still a polynomial of degree at most 2, so it is in V.

2. $p(x) \oplus q(x) = p(x)$ but $q(x) \oplus p(x) = q(x)$.

These are not always equivalent, so this **is not a vector space.**

Exercise 4.1.3.1e

Consider the collection of vectors $V = \{p(x) \in \mathbb{P}_2(\mathbb{R})\}$ with addition and scalar multiplication defined as:

The usual polynomial addition and $t \odot p(x) = p(t)$

Is V a vector space or not? Justify your answer.

Answer

Since we are using the standard polynomial addition, the first five rules will work automatically.

6. $t \odot p(x) = p(t)$ is a constant, which is a polynomial of degree at most 2, so it is in V

7. $t \odot (p(x) \oplus q(x)) = (p+q)(t) = p(t) + q(t) = (t \odot p(x)) \oplus (t \odot q(x))$

8. $(t+s)\odot p(x) = p(t+s)$ but $(t\odot p(x))\oplus(s\odot p(x)) = p(t)+p(s)$.

These are not always equivalent, so this **is not a vector space.**

Exercise 4.1.3.2a

Consider the set $\left\{\begin{bmatrix}1&0\\0&1\end{bmatrix},\begin{bmatrix}1&1\\1&1\end{bmatrix},\begin{bmatrix}0&1\\0&1\end{bmatrix},\begin{bmatrix}0&1\\1&1\end{bmatrix}\right\}$

(i) Is this set linearly independent?

(ii) Is this set a spanning set of $\mathbb{M}_{2\times2}(\mathbb{R})$?

(iii) Is this set a basis of $\mathbb{M}_{2\times2}(\mathbb{R})$?

Answer

(i) We set up $a\begin{bmatrix}1&0\\0&1\end{bmatrix}+b\begin{bmatrix}1&1\\1&1\end{bmatrix}+c\begin{bmatrix}0&1\\0&1\end{bmatrix}+d\begin{bmatrix}0&1\\1&1\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}$

This is equivalent to $\begin{bmatrix}a+b&b+c+d\\b+d&a+b+c+d\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}\leftrightarrow\begin{cases}a+b=0\\b+c+d=0\\b+d=0\\a+b+c+d=0\end{cases}$

We solve this homogeneous system by row reducing: $\begin{bmatrix}1&1&0&0\\0&1&1&1\\0&1&0&1\\1&1&1&1\end{bmatrix}\xrightarrow{RREF}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$

Since there are no free variables, the only solution is $a=b=c=d=0$ and the set **is linearly independent.**

(ii) $a\begin{bmatrix}1&0\\0&1\end{bmatrix}+b\begin{bmatrix}1&1\\1&1\end{bmatrix}+c\begin{bmatrix}0&1\\0&1\end{bmatrix}+d\begin{bmatrix}0&1\\1&1\end{bmatrix}=\begin{bmatrix}\alpha&\beta\\\gamma&\delta\end{bmatrix}$

This is equivalent to $\begin{bmatrix}a+b&b+c+d\\b+d&a+b+c+d\end{bmatrix}=\begin{bmatrix}\alpha&\beta\\\gamma&\delta\end{bmatrix}\leftrightarrow\begin{cases}a+b=\alpha\\b+c+d=\beta\\b+d=\gamma\\a+b+c+d=\delta\end{cases}$

Since we saw in part (i) that the RREF of the matrix has a pivot in every row, this system will be consistent for any $\alpha,\beta,\gamma,\delta\in\mathbb{R}$

Therefore this **is a spanning set for $\mathbb{M}_{2\times2}(\mathbb{R})$**

(iii) Since the set is both linearly independent and spans $\mathbb{M}_{2\times2}(\mathbb{R})$, **it is a basis.**

Exercise 4.1.3.2b

Consider the set $\left\{\begin{bmatrix}2&-1\\0&1\end{bmatrix},\begin{bmatrix}0&1\\0&1\end{bmatrix},\begin{bmatrix}1&1\\1&1\end{bmatrix},\begin{bmatrix}0&1\\1&0\end{bmatrix}\right\}$

(i) Is this set linearly independent?

(ii) Is this set a spanning set of $\mathbb{M}_{2\times2}(\mathbb{R})$?

(iii) Is this set a basis of $\mathbb{M}_{2\times2}(\mathbb{R})$?

Answer

(i) We set up $a\begin{bmatrix}2&-1\\0&1\end{bmatrix}+b\begin{bmatrix}0&1\\0&1\end{bmatrix}+c\begin{bmatrix}1&1\\1&1\end{bmatrix}+d\begin{bmatrix}0&1\\1&0\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}$

This is equivalent to $\begin{bmatrix} 2a+c & -a+b+c+d \\ c+d & a+b+c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \leftrightarrow \begin{cases} 2a+c=0 \\ -a+b+c+d=0 \\ c+d=0 \\ a+b+c=0 \end{cases}$

We solve this homogeneous system by row reducing: $\begin{bmatrix} 2 & 0 & 1 & 0 \\ -1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

Since there is a free variable, the set **is linearly dependent.**

(ii) $a\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} + c\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + d\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$

This is equivalent to $\begin{bmatrix} 2a+c & -a+b+c+d \\ c+d & a+b+c \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \leftrightarrow \begin{cases} 2a+c=\alpha \\ -a+b+c+d=\beta \\ c+d=\gamma \\ a+b+c=\delta \end{cases}$

Since we saw in part (i) that the RREF of the matrix does not have a pivot in every row, this system will be inconsistent for some choices of $\alpha, \beta, \gamma, \delta \in \mathbb{R}$

Therefore this **is not a spanning set for $\mathbb{M}_{2\times 2}(\mathbb{R})$**

(iii) Since the set is neither linearly independent nor spans $\mathbb{M}_{2\times 2}(\mathbb{R})$, **it is not a basis.**

? Exercise 4.1.3.2c

Consider the set $\left\{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

(i) Is this set linearly independent?

(ii) Is this set a spanning set of $\mathbb{M}_{2\times 2}(\mathbb{R})$?

(iii) Is this set a basis of $\mathbb{M}_{2\times 2}(\mathbb{R})$?

Answer

(i) We set up $a\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + b\begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix} + c\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

This is equivalent to $\begin{bmatrix} a+4b+c & 2a+b \\ 3a+3b & 4a+2b+c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \leftrightarrow \begin{cases} a+4b+c=0 \\ 2a+b=0 \\ 3a+3b=0 \\ 4a+2b+c=0 \end{cases}$

We solve this homogeneous system by row reducing: $\begin{bmatrix} 1 & 4 & 1 \\ 2 & 1 & 0 \\ 3 & 3 & 0 \\ 4 & 2 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

Since there are no free variables, the set **is linearly independent.**

(ii) $a\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + b\begin{bmatrix} 4 & 1 \\ 3 & 2 \end{bmatrix} + c\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$

This is equivalent to $\begin{bmatrix} a+4b+c & 2a+b \\ 3a+3b & 4a+2b+c \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \leftrightarrow \begin{cases} a+4b+c=\alpha \\ 2a+b=\beta \\ 3a+3b=\gamma \\ 4a+2b+c=\delta \end{cases}$

Since we saw in part (i) that the RREF of the matrix does not have a pivot in every row, this system will be inconsistent for some choices of $\alpha, \beta, \gamma, \delta \in \mathbb{R}$

Therefore this **is not a spanning set for $\mathbb{M}_{2\times 2}(\mathbb{R})$**

(iii) Since the set does not span $\mathbb{M}_{2\times 2}(\mathbb{R})$, **it is not a basis.**

Exercise 4.1.3.3a

Consider the set $\{x^2+x+1, x^2-x+1, 2x\}$

(i) Is this set linearly independent?

(ii) Is this set a spanning set of $\mathbb{P}_2(\mathbb{R})$?

(iii) Is this set a basis of $\mathbb{P}_2(\mathbb{R})$?

Answer

(i) We set up $a(x^2+x+1)+b(x^2-x+1)+c(2x)=0x^2+0x+0$

This is equivalent to $(a+b)x^2+(a-b+2c)x+(a+b)=0x^2+0x+0 \longleftrightarrow \begin{cases} a+b=0 \\ a-b+2c=0 \\ a+b=0 \end{cases}$

We solve this homogeneous system by row reducing: $\begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 2 \\ 1 & 1 & 0 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$

Since there is a free variable, the set **is linearly dependent.**

(ii) $a(x^2+x+1)+b(x^2-x+1)+c(2x)=\alpha x^2+\beta x+\gamma$

This is equivalent to $(a+b)x^2+(a-b+2c)x+(a+b)=\alpha x^2+\beta x+\gamma \longleftrightarrow \begin{cases} a+b=\alpha \\ a-b+2c=\beta \\ a+b=\gamma \end{cases}$

Since we saw in part (i) that the RREF of the matrix does not have a pivot in every row, this system will be inconsistent for some choices of $\alpha, \beta, \gamma \in \mathbb{R}$

Therefore this **is not a spanning set for $\mathbb{P}_2(\mathbb{R})$**

(iii) Since the set is neither linearly independent nor spans $\mathbb{P}_2(\mathbb{R})$, **it is not a basis.**

Exercise 4.1.3.3b

Consider the set $\{x^2-2x, x-3, x^2-6, 3x+1\}$

(i) Is this set linearly independent?

(ii) Is this set a spanning set of $\mathbb{P}_2(\mathbb{R})$?

(iii) Is this set a basis of $\mathbb{P}_2(\mathbb{R})$?

Answer

(i) We set up $a(x^2-2x)+b(x-3)+c(x^2-6)+d(3x+1)=0x^2+0x+0$

This is equivalent to $(a+c)x^2+(-2a+b+3d)x+(-3b-6c+d)=0x^2+0x+0 \longleftrightarrow \begin{cases} a+c=0 \\ -2a+b+3d=0 \\ -3b-6c+d=0 \end{cases}$

We solve this homogeneous system by row reducing: $\begin{bmatrix} 1 & 0 & 1 & 0 \\ -2 & 1 & 0 & 3 \\ 0 & -3 & -6 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Since there is a free variable, the set **is linearly dependent.**

(ii) $a(x^2-2x)+b(x-3)+c(x^2-6)+d(3x+1)=\alpha x^2+\beta x+\gamma$

This is equivalent to $(a+c)x^2+(-2a+b+3d)x+(-3b-6c+d)=\alpha x^2+\beta x+\gamma \leftrightarrow \begin{cases} a+c=\alpha \\ -2a+b+3d=\beta \\ -3b-6c+d=\gamma \end{cases}$

Since we saw in part (i) that the RREF of the matrix has a pivot in every row, so this system will be consistent for all possible $\alpha, \beta, \gamma \in \mathbb{R}$

Therefore this **is a spanning set for $\mathbb{P}_2(\mathbb{R})$**

(iii) Since the set is not linearly independent, **it is not a basis.**

Exercise 4.1.3.3c

Consider the set $\{3x^2+x, x+3, x^2+1\}$

(i) Is this set linearly independent?

(ii) Is this set a spanning set of $\mathbb{P}_2(\mathbb{R})$?

(iii) Is this set a basis of $\mathbb{P}_2(\mathbb{R})$?

Answer

(i) We set up $a(3x^2+x)+b(x+3)+c(x^2+1)=0x^2+0x+0$

This is equivalent to $(3a+c)x^2+(a+b)x+(3b+c)=0x^2+0x+0 \leftrightarrow \begin{cases} 3a+c=0 \\ a+b=0 \\ 3b+c=0 \end{cases}$

We solve this homogeneous system by row reducing: $\begin{bmatrix} 3 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 3 & 1 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Since there are no free variables, the set **is linearly independent.**

(ii) $a(3x^2+x)+b(x+3)+c(x^2+1)=\alpha x^2+\beta x+\gamma$

This is equivalent to $(3a+c)x^2+(a+b)x+(3b+c)=\alpha x^2+\beta x+\gamma \leftrightarrow \begin{cases} 3a+c=\alpha \\ a+b=\beta \\ 3b+c=\gamma \end{cases}$

Since we saw in part (i) that the RREF of the matrix has a pivot in every row, so this system will be consistent for all $\alpha, \beta, \gamma \in \mathbb{R}$

Therefore this **is a spanning set for $\mathbb{P}_2(\mathbb{R})$**

(iii) Since the set is linearly independent and spans $\mathbb{P}_2(\mathbb{R})$, **it is a basis.**

Exercise 4.1.3.4a

Prove or disprove the subset $A=\{p(x)\in\mathbb{P}_3(\mathbb{R}) : p''(0)=p(0)\}$ is a subspace.

If it is a subspace, find its dimension.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{P}_3(\mathbb{R})$ is $z(x)=0x^3+0x^2+0x+0$

We can see that $z(0)=0$ and $z''(0)=0$ so $z(x)\in A$.

(ii) If p and q are in A, then we know they are at most cubic polynomials that satisfy $p(0)=p''(0)$ and $q(0)=q''(0)$

If r, s are any real numbers, then the function $(rp+sq)(x)$ is also an at most cubic polynomial and

$(rp+sq)''(0)=(rp)''(0)+(sq)''(0)=r\cdot p''(0)+s\cdot q''(0)=r\cdot p(0)+s\cdot q(0)=(rp)(0)+(sq)(0)=(rp+sq)(0)$ so $rp+sq\in A$

Therefore A **is a subspace** of $\mathbb{P}_3(\mathbb{R})$.

To find its dimension, we construct a basis. Notice that if $p(x) = ax^3 + bx^2 + cx + d$, then $p''(0) = 2b$ and $p(0) = d$

Therefore $A = \{ax^3 + bx^2 + cx + d : 2b = d\} = \{ax^3 + bx^2 + cx + 2b\} = \{a(x^3) + b(x^2 + 2) + c(x)\}$

Thus we know $A = \text{Span}\{x^3, x^2 + 2, x\}$. This spanning set is clearly linearly independent, so it is a basis of A

Therefore we have $\dim(A) = 3$

? Exercise 4.1.3.4*b*

Prove or disprove the subset $B = \{p(x) \in \mathbb{P}_4(\mathbb{R}) : p(x) = p(-x)\}$ is a subspace.

If it is a subspace, find its dimension.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{P}_4(\mathbb{R})$ is $z(x) = 0x^4 + 0x^3 + 0x^2 + 0x + 0$

We can see that $z(-x) = 0(-x)^4 + 0(-x)^3 + 0(-x)^2 + 0(-x) + 0 = 0x^4 + 0x^3 + 0x^2 + 0x + 0 = z(x)$ so $z(x) \in B$.

(ii) If p and q are in B, then we know they are at most quartic polynomials that satisfy $p(x) = p(-x)$ and $q(x) = q(-x)$

If r, s are any real numbers, then the function $(rp + sq)(x)$ is also an at most quartic polynomial and

$(rp + sq)(-x) = (rp)(-x) + (sq)(-x) = r \cdot p(-x) + s \cdot q(-x) = r \cdot p(x) + s \cdot q(x) = (rp)(x) + (sq)(x) = (rp + sq)(x)$ so $rp + sq \in B$

Therefore B **is a subspace** of $\mathbb{P}_4(\mathbb{R})$.

To find its dimension, we construct a basis. Notice that if $p(x) = ax^4 + bx^3 + cx^2 + dx + e$, then $p(-x) = ax^4 - bx^3 + cx^2 - dx + e$

Therefore $B = \{ax^4 + bx^3 + cx^2 + dx + e : ax^4 + bx^3 + cx^2 + dx + e = ax^4 - bx^3 + cx^2 - dx + e\}$

$= \{ax^4 + bx^3 + cx^2 + dx + e : b = -b, d = -d\} = \{ax^4 + bx^3 + cx^2 + dx + e : b = d = 0\}$

$= \{ax^4 + 0x^3 + cx^2 + 0x + e\} = \{a(x^4) + c(x^2) + e(1)\}$

Thus we know $B = \text{Span}\{x^4, x^2, 1\}$. This spanning set is clearly linearly independent, so it is a basis of B

Therefore we have $\dim(B) = 3$

? Exercise 4.1.3.4*c*

Prove or disprove the subset $C = \{p(x) \in \mathbb{P}_2(\mathbb{R}) : p(\pi) = 1\}$ is a subspace.

If it is a subspace, find its dimension.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{P}_2(\mathbb{R})$ is $z(x) = 0x^2 + 0x + 0$

We can see that $z(\pi) = 0 \cdot \pi^2 + 0 \cdot \pi + 0 = 0 \neq 1$ so $z(x) \notin C$.

Therefore C **is not a subspace** of $\mathbb{P}_2(\mathbb{R})$.

Exercise 4.1.3.4d

Prove or disprove the subset $D = \{M \in \mathbb{M}_{2\times 2}(\mathbb{R}) : M^T = M\}$ is a subspace.

If it is a subspace, find its dimension.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{M}_{2\times 2}(\mathbb{R})$ is $Z = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

$Z^T = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = Z$ so $Z \in D$.

(ii) If A and B are in D, then we know they are 2×2 matrices that satisfy $A^T = A$ and $B^T = B$

If r, s are any real numbers, then the matrix $rA + sB$ is also 2×2 and

$(rA + sB)^T = (rA)^T + (sB)^T = r \cdot A^T + s \cdot B^T = r \cdot A + s \cdot B = rA + sB$ so $rA + sB \in D$

Therefore D **is a subspace** of $\mathbb{M}_{2\times 2}(\mathbb{R})$.

To find its dimension, we construct a basis. Notice that if $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then $M^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$

Therefore $D = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \right\}$

$= \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : b = c \right\} = \left\{ \begin{bmatrix} a & c \\ c & d \end{bmatrix} \right\} = \left\{ a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

Thus we know $D = \text{Span}\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$. This spanning set is clearly linearly independent, so it is a basis of D

Therefore we have $\dim(D) = 3$

Exercise 4.1.3.4e

Prove or disprove the subset $E = \{M \in \mathbb{M}_{2\times 2}(\mathbb{R}) : \det(M) = 0\}$ is a subspace.

If it is a subspace, find its dimension.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{M}_{2\times 2}(\mathbb{R})$ is $Z = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

$\det(Z) = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix} = 0 \cdot 0 - 0 \cdot 0 = 0$ so $Z \in E$.

(ii) If A and B are in E, then we know they are 2×2 matrices that satisfy $\det(A) = 0$ and $\det(B) = 0$

If r, s are any real numbers, then the matrix $rA + sB$ is also 2×2 but $\det(rA + sB)$ does not have to be zero.

Consider $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, each has determinant zero so $A, B \in E$,

but $1A + 1B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$ and $\det(I) = 1$, so $1A + 1B \notin E$

Therefore E **is not a subspace** of $\mathbb{M}_{2\times 2}(\mathbb{R})$.

Exercise 4.1.3.4f

Prove or disprove the subset $F = \left\{ M \in \mathbb{M}_{2\times 3}(\mathbb{R}) : M \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right\}$ is a subspace.

If it is a subspace, find its dimension.

Answer

Define $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$. We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{M}_{2\times 3}(\mathbb{R})$ is $Z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$ZC = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ so $Z \in F$.

(ii) If A and B are in F, then we know they are 2×3 matrices that satisfy $AC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $BC = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

If r, s are any real numbers, then the matrix $rA + sB$ is also 2×3 and

$(rA + sB)C = (rA)C + (sB)C = r(AC) + s(BC) = r \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + s \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, so $rA + sB \in F$,

Therefore F **is a subspace** of $\mathbb{M}_{2\times 3}(\mathbb{R})$.

To find its dimension, we construct a basis. Notice that if $M = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$, then $MC = \begin{bmatrix} a+3b+5c & 2a+4b+6c \\ d+3e+5f & 2d+4e+6f \end{bmatrix}$

Therefore $F = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} : \begin{bmatrix} a+3b+5c & 2a+4b+6c \\ d+3e+5f & 2d+4e+6f \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \right\}$

We can solve this homogeneous system with the matrix $\begin{bmatrix} 1 & 3 & 5 & 0 & 0 & 0 \\ 2 & 4 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 2 & 4 & 6 \end{bmatrix} \xrightarrow{RREF} \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{bmatrix}$

So $F = \left\{ \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} : a = c, b = -2c, d = f, e = -2f \right\} = \left\{ \begin{bmatrix} c & -2c & c \\ f & -2f & f \end{bmatrix} \right\} = \left\{ c \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} + f \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \end{bmatrix} \right\}$

Thus we know $F = \text{Span} \left\{ \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \end{bmatrix} \right\}$. This spanning set is clearly linearly independent, so it is a basis of F

Therefore we have $\dim(F) = 2$

Exercise 4.1.3.5a

Prove or disprove the subset $A = \left\{ (a_n) \in \mathbb{S} : \lim_{n\to\infty} a_n = 1 \right\}$ is a subspace.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{S}$ is $(z_n) = (0, 0, 0, 0, \ldots)$

Notice $\lim_{n\to\infty} z_n = 0 \neq 1$ so $(z_n) \notin A$.

Therefore A **is not a subspace** of $\mathbb{S}$.

? Exercise 4.1.3.5b

Prove or disprove the subset $B = \left\{(a_n) \in \mathbb{S} : \lim_{n\to\infty} a_n = 0\right\}$ is a subspace.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{S}$ is $(z_n) = (0, 0, 0, 0, ...)$

We see $\lim_{n\to\infty} z_n = 0$ so $(z_n) \in B$.

(ii) If (a_n) and (b_n) are in B, then we know they are sequences that satisfy $\lim_{n\to\infty} a_n = 0$ and $\lim_{n\to\infty} b_n = 0$

If r, s are any real numbers, then the sequence $ra_n + sb_n$ will have

$\lim_{n\to\infty}(ra_n + sb_n) = \left(\lim_{n\to\infty} ra_n\right) + \left(\lim_{n\to\infty} sb_n\right) = r\left(\lim_{n\to\infty} a_n\right) + s\left(\lim_{n\to\infty} b_n\right) = r \cdot 0 + s \cdot 0 = 0$, so $ra_n + sb_n \in B$,

Therefore B **is a subspace** of $\mathbb{S}$.

? Exercise 4.1.3.5c

Prove or disprove the subset $C = \left\{(a_n) \in \mathbb{S} : \lim_{n\to\infty} a_n \text{ diverges}\right\}$ is a subspace.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{S}$ is $(z_n) = (0, 0, 0, 0, ...)$

We see $\lim_{n\to\infty} z_n = 0$ so $(z_n) \notin C$.

Therefore C **is not a subspace** of $\mathbb{S}$.

? Exercise 4.1.3.5d

Prove or disprove the subset $D = \left\{(a_n) \in \mathbb{S} : \sum_{n=1}^{\infty} a_n \text{ converges}\right\}$ is a subspace.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{S}$ is $(z_n) = (0, 0, 0, 0, ...)$

We see $\sum_{n=1}^{\infty} z_n = \sum_{n=1}^{\infty} 0 = 0$ so $(z_n) \in D$.

(ii) If (a_n) and (b_n) are in D, then we know they are sequences that satisfy $\sum_{n=1}^{\infty} a_n = \alpha$ and $\sum_{n=1}^{\infty} b_n = \beta$ for some $\alpha, \beta \in \mathbb{R}$

If r, s are any real numbers, then the sequence $ra_n + sb_n$ will have

$\sum_{n=1}^{\infty}(ra_n + sb_n) = \left(\sum_{n=1}^{\infty} ra_n\right) + \left(\sum_{n=1}^{\infty} sb_n\right) = r\left(\sum_{n=1}^{\infty} a_n\right) + s\left(\sum_{n=1}^{\infty} b_n\right) = r\alpha + s\beta \in \mathbb{R}$, so $ra_n + sb_n \in D$,

Therefore D **is a subspace** of $\mathbb{S}$.

? Exercise 4.1.3.5e

Prove or disprove the subset $E = \left\{f(x) \in \mathbb{F}(\mathbb{R}, \mathbb{R}) : \int_0^{2\pi} f(x)dx = 0\right\}$ is a subspace.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $\mathbb{F}(\mathbb{R},\mathbb{R})$ is the constant function $z(x)=0$

We see $\int_0^{2\pi} z(x)dx = 0$ so $z(x) \in E$.

(ii) If f and g are in E, then we know they are functions that satisfy $\int_0^{2\pi} f(x)dx = 0$ and $\int_0^{2\pi} g(x)dx = 0$

If r, s are any real numbers, then the function $rf+sg$ will have

$\int_0^{2\pi} (rf+sg)(x)dx = \int_0^{2\pi} (r\cdot f(x)+s\cdot g(x))\,dx = r\cdot\int_0^{2\pi} f(x)dx + s\cdot\int_0^{2\pi} g(x)dx = r\cdot 0 + s\cdot 0 = 0$, so $rf+sg \in E$,

Therefore E **is a subspace** of $\mathbb{F}(\mathbb{R},\mathbb{R})$.

? Exercise 4.1.3.5f

Prove or disprove the subset $F = \{f(x) \in C'(\mathbb{R},\mathbb{R}) : f'(x) = f(x+1)\}$ is a subspace.

Answer

We check the requirements of being a subspace:

(i) The zero vector of $C'(\mathbb{R},\mathbb{R})$ which is a subspace of $\mathbb{F}(\mathbb{R},\mathbb{R})$ is the constant function $z(x)=0$

We see $z'(x) = 0 = z(x+1)$ so $z(x) \in F$.

(ii) If f and g are in F, then we know they are functions that satisfy $f'(x) = f(x+1)$ and $g'(x) = g(x+1)$

If r, s are any real numbers, then the function $rf+sg$ will have

$(rf+sg)'(x) = (rf)'(x) + (sg)'(x) = r\cdot f'(x) + s\cdot g'(x) = r\cdot f(x+1) + s\cdot g(x+1) = (rf+sg)(x+1)$, so $rf+sg \in F$,

Therefore F **is a subspace** of $C'(\mathbb{R},\mathbb{R})$

4.1.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.2: Linear Transformations in General Vector Spaces

Learning Objectives

1. Understand the definition of a linear transformation in the context of vector spaces.
2. Recognize when a transformation between vector spaces is linear, and when objects are in their kernel or range.
3. *Recipes:* Find the matrix representation with respect to a basis, kernel, and range of a linear transformation.
4. *Vocabulary Words:* **kernel, range, zero transformation, identity transformation**

Recall that a function is simply a transformation of a vector to result in a new vector. Consider the following definition.

Definition: Linear Transformation

A ***linear transformation*** is a transformation $T: V \to W$ satisfying

$$T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$$
$$T(c\vec{x}) = cT(\vec{x})$$

for all vectors $\vec{x}, \vec{y}$ in V and all scalars c.

This is the same definition we saw for linear transformations $T: \mathbb{R}^n \to \mathbb{R}^m$, but generalized.

This page titled 4.2: Linear Transformations in General Vector Spaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber via source content that was edited to the style and standards of the LibreTexts platform.

- **9.6: Linear Transformations** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.
- **3.3: Linear Transformations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

4.2.1: Linear Transformations

Several important examples of linear transformations include the zero transformation, the identity transformation, and the scalar transformation.

Illustration 4.2.1.1: Linear Transformations

Let V and W be vector spaces.

1. **The zero transformation:**
   $\mathbf{0} : V \to W$ is defined by $\mathbf{0}(\vec{v}) = \vec{0}$ for all $\vec{v} \in V$
2. **The identity transformation:**
   $\mathbf{1}_V : V \to V$ is defined by $\mathbf{1}_V(\vec{v}) = \vec{v}$ for all $\vec{v} \in V$
3. **The scalar transformation:**
   Let $a \in \mathbb{R}$, then $s_a : V \to V$ is defined by $s_a(\vec{v}) = a\vec{v}$ for all $\vec{v} \in V$

We will show that each transformation above is linear.

For the zero transformation:

$$\mathbf{0}(\vec{x} + \vec{y}) = \vec{0} = \vec{0} + \vec{0} = \mathbf{0}(\vec{x}) + \mathbf{0}(\vec{y})$$

$$\mathbf{0}(c\vec{x}) = \vec{0} = c\vec{0} = c\mathbf{0}(\vec{x})$$

Clearly it is linear.

For the identity transformation:

$$\mathbf{1}_V(\vec{x} + \vec{y}) = \vec{x} + \vec{y} = \mathbf{1}_V(\vec{x}) + \mathbf{1}_V(\vec{y})$$

$$\mathbf{1}_V(c\vec{x}) = c\vec{x} = c\mathbf{1}_V(\vec{x})$$

Again this is linear.

Lastly for the scalar transformation:

$$s_a(\vec{x} + \vec{y}) = a(\vec{x} + \vec{y}) = a\vec{x} + a\vec{y} = s_a(\vec{x}) + s_a(\vec{y})$$

$$s_a(c\vec{x}) = a(c\vec{x}) = c(a\vec{x}) = cs_a(\vec{x})$$

Therefore, s_a is a linear transformation.

Note

The same properties of linear transformations still hold. If $T: V \to W$ be a linear transformation. Then:

1. $T(\vec{0}) = \vec{0}$.
2. For any vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ in V and scalars $c_1, c_2, \ldots, c_k$, we have

$$T(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k) = c_1T(\vec{v}_1) + c_2T(\vec{v}_2) + \cdots + c_kT(\vec{v}_k)$$

Let's explore a few popular operations and their linearity.

Example 4.2.1.1: Nonexample

Let $T : \mathbb{M}_{2\times 2}(\mathbb{R}) \to \mathbb{R}$ be the determinant function defined by $T(A) = \det(A)$. Is T a linear transformation?

Solution:

We need to check if $T(A + B) = T(A) + T(B)$ and if $T(cA) = cT(A)$.

Neither of these properties are true! For a counterexample let $A = B = I$:

$T(A) = T(B) = \det(I) = 1$, but $T(A+B) = \det(2I) = 4$

Similarly, $T(cI) = c^2 \det(I) = c^2 \neq c = c\det(I) = cT(I)$

The determinant is **not a linear transformation**, but it is still a useful function.

Example 4.2.1.2: Derivative Example

Let $T : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ be the function defined by $T(p(x)) = p'(x)$. Is T a linear transformation?

Solution:

We need to check if $T(p(x) + q(x)) = T(p(x)) + T(q(x))$ and if $T(cp(x)) = cT(p(x))$.

Anyone who has taken a calculus course knows that these properties are true:

$$T(p(x) + q(x)) = (p(x) + q(x))' = p'(x) + q'(x) = T(p(x)) + T(q(x))$$

$$T(cp(x)) = (cp(x))' = cp'(x) = cT(p(x))$$

The derivative **is a linear transformation**

Note

Note that the domain and codomain did not need to be restricted to polynomials for this transformation to be defined and linear.

If we define $C'([a,b]) = \{f(x) \in \mathbb{F} : f(x) \text{ is differentiable on the interval } [a,b]\}$, we can prove that it is a subspace by once again leaning on our calculus experience: The sum of two differentiable functions must be differentiable, and a scalar multiple of a differentiable function is always differentiable.

So instead, we could have said $T : C'([a,b]) \to \mathbb{F}(\mathbb{R})$ is a linear transformation with the same argument seen above.

Example 4.2.1.3: An Uglier Example

Let $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{M}_{2\times 2}(\mathbb{R})$ be the function defined by $T(p(x)) = \begin{bmatrix} p(1) & p(-1) \\ p'(0) & p'''(2) \end{bmatrix}$. Is T a linear transformation?

Solution:

We need to check if $T(p(x) + q(x)) = T(p(x)) + T(q(x))$ and if $T(cp(x)) = cT(p(x))$.

$$T(p(x)+q(x)) = \begin{bmatrix} (p+q)(1) & (p+q)(-1) \\ (p+q)'(0) & (p+q)'''(2) \end{bmatrix} = \begin{bmatrix} p(1)+q(1) & p(-1)+q(-1) \\ p'(0)+q'(0) & p'''(2)+q'''(2) \end{bmatrix}$$
$$= \begin{bmatrix} p(1) & p(-1) \\ p'(0) & p'''(2) \end{bmatrix} + \begin{bmatrix} q(1) & q(-1) \\ q'(0) & q'''(2) \end{bmatrix} = T(p(x)) + T(q(x))$$

$$T(cp(x)) = \begin{bmatrix} (cp)(1) & (cp)(-1) \\ (cp)'(0) & (cp)'''(2) \end{bmatrix} = \begin{bmatrix} cp(1) & cp(-1) \\ cp'(0) & cp'''(2) \end{bmatrix} = c\begin{bmatrix} p(1) & p(-1) \\ p'(0) & p'''(2) \end{bmatrix} = cT(p(x))$$

Therefore, **T is a linear transformation**

Now that we can recognize if a function between vector spaces is a linear transformation, we will study two familiar sets.

Definition: Kernel and Range

Let V and W be vector spaces and let $T : V \to W$ be a linear transformation. Then the range of T denoted as $\textbf{Range}(T)$ is defined to be the set

$$\textbf{Range}(T) = \{T(\vec{v}) : \vec{v} \in V\}$$

In words, it consists of all vectors in W which equal $T(\vec{v})$ for some $\vec{v} \in V$, just like the standard definition of range.

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 $\text{Range}(T)$ and $\ker(T)$ are subspaces of W and V respectively.

Proposition 4.2.1.1: 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 $\text{Range}(T) \subseteq W$. 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 $\text{Range}(T)$. Then if a, b are scalars,

$$aT(\vec{v}_1) + bT(\vec{v}_2) = T\left(a\vec{v}_1 + b\vec{v}_2\right)$$

and this last vector is in $\text{Range}(T)$ by definition.

Illustration 4.2.1.2: Linear Transformations

Let V and W be vector spaces. We will find the kernel and range of the zero and identity transformations.

For the zero transformation, every vector is mapped to $\vec{0}$ in W so:

$$\ker(0) = \left\{\vec{v} \in V : 0(\vec{v}) = \vec{0}\right\} = \{\vec{v} \in V\}$$

$$\text{Range}(0) = \{0(\vec{v}) : \vec{v} \in V\} = \left\{\vec{0} : \vec{v} \in V\right\} = \{\vec{0}\}$$

For the identity transformation:

$$\ker(1_V) = \left\{\vec{v} \in V : 1_V(\vec{v}) = \vec{0}\right\} = \{\vec{v} \in V : \vec{v} = \vec{0}\} = \{\vec{0}\}$$

$$\text{Range}(1_V) = \{1_V(\vec{v}) : \vec{v} \in V\} = \{\vec{v} : \vec{v} \in V\} = V$$

These two transformations represent the extreme cases for the kernel and range.

Note

Once again we use the terminology **one-to-one** when a linear transformation does not repeat the same output for different input vectors, and **onto** when the linear transformation reaches every possible vector in its codomain. So if $T : V \to W$:

T is one-to-one when $\ker(T) = \{\vec{0}\}$ and T is onto when $\text{Range}(T) = W$

Example 4.2.1.4: Derivative Example

Let $T : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ be the function defined by $T(p(x)) = p'(x)$. Find the kernel and range of T.

Solution:

$$\ker(T) = \left\{p(x) \in \mathbb{P}_3 : T(p(x)) = \vec{0}\right\} = \{p(x) \in \mathbb{P}_3 : p'(x) = 0x^2 + 0x + 0\}$$

Once again, we may recall that only constant functions have derivatives which are identically zero. But in a more general sense, for any polynomial in our domain $p(x) = ax^3 + bx^2 + cx + d$ we have $p'(x) = 3ax^2 + 2bx + c$

$$p'(x) = 0x^2 + 0x + 0 \longrightarrow \begin{cases} 3a = 0 \\ 2b = 0 \\ c = 0 \end{cases} \longrightarrow a = b = c = 0$$

Therefore, we know $\ker(T) = \{ax^3 + bx^2 + cx + d : a = b = c = 0\} = \{0x^3 + 0x^2 + 0x + d : d \in \mathbb{R}\}$ is indeed the subspace of $\mathbb{P}_3$ containing all constant functions.

$$\begin{aligned} \text{Range}(T) = \{T(p(x)) : p(x) \in \mathbb{P}_3\} &= \left\{3ax^2 + 2bx + c : ax^3 + bx^2 + cx + d \in \mathbb{P}_3\right\} \\ &= \{a(3x^2) + b(2x) + c(1) : a, b, c \in \mathbb{R}\} \end{aligned}$$

This looks like a span of three linearly independent vectors (polynomials) in $\mathbb{P}_2$. Clearly these should span all of $\mathbb{P}_2$, so

$$\text{Range}(T) = \text{Span}\{3x^2, 2x, 1\} = \mathbb{P}_2$$

In this case, $\boxed{T \text{ is not one-to-one, but it is onto}}$

Example 4.2.1.5: An Uglier Example

Let $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{M}_{2\times 2}(\mathbb{R})$ be the linear transformation defined by $T(p(x)) = \begin{bmatrix} p(1) & p(-1) \\ p'(0) & p'''(2) \end{bmatrix}$. Find the kernel and range of T.

Solution:

For any polynomial in our domain $p(x) = ax^2 + bx + c$ we have $\begin{cases} p(1) = a + b + c \\ p(-1) = a - b + c \\ p'(0) = b \\ p'''(2) = 0 \end{cases}$ So

$$T(ax^2 + bx + c) = \begin{bmatrix} a+b+c & a-b+c \\ b & 0 \end{bmatrix}$$

$$\ker(T) = \left\{p(x) \in \mathbb{P}_2 : T(p(x)) = \vec{0}\right\} = \{p(x) \in \mathbb{P}_2 : \begin{bmatrix} a+b+c & a-b+c \\ b & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\}$$

We can set up a system of equations for a, b, c and use matrices to solve it:

$$\begin{cases} a+b+c=0 \\ a-b+c=0 \\ b=0 \end{cases} \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \xrightarrow{RREF} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

In parametric form, this becomes $a = -c, b = 0$. Putting this into our kernel expression above gives:

$$\ker(T) = \left\{ax^2 + bx + c : a = -c, b = 0\right\} = \{-cx^2 + 0x + c : c \in \mathbb{R}\} = \{c(-x^2 + 1) : c \in \mathbb{R}\}$$

Once again this is a span, this time of a single vector: $ker(T) = \text{Span}\{-x^2 + 1\}$

$$\begin{aligned} \text{Range}(T) = \{T(p(x)) : p(x) \in \mathbb{P}_2\} &= \left\{\begin{bmatrix} a+b+c & a-b+c \\ b & 0 \end{bmatrix} : ax^2 + bx + c \in \mathbb{P}_2\right\} \\ &= \left\{a\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} : a, b, c \in \mathbb{R}\right\} \end{aligned}$$

This looks like a span of three linearly dependent vectors (matrices) in $\mathbb{M}_{2\times 2}$. Since the first and last of these vectors is identical, we can omit one of them and still span our entire range:

$$\text{Range}(T) = \text{Span}\left\{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right\} = \text{Span}\left\{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}\right\}$$

In this case, $\boxed{T \text{ is not one-to-one or onto}}$

4.2.1: Linear Transformations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **9.6: Linear Transformations** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.
- **3.3: Linear Transformations** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.
- **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.

4.2.2: Coordinatization

Recall that for any transformation $T:\mathbb{R}^n \to \mathbb{R}^n$ and any basis $\mathcal{B}$ of $\mathbb{R}^n$ we can find a matrix A that satisfies $[T(\vec{x})]_{\mathcal{B}} = A[\vec{x}]_{\mathcal{B}}$.

Similarly, for any $T:V \to W$ we can use a basis of the domain V and a basis of the codomain W to coordinatize each vector in V and replace it with a vector in $\mathbb{R}^n$ and coordinatize each output vector in W and replace it with a vector in $\mathbb{R}^m$. This allows us to once again find a $m \times n$ **matrix** A that emulates the action of T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$.

Proposition 4.2.2.2: Matrix Representation of General Linear Transformations

Let V, W be vector spaces and let $T:V \to W$ be a linear transformation. Assume $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for V and $\mathcal{C} = \{\vec{w}_1, \dots, \vec{w}_m\}$ is a basis for W.

If we define $B_{\mathcal{B}}: V \to \mathbb{R}^n$ and $C_{\mathcal{C}}: W \to \mathbb{R}^m$, to be the coordinatization functions, then the function $T': \mathbb{R}^n \to \mathbb{R}^m$ defined by $T' = C_{\mathcal{C}} \circ T \circ B_{\mathcal{B}}^{-1}$ is a linear transformation.

Proof

First note that $B_{\mathcal{B}}(a_1\vec{v}_1 + \cdots + a_n\vec{v}_n) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$ is clearly a linear transformation with kernel and range:

$$\ker(B_{\mathcal{B}}) = \{\vec{x} \in V : B_{\mathcal{B}}(\vec{x}) = \vec{0}\} = \{\vec{0}\}$$

$$\text{Range}(B_{\mathcal{B}}) = \{B_{\mathcal{B}}(\vec{x}) : \vec{x} \in V\} = \left\{ \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} : a_1, \dots, a_n \in \mathbb{R} \right\} = \mathbb{R}^n$$

Since this transformation is one-to-one and onto, it is invertible with $B_{\mathcal{B}}^{-1}\left(\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \right) = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n$

Similarly, we can see that $C_{\mathcal{C}}$ is invertible with $C_{\mathcal{C}}(a_1\vec{w}_1 + \cdots + a_m\vec{w}_m) = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$

Since T itself is a linear transformation, and the composition of linear transformations must be a linear transformation, we have proven that T' is a linear transformation.

Recipe: Matrix Representation with respect to Bases

Since $T': \mathbb{R}^n \to \mathbb{R}^m$ above is a linear transformation, we know it is a matrix transformation for the matrix

$$[T]_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} | & & | \\ T'(\vec{e}_1) & \cdots & T'(\vec{e}_n) \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ (C_{\mathcal{C}} \circ T \circ B_{\mathcal{B}}^{-1})(\vec{e}_1) & \cdots & (C_{\mathcal{C}} \circ T \circ B_{\mathcal{B}}^{-1})(\vec{e}_n) \\ | & & | \end{bmatrix}$$

We have a straightforward definition of $B_{\mathcal{B}}^{-1}$, so we can easily find $B_{\mathcal{B}}^{-1}(\vec{e}_k)$:

$$B_{\mathcal{B}}^{-1}(\vec{e}_k) = 0\vec{v}_1 + \cdots + 0\vec{v}_{k-1} + 1\vec{v}_k + 0\vec{v}_{k+1} + \cdots + 0\vec{v}_n = \vec{v}_k$$

Therefore our matrix should be

$$\begin{bmatrix} | & & | \\ (C_{\mathcal{C}} \circ T)(\vec{v}_1) & \cdots & (C_{\mathcal{C}} \circ T)(\vec{v}_n) \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ C_{\mathcal{C}}(T(\vec{v}_1)) & \cdots & C_{\mathcal{C}}(T(\vec{v}_n)) \\ | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & & | \\ [T(\vec{v}_1)]_{\mathcal{C}} & \cdots & [T(\vec{v}_n)]_{\mathcal{C}} \\ | & & | \end{bmatrix}$$

So the matrix representation of T with respect to the bases $\mathcal{B}$ and $\mathcal{C}$ is the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ which has columns which are found by plugging in each basis vector of the domain into T and then coordinatizing the output with respect to the basis of the codomain.

Illustration 4.2.2.3: Linear Transformations

Let V be a vector space with basis $\mathcal{B} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ and W be a vector space with basis $\mathcal{C} = \{\vec{w}_1, \ldots, \vec{w}_m\}$.

We will find the matrix representation of the zero and identity transformations with respect to these bases, then find their null space and column spaces.

For the zero transformation, every vector $\vec{x} \in V$ is mapped to $\vec{0}$ in W, so we have

$$0(\vec{v}_i) = \vec{0} = 0\vec{w}_1 + \cdots + 0\vec{w}_m \implies [0(\vec{v}_i)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{0}$$

This is the same for basis vector of $\mathcal{B}$, so the matrix representation is

$$[0]_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

This is (unsurprisingly) the $m \times n$ matrix filled with zeros, known as the zero matrix.

The null space of this matrix is

$$\text{Null}\left([0]_{\mathcal{B}}^{\mathcal{C}}\right) = \text{Null}\left(\begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}\right) = \mathbb{R}^n$$

The column space is

$$\text{Col}\left([0]_{\mathcal{B}}^{\mathcal{C}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}\right\} = \{\vec{0}\}$$

For the identity transformation, every vector $\vec{x} \in V$is mapped to itself $\vec{x}$ in V, so we have

$$1_V(\vec{v}_i) = \vec{v}_i = 0\vec{v}_1 + \cdots + 0\vec{v}_{i-1} + 1\vec{v}_i + 0\vec{v}_{i+1} + \cdots + 0\vec{v}_n \implies [1_V(\vec{v}_i)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \vec{e}_i$$

This is (again unsurprisingly) the $n \times n$ matrix whose i^{th} column is $\vec{e}_i$, which is the standard identity matrix.

The null space of this matrix is

$$\text{Null}\left([1_V]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Null}\,(I) = \{\vec{0}\}$$

The column space is

$$\text{Col}\left([1_V]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\,\{\vec{e}_1, \ldots, \vec{e}_n\} = \mathbb{R}^n$$

Note

Notice that just like the kernel and range, the null and column spaces of these matrix representations are the zero space or $\mathbb{R}^n$.

In fact, these spaces are analogous. The only difference is that the null and column spaces above are the **coordinatized** version of the kernel and range of their respective linear transformation. This will always hold!

Example 4.2.2.6: Derivative Example

Let $T : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ be the function defined by $T(p(x)) = p'(x)$.

(a) If $\mathcal{B} = \{x^3, x^2, x, 1\}$ and $\mathcal{C} = \{x^2, x, 1\}$, find $[T]_{\mathcal{B}}^{\mathcal{C}}$ and its null and column space.

(b) If $\mathcal{D} = \{x^3 - x^2, x^3 + x^2, x - 1, x + 1\}$ and $\mathcal{E} = \{1, x, x^2\}$, find $[T]_{\mathcal{D}}^{\mathcal{E}}$ and its null and column space.

Solution:

(a) We plug in each basis vector from $\mathcal{B}$ and coordinatize the output with respect to $\mathcal{C}$:

$$T(x^3) = (x^3)' = 3x^2 = 3 \cdot x^2 + 0 \cdot x + 0 \cdot 1 \implies [T(x^3)]_{\mathcal{C}} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

$$T(x^2) = (x^2)' = 2x = 0 \cdot x^2 + 2 \cdot x + 0 \cdot 1 \implies [T(x^2)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}$$

$$T(x) = (x)' = 1 = 0 \cdot x^2 + 0 \cdot x + 1 \cdot 1 \implies [T(x)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$T(1) = (1)' = 0 = 0 \cdot x^2 + 0 \cdot x + 0 \cdot 1 \implies [T(1)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Therefore we have matrix representation $[T]_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, which has RREF $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

The column space of this matrix is

$$\boxed{\text{Col}\left([T]_{\mathcal{B}}^{\mathcal{C}}\right) = \text{Span}\left\{\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\} = \mathbb{R}^3}$$

Solving for the null space gives $a = 0, b = 0, c = 0$ and d is free. Thus,

$$\boxed{\text{Null}\left([T]_{\mathcal{B}}^{\mathcal{C}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}}$$

(b) We plug in each basis vector from $\mathcal{D}$ and coordinatize the output with respect to $\mathcal{E}$:

$$T(x^3 - x^2) = (x^3 - x^2)' = 3x^2 - 2x = 0 \cdot 1 - 2 \cdot x + 3 \cdot x^2 \implies [T(x^3 - x^2)]_{\mathcal{E}} = \begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix}$$

$$T(x^3 + x^2) = (x^3 + x^2)' = 3x^2 + 2x = 0 \cdot 1 + 2 \cdot x + 3 \cdot x^2 \implies [T(x^3 + x^2)]_{\mathcal{E}} = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$$

$$T(x - 1) = (x - 1)' = 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \implies [T(x - 1)]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$T(x+1) = (x+1)' = 1 = 1\cdot 1 + 0\cdot x + 0\cdot x^2 \implies [T(x+1)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Therefore we have matrix representation $[T]_{\mathcal{D}}^{\mathcal{E}} = \begin{bmatrix} 0 & 0 & 1 & 1 \\ -2 & 2 & 0 & 0 \\ 3 & 3 & 0 & 0 \end{bmatrix}$, which has RREF $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

The column space of this matrix is

$$\boxed{\text{Col}\left([T]_{\mathcal{D}}^{\mathcal{E}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\} = \mathbb{R}^3}$$

Solving for the null space gives $a = 0, b = 0, c = -d$ and d is free. Thus,

$$\boxed{\text{Null}\left([T]_{\mathcal{D}}^{\mathcal{E}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}\right\}}$$

Note

Note that each of these matrices had a column space that spanned all of $\mathbb{R}^3$. This means that all possible combinations of $\mathcal{C}$ or $\mathcal{E}$ coordinates can be reached in the span of their columns. This is equivalent to saying that all possible vectors in $\mathbb{P}_2$, the codomain of T, are in its range. This checks out with our earlier work, where we showed $\text{Range}(T) = \mathbb{P}_2$.

On the other hand, each of the null spaces of these matrices are different looking at first:

Notice that all vectors with $\mathcal{B}$ coordinates in the span of $\text{Null}\left([T]_{\mathcal{B}}^{\mathcal{C}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$ is represented by $\text{Span}\left\{0\cdot x^3 + 0\cdot x^2 + 0\cdot x + 1\cdot 1\right\} = \text{Span}\left\{1\right\}$.

Similarly, all of the vectors with $\mathcal{D}$ coordinates in the span of $\text{Null}\left([T]_{\mathcal{D}}^{\mathcal{E}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}\right\}$ is represented by $\text{Span}\left\{0\cdot(x^3 - x^2) + 0\cdot(x^3 + x^2) - 1\cdot(x-1) + 1\cdot(x+1)\right\} = \text{Span}\left\{2\right\}$.

These both represent our $\ker(T)$, just coordinatized with respect to the basis we used in the domain $\mathbb{P}_3$.

Example 4.2.2.7: An Uglier Example

Let $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{M}_{2\times 2}(\mathbb{R})$ be the linear transformation defined by $T(p(x)) = \begin{bmatrix} p(1) & p(-1) \\ p'(0) & p'''(2) \end{bmatrix}$.

If $\mathcal{B} = \{x^2, x, 1\}$ and $\mathcal{C} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$, find $[T]_{\mathcal{B}}^{\mathcal{C}}$ and its null and column space.

Solution:

We plug in each vector from $\mathcal{B}$ and coordinatize the output with respect to $\mathcal{C}$:

$$T(x^2) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \implies [T(x^2)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$T(x) = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} \implies [T(x)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

$$T(1) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \implies [T(1)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Therefore we have matrix representation $[T]_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, which has RREF $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

The column space of this matrix is

$$\text{Col}\left([T]_{\mathcal{B}}^{\mathcal{C}}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}\right\}$$

Solving for the null space gives $a = -c, b = 0$, and c is free. Thus,

$$\text{Null}\left([T]_{\mathcal{B}}^{\mathcal{C}}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$$

Once again, we can see that the column space of $[T]_{\mathcal{B}}^{\mathcal{C}}$ is $\text{Range}(T)$ coordinatized with respect to $\mathcal{C}$:

$$\boxed{\text{Range}(T) = \text{Span}\left\{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}\right\}}$$

And that the null space of $[T]_{\mathcal{B}}^{\mathcal{C}}$ is $\ker(T)$ coordinatized with respect to $\mathcal{B}$:

$$\boxed{\ker(T) = \text{Span}\left\{-1 \cdot x^2 + 0 \cdot x + 1 \cdot 1\right\} = \text{Span}\left\{-x^2 + 1\right\}}$$

Now that we can effectively translate linear transformations between finite dimensional vector spaces into a matrix, we can use them to help find eigenvalues, eigenvectors, and eigenspaces of linear transformations of the type $T : V \to V$.

Example 4.2.2.8: Eigenspaces of a Linear Transformation

Let $T : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ be the linear transformation defined by $T(p(x)) = (x-1)p'(x) + p(0)$.

Find all eigenspaces with non-zero dimension.

Solution:

We begin by picking any basis, such as $\mathcal{B} = \{x^2, x, 1\}$, and find $[T]_{\mathcal{B}}^{\mathcal{B}}$:

$$T(x^2) = (x-1) \cdot 2x + 0 = 2x^2 - 2x \implies [T(x^2)]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}$$

$$T(x) = (x-1) \cdot 1 + 0 = x - 1 \implies [T(x)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$$T(1) = (x-1) \cdot 0 + 1 = 1 \implies [T(1)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Therefore we have matrix representation $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 2 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$. We now look for the eigenvalues by finding the characteristic polynomial:

$$f(\lambda) = \begin{vmatrix} 2-\lambda & 0 & 0 \\ -2 & 1-\lambda & 0 \\ 0 & -1 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda)^2$$

Clearly, we have two eigenvalues: $\lambda_1 = 1$ and $\lambda_2 = 2$. We now find the eigenspaces **of the matrix** $[T]_{\mathcal{B}}^{\mathcal{B}}$:

$$E_1 = \text{Null}\left([T]_{\mathcal{B}}^{\mathcal{B}} - I\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ -2 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$$

$$E_2 = \text{Null}\left([T]_{\mathcal{B}}^{\mathcal{B}} - 2I\right) = \text{Null}\left(\begin{bmatrix} 0 & 0 & 0 \\ -2 & -1 & 0 \\ 0 & -1 & -1 \end{bmatrix}\right)$$
$$= \text{Null}\left(\begin{bmatrix} 1 & 0 & -\frac{1}{2} \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} \frac{1}{2} \\ -1 \\ 1 \end{bmatrix}\right\}$$

These eigenspaces are the coordinatized versions of the eigenspaces of T with respect to our chosen basis $\mathcal{B}$:

$$\boxed{E_1 = \text{Span}\left\{0 \cdot x^2 + 0 \cdot x + 1 \cdot 1\right\} = \text{Span}\left\{1\right\}}$$

$$\boxed{E_2 = \text{Span}\left\{\frac{1}{2} \cdot x^2 - 1 \cdot x + 1 \cdot 1\right\} = \text{Span}\left\{\frac{1}{2}x^2 - x + 1\right\}}$$

It is easy to verify that these are indeed our eigenvalues and eigenvectors by plugging them into T directly:

When $p(x) = c$ is constant (in the span of the polynomial 1), we get: $T(p(x)) = (x-1) \cdot 0 + c = c = 1 \cdot p(x)$

When $p(x) = c\left(\frac{1}{2}x^2 - x + 1\right)$, we get: $T(p(x)) = (x-1) \cdot (2cx - c) + c = c(x^2 - 2x + 2) = 2 \cdot p(x)$

4.2.2: Coordinatization is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

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

4.2.3: Practice Set

Exercise 4.2.3.1

Consider the function $f : \mathbb{M}_{2\times 2}(\mathbb{R}) \to \mathbb{M}_{2\times 2}(\mathbb{R})$ satisfying $f(A) = A + A^T$

(a) Prove f is a linear transformation.

(b) Find a basis for $\ker(f)$. Is f one-to-one?

(c) Find a basis for $\text{Range}(f)$. Is f onto?

(d) Find all eigenvalues and eigenspaces of f

Answer

(a) We check the two properties of linear transformations:

$f(A+B) = (A+B) + (A+B)^T = A + B + A^T + B^T = (A + A^T) + (B + B^T) = f(A) + f(B)$

$f(cA) = (cA) + (cA)^T = cA + cA^T = c(A + A^T) = cf(A)$

This proves f is a linear transformation.

To facilitate the next three parts, we let $\mathcal{B} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ and find $[f]_{\mathcal{B}}^{\mathcal{B}}$

$$f\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$f\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}^T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$f\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$f\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}$$

Therefore, we have $[f]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ which has RREF $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(b) To find $\ker(f)$, we solve for $\text{Null}\left([f]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}\right\}$

This means that $\ker(f) = \text{Span}\left\{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right\}$

This contains nonzero vectors, we know that f **is not one-to-one.**

(c) To find $\text{Range}(f)$, we solve for $\text{Col}\left([f]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}\right\}$

This means that $\text{Range}(f) = \text{Span}\left\{\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}\right\}$

This has dimension 3, and the codomain $\mathbb{M}_{2\times 2}(\mathbb{R})$ has dimension 4, we know that f **is not onto.**

(d) The characteristic polynomial is $p(\lambda) = \det\left([f]_{\mathcal{B}}^{\mathcal{B}} - \lambda I\right) = \lambda(\lambda-2)^3$

Thus, the eigenvalues of $[f]_{\mathcal{B}}^{\mathcal{B}}$ (and also of f) are $\lambda = 0, 2$

The eigenspaces of $[f]_{\mathcal{B}}^{\mathcal{B}}$ are:

$E_0 = \text{Null}\left([f]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}\right\}$ has already been found.

$$E_2 = \text{Null}\left([f]_{\mathcal{B}}^{\mathcal{B}} - 2I\right) = \text{Null}\left(\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$$

Which means the eigenspaces of f will be:

$E_0 = \text{Span}\left\{\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right\}$

$E_2 = \text{Span}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

? Exercise 4.2.3.2

Consider the function $T : \mathbb{M}_{2\times 2}(\mathbb{R}) \to \mathbb{M}_{2\times 2}(\mathbb{R})$ satisfying $T(A) = \begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix} A$

(a) Prove T is a linear transformation.

(b) Find a basis for $\ker(T)$. Is T one-to-one?

(c) Find a basis for $\text{Range}(T)$. Is T onto?

(d) Find all eigenvalues and eigenspaces of T

Answer

(a) We check the two properties of linear transformations:

$$T(A+B) = \begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix}(A+B) = \begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix} A + \begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix} B = T(A) + T(B)$$

$$T(cA) = \begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix}(cA) = c\left(\begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix} A\right) = cT(A)$$

This proves T is a linear transformation.

To facilitate the next three parts, we let $\mathcal{B} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ and find $[T]_{\mathcal{B}}^{\mathcal{B}}$

$$T\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -2 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 \\ 0 \\ -2 \\ 0 \end{bmatrix}$$

$$T\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & -3 \\ -2 & 6 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -2 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \\ -2 \end{bmatrix}$$

$$T\left(\begin{bmatrix}0&0\\1&0\end{bmatrix}\right)=\begin{bmatrix}1&-3\\-2&6\end{bmatrix}\begin{bmatrix}0&0\\1&0\end{bmatrix}=\begin{bmatrix}-3&0\\6&0\end{bmatrix}\longrightarrow\begin{bmatrix}-3\\0\\6\\0\end{bmatrix}$$

$$T\left(\begin{bmatrix}0&0\\0&1\end{bmatrix}\right)=\begin{bmatrix}1&-3\\-2&6\end{bmatrix}\begin{bmatrix}0&0\\0&1\end{bmatrix}=\begin{bmatrix}0&-3\\0&6\end{bmatrix}\longrightarrow\begin{bmatrix}0\\-3\\0\\6\end{bmatrix}$$

Therefore, we have $[T]_{\mathcal{B}}^{\mathcal{B}}=\begin{bmatrix}1&0&-3&0\\0&1&0&-3\\-2&0&6&0\\0&-2&0&6\end{bmatrix}$ which has RREF $\begin{bmatrix}1&0&-3&0\\0&1&0&-3\\0&0&0&0\\0&0&0&0\end{bmatrix}$

(b) To find $\ker(T)$, we solve for $\text{Null}\left([T]_{\mathcal{B}}^{\mathcal{B}}\right)=\text{Span}\left\{\begin{bmatrix}3\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\3\\0\\1\end{bmatrix}\right\}$

This means that $\ker(T)=\text{Span}\left\{\begin{bmatrix}3&0\\1&0\end{bmatrix},\begin{bmatrix}0&3\\0&1\end{bmatrix}\right\}$

This contains nonzero vectors, we know that T **is not one-to-one.**

(c) To find $\text{Range}(T)$, we solve for $\text{Col}\left([T]_{\mathcal{B}}^{\mathcal{B}}\right)=\text{Span}\left\{\begin{bmatrix}1\\0\\-2\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\-2\end{bmatrix}\right\}$

This means that $\text{Range}(T)=\text{Span}\left\{\begin{bmatrix}1&0\\-2&0\end{bmatrix},\begin{bmatrix}0&1\\0&-2\end{bmatrix}\right\}$

This has dimension 2, and the codomain $\mathbb{M}_{2\times2}(\mathbb{R})$ has dimension 4, we know that T **is not onto.**

(d) The characteristic polynomial is $p(\lambda)=\det\left([T]_{\mathcal{B}}^{\mathcal{B}}-\lambda I\right)=\lambda^2(\lambda-7)^2$

Thus, the eigenvalues of $[T]_{\mathcal{B}}^{\mathcal{B}}$ (and also of T) are $\lambda=0,7$

The eigenspaces of $[T]_{\mathcal{B}}^{\mathcal{B}}$ are:

$E_0=\text{Null}\left([T]_{\mathcal{B}}^{\mathcal{B}}\right)=\text{Span}\left\{\begin{bmatrix}3\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\3\\0\\1\end{bmatrix}\right\}$ has already been found.

$$E_7=\text{Null}\left([T]_{\mathcal{B}}^{\mathcal{B}}-7I\right)=\text{Null}\left(\begin{bmatrix}-6&0&-3&0\\0&-6&0&-3\\-2&0&-1&0\\0&-2&0&-1\end{bmatrix}\right)=\text{Span}\left\{\begin{bmatrix}-1\\0\\2\\0\end{bmatrix},\begin{bmatrix}0\\-1\\0\\2\end{bmatrix}\right\}$$

Which means the eigenspaces of T will be:

$E_0=\text{Span}\left\{\begin{bmatrix}3&0\\1&0\end{bmatrix},\begin{bmatrix}0&3\\0&1\end{bmatrix}\right\}$

$E_7=\text{Span}\left\{\begin{bmatrix}-1&0\\2&0\end{bmatrix},\begin{bmatrix}0&-1\\0&2\end{bmatrix}\right\}$

? Exercise 4.2.3.3

Consider the function $U:\mathbb{M}_{2\times2}(\mathbb{R})\to\mathbb{M}_{2\times2}(\mathbb{R})$ satisfying $U\left(\begin{bmatrix}a&b\\c&d\end{bmatrix}\right)=\begin{bmatrix}a+b&b+c\\c+d&2d\end{bmatrix}$

(a) Prove U is a linear transformation.

(b) Find a basis for $\ker(U)$. Is U one-to-one?

(c) Find a basis for $\mathbf{Range}(U)$. Is U onto?

(d) Find all eigenvalues and eigenspaces of U

Answer

(a) We check the two properties of linear transformations:

$$U\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right) = U\left(\begin{bmatrix} a+a' & b+b' \\ c+c' & d+d' \end{bmatrix}\right) = \begin{bmatrix} (a+a')+(b+b') & (b+b')+(c+c') \\ (c+c')+(d+d') & 2(d+d') \end{bmatrix}$$

$$= \begin{bmatrix} a+b & b+c \\ c+d & 2d \end{bmatrix} + \begin{bmatrix} a'+b' & b'+c' \\ c'+d' & 2d' \end{bmatrix} = U\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) + U\left(\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right)$$

$$U\left(r\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = U\left(\begin{bmatrix} ra & rb \\ rc & rd \end{bmatrix}\right) = \begin{bmatrix} ra+rb & rb+rc \\ rc+rd & 2rd \end{bmatrix} = r\begin{bmatrix} a+b & b+c \\ c+d & 2d \end{bmatrix} = rU\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)$$

This proves U is a linear transformation.

To facilitate the next three parts, we let $\mathcal{B} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$ and find $[U]_{\mathcal{B}}^{\mathcal{B}}$

$$U\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$U\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$U\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$U\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

Therefore, we have $[U]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ which has RREF $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

(b) To find $\mathbf{ker}(U)$, we solve for $\mathbf{Null}\left([U]_{\mathcal{B}}^{\mathcal{B}}\right) = \{\vec{0}\}$

This means that $\mathbf{ker}(U) = \left\{\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right\}$

This contains only the zero vector, we know that U **is one-to-one.**

(c) To find $\mathbf{Range}(U)$, we solve for $\mathbf{Col}\left([U]_{\mathcal{B}}^{\mathcal{B}}\right) = \mathbf{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}\right\}$

This means that $\mathbf{Range}(U) = \mathbf{Span}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 2 \end{bmatrix}\right\}$

This has dimension 4, and the codomain $\mathbb{M}_{2\times 2}(\mathbb{R})$ has dimension 4, we know that U **is onto.**

(d) The characteristic polynomial is $p(\lambda) = \det\left([U]_{\mathcal{B}}^{\mathcal{B}} - \lambda I\right) = (\lambda-1)^3(\lambda-2)$

Thus, the eigenvalues of $[U]_{\mathcal{B}}^{\mathcal{B}}$ (and also of U) are $\lambda = 1, 2$

The eigenspaces of $[U]_{\mathcal{B}}^{\mathcal{B}}$ are:

$$E_1 = \text{Null}\left([U]_{\mathcal{B}}^{\mathcal{B}} - I\right) = \text{Null}\left(\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right\}$$

$$E_2 = \text{Null}\left([U]_{\mathcal{B}}^{\mathcal{B}} - 2I\right) = \text{Null}\left(\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right\}$$

Which means the eigenspaces of U will be:

$$E_1 = \text{Span}\left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right\}$$

$$E_2 = \text{Span}\left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$$

Exercise 4.2.3.4

Consider the function $K : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ satisfying $K(p(x)) = p(1)x^2 + p(-1)x + p'(0)$

(a) Prove K is a linear transformation.

(b) Find a basis for $\ker(K)$. Is K one-to-one?

(c) Find a basis for $\text{Range}(K)$. Is K onto?

(d) Find all eigenvalues and eigenspaces of K

Answer

(a) We check the two properties of linear transformations:

$$K((p+q)(x)) = (p+q)(1)x^2 + (p+q)(-1)x + (p+q)'(0) = (p(1)+q(1))x^2 + (p(-1)+q(-1))x + (p'(0)+q'(0))$$

$$= (p(1)x^2 + p(-1)x + p'(0)) + (q(1)x^2 + q(-1)x + q'(0)) = K(p(x)) + K(q(x))$$

$$K((rp)(x)) = (rp)(1)x^2 + (rp)(-1)x + (rp)'(0) = r \cdot p(1)x^2 + r \cdot p(-1)x + r \cdot p'(0) = r \cdot (p(1)x^2 + p(-1)x + p'(0)) = r \cdot K(p(x))$$

This proves K is a linear transformation.

To facilitate the next three parts, we let $\mathcal{B} = \{x^2, x, 1\}$ and find $[K]_{\mathcal{B}}^{\mathcal{B}}$

$$K(x^2) = 1x^2 + 1x + 0 \longrightarrow \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$K(x) = 1x^2 - 1x + 1 \longrightarrow \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$K(1) = 1x^2 + 1x + 0 \longrightarrow \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Therefore, we have $[K]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ which has RREF $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(b) To find $\ker(K)$, we solve for $\text{Null}\left([K]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$

This means that $\ker(K) = \text{Span}\{-x^2 + 1\}$

This contains nonzero vectors, we know that **K is not one-to-one.**

(c) To find $\text{Range}(K)$, we solve for $\text{Col}\left([K]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\-1\\1\end{bmatrix}\right\}$

This means that $\text{Range}(K) = \text{Span}\left\{x^2+x, x^2-x+1\right\}$

This has dimension 2, and the codomain $\mathbb{P}_2(\mathbb{R})$ has dimension 3, we know that K **is not onto.**

(d) The characteristic polynomial is $p(\lambda) = \det\left([K]_{\mathcal{B}}^{\mathcal{B}} - \lambda I\right) = \lambda(\lambda^2 - 3)$

Thus, the eigenvalues of $[K]_{\mathcal{B}}^{\mathcal{B}}$ (and also of K) are $\lambda = 0, \pm\sqrt{3}$

The eigenspaces of $[K]_{\mathcal{B}}^{\mathcal{B}}$ are:

$E_0 = \text{Null}\left([K]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix}-1\\0\\1\end{bmatrix}\right\}$ as we calculated in part (b)

$$E_{\sqrt{3}} = \text{Null}\left([K]_{\mathcal{B}}^{\mathcal{B}} - \sqrt{3}I\right) = \text{Null}\left(\begin{bmatrix}1-\sqrt{3} & 1 & 1\\ 1 & -1-\sqrt{3} & 1\\ 0 & 1 & -\sqrt{3}\end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix}2+\sqrt{3}\\ \sqrt{3}\\ 1\end{bmatrix}\right\}$$

$$E_{-\sqrt{3}} = \text{Null}\left([K]_{\mathcal{B}}^{\mathcal{B}} + \sqrt{3}I\right) = \text{Null}\left(\begin{bmatrix}1+\sqrt{3} & 1 & 1\\ 1 & -1+\sqrt{3} & 1\\ 0 & 1 & \sqrt{3}\end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix}2-\sqrt{3}\\ -\sqrt{3}\\ 1\end{bmatrix}\right\}$$

Which means the eigenspaces of K will be:

$E_0 = \text{Span}\left\{-x^2+1\right\}$

$E_{\sqrt{3}} = \text{Span}\left\{(2+\sqrt{3})x^2 + x\sqrt{3} + 1\right\}$

$E_{-\sqrt{3}} = \text{Span}\left\{(2-\sqrt{3})x^2 - x\sqrt{3} + 1\right\}$

? Exercise 4.2.3.5

Consider the function $\Psi : \mathbb{P}_2(\mathbb{R}) \to \mathbb{P}_2(\mathbb{R})$ satisfying $\Psi\left(ax^2+bx+c\right) = (3a-b-2c)x^2 + (a+b-2c)x + (a-b)$

(a) Prove Ψ is a linear transformation.

(b) Find a basis for $\ker(\Psi)$. Is Ψ one-to-one?

(c) Find a basis for $\text{Range}(\Psi)$. Is Ψ onto?

(d) Find all eigenvalues and eigenspaces of Ψ

Answer

(a) We check the two properties of linear transformations:

$\Psi\left((ax^2+bx+c)+(dx^2+ex+f)\right)) = \Psi\left((a+d)x^2+(b+e)x+(c+f)\right)$

$= (3(a+d)-(b+e)-2(c+f))\,x^2 + ((a+d)+(b+e)-2(c+f))\,x + ((a+d)-(b+e))$

$= ((3a-b-2c)x^2+(a+b-2c)x+(a-b)) + ((3d-e-2f)x^2+(d+e-2f)x+(d-e)) = \Psi(ax^2+bx+c) + \Psi(dx^2+ex+f)$

$\Psi\left(r(ax^2+bx+c)\right) = \Psi(rax^2+rbx+rc) = (3ra-rb-2rc)x^2 + (ra+rb-2rc)x + (ra-rb)$

$= r\left((3a-b-2c)x^2+(a+b-2c)x+(a-b)\right) = r\cdot\Psi(ax^2+bx+c)$

This proves Ψ is a linear transformation.

To facilitate the next three parts, we let $\mathcal{B} = \{x^2, x, 1\}$ and find $[\Psi]_{\mathcal{B}}^{\mathcal{B}}$

$$\Psi\left(x^2\right) = 3x^2+1x+1 \longrightarrow \begin{bmatrix}3\\1\\1\end{bmatrix}$$

$$\Psi(x) = -1x^2 + 1x - 1 \longrightarrow \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$$\Psi(1) = -2x^2 - 2x + 0 \longrightarrow \begin{bmatrix} -2 \\ -2 \\ 0 \end{bmatrix}$$

Therefore, we have $[\Psi]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 3 & -1 & -2 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix}$ which has RREF $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$

(b) To find $\ker(\Psi)$, we solve for $\text{Null}\left([\Psi]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$

This means that $\ker(\Psi) = \text{Span}\left\{x^2 + x + 1\right\}$

This contains nonzero vectors, we know that Ψ **is not one-to-one.**

(c) To find $\text{Range}(\Psi)$, we solve for $\text{Col}\left([\Psi]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}\right\}$

This means that $\text{Range}(\Psi) = \text{Span}\left\{3x^2 + x + 1, -x^2 + x - 1\right\}$

This has dimension 2, and the codomain $\mathbb{P}_2(\mathbb{R})$ has dimension 3, we know that Ψ **is not onto.**

(d) The characteristic polynomial is $p(\lambda) = \det\left([\Psi]_{\mathcal{B}}^{\mathcal{B}} - \lambda I\right) = \lambda(\lambda - 2)^2$

Thus, the eigenvalues of $[\Psi]_{\mathcal{B}}^{\mathcal{B}}$ (and also of Ψ) are $\lambda = 0, 2$

The eigenspaces of $[\Psi]_{\mathcal{B}}^{\mathcal{B}}$ are:

$E_0 = \text{Null}\left([\Psi]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$ as we calculated in part (b)

$$E_2 = \text{Null}\left([\Psi]_{\mathcal{B}}^{\mathcal{B}} - 2I\right) = \text{Null}\left(\begin{bmatrix} 1 & -1 & -2 \\ 1 & -1 & -2 \\ 1 & -1 & -2 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}\right\}$$

Which means the eigenspaces of Ψ will be:

$E_0 = \text{Span}\left\{x^2 + x + 1\right\}$

$E_2 = \text{Span}\left\{x^2 + x, 2x^2 + 1\right\}$

? Exercise 4.2.3.6

Consider the function $\Phi : \mathbb{P}_3(\mathbb{R}) \to \mathbb{P}_3(\mathbb{R})$ satisfying $\Phi(p(x)) = (x^2 - 1)p''(x)$

(a) Prove Φ is a linear transformation.

(b) Find a basis for $\ker(\Phi)$. Is Φ one-to-one?

(c) Find a basis for $\text{Range}(\Phi)$. Is Φ onto?

(d) Find all eigenvalues and eigenspaces of Φ

Answer

(a) We check the two properties of linear transformations:

$$\Phi((p+q)(x)) = (x^2-1)(p+q)''(x) = (x^2-1)(p''(x) + q''(x)) = (x^2-1)p''(x) + (x^2-1)q''(x) = \Phi(p(x)) + \Phi(q(x))$$

$$\Phi((rp)(x)) = (x^2-1)(rp)''(x) = (x^2-1)(r \cdot p''(x)) = r \cdot (x^2-1)p''(x) = r \cdot \Phi(p(x))$$

This proves Φ is a linear transformation.

To facilitate the next three parts, we let $\mathcal{B} = \{x^3, x^2, x, 1\}$ and find $[\Phi]_{\mathcal{B}}^{\mathcal{B}}$

$$\Phi(x^3) = (x^2-1)(6x) = 6x^3 - 6x \longrightarrow \begin{bmatrix} 6 \\ 0 \\ -6 \\ 0 \end{bmatrix}$$

$$\Phi(x^2) = (x^2-1)(2) = 2x^2 - 2 \longrightarrow \begin{bmatrix} 0 \\ 2 \\ 0 \\ -2 \end{bmatrix}$$

$$\Phi(x) = -(x^2-1)(0) = 0 \longrightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Phi(1) = -(x^2-1)(0) = 0 \longrightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Therefore, we have $[\Phi]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 6 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ -6 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \end{bmatrix}$ which has RREF $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(b) To find $\ker(\Phi)$, we solve for $\text{Null}\left([\Phi]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$

This means that $\ker(\Phi) = \text{Span}\{x, 1\}$

This contains nonzero vectors, we know that Φ **is not one-to-one.**

(c) To find $\text{Range}(\Phi)$, we solve for $\text{Col}\left([\Phi]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} 6 \\ 0 \\ -6 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 0 \\ -2 \end{bmatrix}\right\}$

This means that $\text{Range}(\Phi) = \text{Span}\{6x^3 - 6x, 2x^2 - 2\}$

This has dimension 2, and the codomain $\mathbb{P}_3(\mathbb{R})$ has dimension 4, we know that Φ **is not onto.**

(d) The characteristic polynomial is $p(\lambda) = \det\left([\Phi]_{\mathcal{B}}^{\mathcal{B}} - \lambda I\right) = \lambda^2(\lambda-2)(\lambda-6)$

Thus, the eigenvalues of $[\Phi]_{\mathcal{B}}^{\mathcal{B}}$ (and also of Φ) are $\lambda = 0, 2, 6$

The eigenspaces of $[\Phi]_{\mathcal{B}}^{\mathcal{B}}$ are:

$$E_0 = \text{Null}\left([\Phi]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\} \text{ as we calculated in part (b)}$$

$$E_2 = \text{Null}\left([\Phi]_{\mathcal{B}}^{\mathcal{B}} - 2I\right) = \text{Null}\left(\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -6 & 0 & -2 & 0 \\ 0 & -2 & 0 & -2 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}\right\}$$

$$E_6 = \text{Null}\left([\Phi]_{\mathcal{B}}^{\mathcal{B}} - 6I\right) = \text{Null}\left(\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 \\ -6 & 0 & -6 & 0 \\ 0 & -2 & 0 & -6 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}\right\}$$

Which means the eigenspaces of Φ will be:

$E_0 = \text{Span}\{x, 1\}$

$E_2 = \text{Span}\{-x^2 + 1\}$

$E_6 = \text{Span}\{-x^3 + x\}$

? Exercise 4.2.3.7

Consider the subspace $W = \text{Span}\{e^x, e^{-x}, \sin(x), \cos(x)\}$ of $\mathbb{F}(\mathbb{R})$.

Define the function $T : W \to W$ satisfying $T(f(x)) = f'(x)$

(a) Prove T is a linear transformation.

(b) Find a basis for $\ker(T)$. Is T one-to-one?

(c) Find a basis for $\text{Range}(T)$. Is T onto?

(d) Find all real eigenvalues and eigenspaces of T

(e) Find all real and complex eigenvalues and eigenspaces of T

Answer

(a) We check the two properties of linear transformations:

$T((f+g)(x)) = (f+g)'(x) = f'(x) + g'(x) = T(f(x)) + T(g(x))$

$T((rf)(x)) = (rf)'(x) = r \cdot f'(x) = r \cdot T(f(x))$

This proves T is a linear transformation.

To facilitate the next three parts, we let $\mathcal{B} = \{e^x, e^{-x}, \sin(x), \cos(x)\}$ and find $[T]_{\mathcal{B}}^{\mathcal{B}}$

$$T(e^x) = e^x \longrightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$T(e^{-x}) = -e^{-x} \longrightarrow \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

$$T(\sin(x)) = \cos(x) \longrightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$T(\cos(x)) = -\sin(x) \longrightarrow \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}$$

Therefore, we have $[T]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ which has RREF $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

(b) To find $\ker(T)$, we solve for $\text{Null}\left([T]_{\mathcal{B}}^{\mathcal{B}}\right) = \left\{\vec{0}\right\}$

This means that $\ker(T) = \left\{\vec{0}\right\}$ (the constant zero function)

This contains only the zero vector, we know that **T is one-to-one.**

(c) To find $\text{Range}(T)$, we solve for $\text{Col}\left([T]_{\mathcal{B}}^{\mathcal{B}}\right) = \text{Span}\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\-1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\0\\-1\\0\end{bmatrix}\right\}$

This means that $\text{Range}(T) = \text{Span}\left\{e^x, -e^{-x}, \cos(x), -\sin(x)\right\}$

This has dimension 4, and the codomain W has dimension 4, we know that T **is onto.**

(d) The characteristic polynomial is $p(\lambda) = \det\left([T]_{\mathcal{B}}^{\mathcal{B}} - \lambda I\right) = (\lambda - 1)(\lambda + 1)(\lambda^2 + 1)$

Thus, the real eigenvalues of $[T]_{\mathcal{B}}^{\mathcal{B}}$ (and also of T) are $\lambda = -1, 1$

The real eigenspaces of $[T]_{\mathcal{B}}^{\mathcal{B}}$ are:

$$E_1 = \text{Null}\left([T]_{\mathcal{B}}^{\mathcal{B}} - I\right) = \text{Null}\left(\begin{bmatrix}0&0&0&0\\0&-2&0&0\\0&0&-1&-1\\0&0&1&-1\end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}\right\}$$

$$E_{-1} = \text{Null}\left([T]_{\mathcal{B}}^{\mathcal{B}} + I\right) = \text{Null}\left(\begin{bmatrix}2&0&0&0\\0&0&0&0\\0&0&1&-1\\0&0&1&1\end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix}0\\1\\0\\0\end{bmatrix}\right\}$$

Which means the real eigenspaces of T will be:

$E_1 = \text{Span}\left\{e^x\right\}$

$E_{-1} = \text{Span}\left\{e^{-x}\right\}$

(e) The complex eigenvalues of $[T]_{\mathcal{B}}^{\mathcal{B}}$ (and also of T) are $\lambda = -i, i$

These eigenspaces of $[T]_{\mathcal{B}}^{\mathcal{B}}$ are:

$$E_i = \text{Null}\left([T]_{\mathcal{B}}^{\mathcal{B}} - iI\right) = \text{Null}\left(\begin{bmatrix}1-i&0&0&0\\0&-1-i&0&0\\0&0&-i&-1\\0&0&1&-i\end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix}0\\0\\i\\1\end{bmatrix}\right\}$$

$E_{-i} = \text{Span}\left\{\begin{bmatrix}0\\0\\-i\\1\end{bmatrix}\right\}$ since it is the conjugate eigenspace

Which means the complex eigenspaces of T will be:

$E_i = \text{Span}\left\{\cos(x) + i\sin(x)\right\}$ (look familiar? This is Euler's formula for e^{ix})

$E_{-i} = \text{Span}\left\{\cos(x) - i\sin(x)\right\}$ (Euler's formula for e^{-ix})

? Exercise 4.2.3.8

Consider the subspace $Y = \text{Span}\left\{x\sin(x), \sin(x), x\cos(x), \cos(x)\right\}$ of $\mathbb{F}(\mathbb{R})$.

Define the function $\sigma : Y \to Y$ satisfying $\sigma\left(f(x)\right) = f'(x)$

(a) Prove σ is a linear transformation.

(b) Find a basis for $\ker(\sigma)$. Is σ one-to-one?

(c) Find a basis for $\text{Range}(\sigma)$. Is σ onto?

(d) Find all real and complex eigenvalues of σ

Answer

(a) We check the two properties of linear transformations:

$\sigma((f+g)(x)) = (f+g)'(x) = f'(x) + g'(x) = \sigma(f(x)) + \sigma(g(x))$

$\sigma((rf)(x)) = (rf)'(x) = r \cdot f'(x) = r \cdot \sigma(f(x))$

This proves σ is a linear transformation.

To facilitate the next three parts, we let $\mathcal{B} = \{x\sin(x), \sin(x), x\cos(x), \cos(x)\}$ and find $[\sigma]_{\mathcal{B}}^{\mathcal{B}}$

$$\sigma(x\sin(x)) = \sin(x) + x\cos(x) \longrightarrow \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\sigma(\sin(x)) = \cos(x) \longrightarrow \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\sigma(x\cos(x)) = \cos(x) - x\sin(x) \longrightarrow \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\sigma(\cos(x)) = -\sin(x) \longrightarrow \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

Therefore, we have $[\sigma]_{\mathcal{B}}^{\mathcal{B}} = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$ which has RREF $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

(b) To find $\mathbf{ker}(\sigma)$, we solve for $\mathbf{Null}\left([\sigma]_{\mathcal{B}}^{\mathcal{B}}\right) = \left\{\vec{0}\right\}$

This means that $\mathbf{ker}(\sigma) = \left\{\vec{0}\right\}$ (the constant zero function)

This contains only the zero vector, we know that σ **is one-to-one.**

(c) To find $\mathbf{Range}(\sigma)$, we solve for $\mathbf{Col}\left([\sigma]_{\mathcal{B}}^{\mathcal{B}}\right) = \mathbf{Span}\left\{\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}\right\}$

This means that $\mathbf{Range}(\sigma) = \mathbf{Span}\{\sin(x) + x\cos(x), \cos(x), \cos(x) - x\sin(x), -\sin(x)\}$

This has dimension 4, and the codomain Y has dimension 4, we know that σ **is onto.**

(d) The characteristic polynomial is $p(\lambda) = \det\left([\sigma]_{\mathcal{B}}^{\mathcal{B}} - \lambda I\right) = \lambda^4 + 2\lambda^2 + 1 = (\lambda^2 + 1)^2$

Thus, the eigenvalues of $[\sigma]_{\mathcal{B}}^{\mathcal{B}}$ (and also of σ) are $\lambda = \pm i$

4.2.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.3: Examples and Applications

Learning Objectives

1. Understand connections between linear algebra and other mathematical/real world applications.
2. Examine anti-derivatives and differential equations in relation to the derivative operator.
3. Define famous linear transformations used in other math courses.
4. Consider the inner product as a generalization of the dot product in a vector space.
5. See a useful case of repeated matrix multiplication and eigenvalues in probability.
6. Examine how linear algebra tools are used by computers to affect daily life.

While most of linear algebra focused on the study of matrices and their operations, we can see that these skills can be applied to general vector spaces to give useful real life results.

This page titled 4.3: Examples and Applications is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Kenn Huber.

4.3.1: Differential Equations

In the previous section, we discussed the derivative transformation: $T : C'([a,b]) \to \mathbb{F}$, and its kernel: $\ker(T) = \text{Span}\{1\}$

Since this transformation has non-zero vectors in its kernel, it is clearly not invertible. That didn't stop us from defining anti-derivatives though.

Illustration 4.3.1.1: Antiderivatives

For any function $f(x) \in \mathbb{F}$ defined on the interval $[a,b]$, we define an **anti-derivative** to be any function $g(x)$ in $C'([a,b])$ satisfying $T(g(x)) = f(x)$.

Since any constant function $h(x) = C$ is in the kernel of T, we have $T(h(x)) = 0$.

This means that if we find one anti-derivative $g(x)$ satisfying $T(g(x)) = f(x)$, then

$$T(g(x) + h(x)) = T(g(x)) + T(h(x)) = f(x) + 0 = f(x)$$

This means that $g(x) + h(x)$ is also an anti-derivative of $f(x)$.

This is the **reason** that we use +C when taking anti-derivatives in calculus!

For example, we know that $f(x) = 2x$ has a **particular solution** of $g(x) = x^2$, and so all possible anti-derivatives are of the form $F(x) = x^2 + C$

Illustration 4.3.1.2: More Antiderivatives

Consider two functions $f_1(x) = \sin^{-1}(x)$ and $f_2(x) = -\cos^{-1}(x)$ defined on $[-1,1]$.

It is easy to see that $f_1'(x) = \dfrac{1}{\sqrt{1-x^2}} = f_2'(x)$ for all $x \in (-1,1)$

This means that $T(f_1(x)) = T(f_2(x))$, and using the property of linear transformations:

$$T(f_1(x) - f_2(x)) = T(f_1(x)) - T(f_2(x)) = 0 \implies f_1(x) - f_2(x) \in \ker(T)$$

Since the kernel of T only contains constant functions, this tells us that $f_1(x) - f_2(x) = C$ for some constant C.

We can even evaluate this constant by plugging in any x in our interval, like $x = 0$:

$$C = f_1(0) - f_2(0) = \sin^{-1}(0) + \cos^{-1}(0) = 0 + \frac{\pi}{2} = \frac{\pi}{2}$$

This means that $\sin^{-1}(x) + \cos^{-1}(x) = \dfrac{\pi}{2}$ holds for all $x \in (-1,1)$ (and even at the endpoints thanks to continuity).

Since we have proven that linear combinations and compositions of linear transformations are always linear transformations, we can apply the derivative operator any number of times to our input function. In this subsection, we use the subspace $C^n([a,b]) = \{f(x) \in \mathbb{F} : f \text{ is n times differentiable on } [a,b]\}$

Definition: Linear Homogeneous Differential Equation

For any function $y(x) \in C^n([a,b])$ and functions $a_0(x), a_1(x), \ldots, a_n(x) \in \mathbb{F}$, define $T : C^n([a,b]) \to \mathbb{F}$ by

$$T(y(x)) = a_n(x)y^{(n)}(x) + a_{n-1}(x)y^{(n-1)}(x) + \cdots + a_1(x)y'(x) + a_0(x)y(x)$$

This is a linear transformation since:

$$\begin{gathered}T(y_1+y_2)\\ =a_n(x)(y_1+y_2)^{(n)}(x)+a_{n-1}(x)(y_1+y_2)^{(n-1)}(x)+\cdots+a_1(x)(y_1+y_2)'(x)+a_0(x)(y_1+y_2)(x)\\ =a_n(x)(y_1^{(n)}(x)+y_2^{(n)}(x))+a_{n-1}(x)(y_1^{(n-1)}(x)+y_2^{(n-1)}(x))\\ +\cdots+a_1(x)(y_1'(x)+y_2'(x))+a_0(x)(y_1(x)+y_2(x))\\ =a_n(x)y_1^{(n)}(x)+a_{n-1}(x)y_1^{(n-1)}(x)\\ +\cdots+a_1(x)y_1'(x)+a_0(x)y_1(x)+a_n(x)y_2^{(n)}(x)+a_{n-1}(x)y_2^{(n-1)}(x)+\cdots+a_1(x)y_2'(x)+a_0(x)y_2(x)\\ =T(y_1)+T(y_2)\end{gathered}$$

$$\begin{gathered}T(cy)=a_n(x)(cy)^{(n)}(x)+a_{n-1}(x)(cy)^{(n-1)}(x)+\cdots+a_1(x)(cy)'(x)+a_0(x)(cy)(x)\\ =ca_n(x)y^{(n)}(x)+ca_{n-1}(x)y^{(n-1)}(x)+\cdots+ca_1(x)y'(x)+ca_0(x)y(x)\\ =c\left(a_n(x)y^{(n)}(x)+a_{n-1}(x)y^{(n-1)}(x)+\cdots+a_1(x)y'(x)+a_0(x)y(x)\right)=cT(y)\end{gathered}$$

We say that a **linear homogeneous differential equation of order n** is a differential equation that looks like:

$$a_n(x)y^{(n)}(x)+a_{n-1}(x)y^{(n-1)}(x)+\cdots+a_1(x)y'(x)+a_0(x)y(x)=0$$

When asked to solve this, we are simply being asked to find the kernel of this type of transformation, which will always be n-dimensional.

Example 4.3.1.1: Example DE

Find a solution to the differential equation $xy''-y'=0$ which satisfies $y(1)=1,\ y'(1)=4$

Solution

We define T to be the transformation $T(y)=xy''-y'$, and look for the element in its kernel with $y(1)=1,\ y'(1)=4$.

This begins with us finding a spanning set for the kernel:

If we guess that $y(x)=x^n$ is in $\ker(T)$ for some n, we see that $y'=nx^{n-1},\ y''=n(n-1)x^{n-2}$.

In order to satisfy the differential equation above, that would mean that:

$$T(x^n)=x\cdot n(n-1)x^{n-2}-nx^{n-1}=n(n-2)x^{n-1}$$

Clearly this equals zero for all possible x values when $n=0,2$.

Thus, we have found two **linearly independent** solutions in our kernel, which will be two dimensional. Thus:

$$\ker(T)=\text{Span}\{x^2,1\}=\{ax^2+b:a,b\in\mathbb{R}\}$$

We now focus on finding exactly which linear combination of these vectors satisfies the initial conditions $y(1)=1,\ y'(1)=4$:

$$\begin{cases}y(1)=a+b=1\\ y'(1)=2a=4\end{cases}\implies a=2,b=-1$$

Thus, we have shown that our solution must be $\boxed{y(x)=2x^2-1}$

Example 4.3.1.2: Example Non-homogeneous DE

Find the general solution to the non-homogeneous differential equation $xy''-y'=1$

Solution

We define T to be the transformation $T(y)=xy''-y'$, and use the kernel we calculated above:

$$\ker(T)=\text{Span}\{x^2,1\}=\{ax^2+b:a,b\in\mathbb{R}\}$$

Since this equation is non-homogeneous, our solution will be in the familiar form: $y=y_p+y_h$, where y_p is a particular solution and y_h is all possible homogeneous solutions (which is our kernel)

Thankfully, it is pretty easy to guess *one* function that satisfies $xy''-y'=1$, for example $y_p=-x$ with $y_p'=-1,y_p''=0$ works.

This means all of our solutions will be of the form

$$\boxed{y(x) = -x + ax^2 + b}$$

Note

"How did we know to guess $y = x^n$ was in the kernel?"

That is a completely valid question, and the answer is found in a differential equations course (where you learn how to recognize which guess to make based on the shape of the original differential equation given).

4.3.1: Differential Equations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.3.2: Famous Transformations

Some differential equations are particularly difficult to answer with ordinary means. Linear algebra gives us help here too.

In a finite dimensional vector space V, we have learned that with any basis $\mathcal{B}$ of V, we can use a linear transformation to coordinatize each vector in V and "turn it into" a vector in $\mathbb{R}^n$. This is done because we have powerful tools that make solving questions easier in $\mathbb{R}^n$, and when we are done we un-coordinatize our answers to transform them back into V.

This is a common mathematical trick. Transform objects into another vector space (where calculations are easier), solve the problem, and then transform them back.

Definition: Laplace Transformations

The Laplace Transformation of a function $f(t)$ is defined to be the function $\mathcal{L}(f) = \int_0^\infty f(t)e^{-st}dt$

This is a linear transformation:

$$\mathcal{L}(f+g) = \int_0^\infty [f(t)+g(t)]e^{-st}dt = \int_0^\infty f(t)e^{-st}dt + \int_0^\infty g(t)e^{-st}dt = \mathcal{L}(f) + \mathcal{L}(g)$$

$$\mathcal{L}(cf) = \int_0^\infty [cf(t)]e^{-st}dt = c\int_0^\infty f(t)e^{-st}dt = c\mathcal{L}(f)$$

We restrict the domain of $\mathcal{L}$ to be $V = \{f \in \mathbb{F} : |f(x)| \leq Me^{cx} \text{ for some } M, c \in \mathbb{R}\}$ (so that the integral always converges), and then define $W = \text{Range}(\mathcal{L})$ so that $\mathcal{L} : V \to W$ is invertible (but the inverse is ugly to look at)

It is (easily) verified that the following formulas (which we will use in the example below) hold for the Laplace transformation:

$$\mathcal{L}(f'(t)) = s\mathcal{L}(f) - f(0) \text{ and } \mathcal{L}(e^{at}) = \frac{1}{s-a} \leftrightarrow \mathcal{L}^{-1}\left(\frac{1}{s-a}\right) = e^{at}$$

Illustration 4.3.2.3: Eigenvalues of the Derivative

We will find all real eigenvalues and eigenvectors of the derivative transformation $T : C'(\mathbb{R}) \to C'(\mathbb{R})$ with $T(f) = f'$

We want to find the functions $f \in C'(\mathbb{R})$ which satisfy $T(f) = \lambda f$

Ultimately, we are looking to solve the differential equation $y' = \lambda y$. We will do so by applying the Laplace Transformation to both sides:

$$\mathcal{L}(y') = \mathcal{L}(\lambda y) \implies s\mathcal{L}(y) - y(0) = \lambda\mathcal{L}(y)$$

We can solve for $\mathcal{L}(y)$ to get $\mathcal{L}(y) = \frac{y(0)}{s-\lambda}$

Now we apply the inverse transformation to both sides:

$$\mathcal{L}^{-1}(\mathcal{L}(y)) = \mathcal{L}^{-1}\left(\frac{y(0)}{s-\lambda}\right) \implies y = y(0)e^{\lambda t}$$

Therefore **all real numbers** are eigenvalues of T, and each has a one dimensional eigenspace $E_\lambda = \text{Span}\{e^{\lambda t}\}$

Definition: Fourier Transformations

The Fourier Transformation of a function $f(t)$ is defined to be the function $\mathcal{F}(f) = \int_{-\infty}^\infty f(t)e^{-2\pi i\, st}dt$

This is also a linear transformation:

$$\mathcal{F}(f+g) = \int_{-\infty}^\infty [f(t)+g(t)]e^{-2\pi i\, st}dt = \int_{-\infty}^\infty f(t)e^{-2\pi i\, st}dt + \int_{-\infty}^\infty g(t)e^{-2\pi i\, st}dt = \mathcal{F}(f) + \mathcal{F}(g)$$

$$\mathcal{F}(cf) = \int_{-\infty}^\infty [cf(t)]e^{-2\pi i\, st}dt = c\int_{-\infty}^\infty f(t)e^{-2\pi i\, st}dt = c\mathcal{F}(f)$$

We restrict the domain of $\mathcal{F}$ to be $V = \left\{ f \in \mathbb{F} : \lim_{|x| \to \infty} |f^{(n)}(x)| = 0 \text{ for all } n \right\}$ (so that the integral always converges), and then define $W = \text{Range}(\mathcal{F})$ so that $\mathcal{F} : V \to W$ is invertible (the inverse is again ugly to look at)

Fourier claimed (in 1822) that almost every function can be decomposed into a (possibly infinite) **linear combination** of sine and/or cosine functions of varying frequencies (later dubbed a *Fourier Series*). So, instead of being used to solve differential equations, Fourier Transformations are used to describe a given function f in terms of a complex valued function of s which describes its frequency.

4.3.2: Famous Transformations is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.3.3: Inner Product Spaces

When we wanted to find distances in $\mathbb{R}^n$, we defined the dot product. But what if we wanted to define distance in $\mathbb{P}_n$ or $\mathbb{M}_{m\times n}$? The dot product does not make sense with these objects.

That is because the dot product is really just an *example* of a more general concept called an **inner product space.**

Definition: Inner Product Spaces

For any vector space V, an **inner product**, $\bullet$ is a function that maps two vectors in V to a scalar such that:

$$\vec{u}\bullet\vec{v}=\vec{v}\bullet\vec{u}$$

$$(a\vec{u}+b\vec{v})\bullet\vec{w}=a\vec{u}\bullet\vec{w}+b\vec{v}\bullet\vec{w}$$

$$\vec{v}\bullet\vec{v}>0 \text{ for } \vec{v}\neq\vec{0}$$

Notice these are the exact rules that the dot product satisfies for vectors in $\mathbb{R}^n$.

Every inner product induces a concept of distance: $\|\vec{v}\|=\sqrt{\vec{v}\bullet\vec{v}}$

And it also defines a new form of orthogonality: $\vec{u}\bullet\vec{v}=0 \implies \vec{u}\perp\vec{v}$

Besides the definition of how it is calculated, an inner product has all of the same properties that our classic dot product has.

Example 4.3.3.3: An Inner Product on $\mathbb{P}_2$

Consider the vector space $V=\mathbb{P}_2$ and define the inner product $p\bullet q=\int_0^1 p(x)q(x)dx$

(a) Prove that this inner product space satisfies the three rules defined above.

(b) Find the length of the general vector $p(x)=ax^2+bx+c$

(c) Find an orthogonal basis for V.

Solution

(a) Let $p(x),q(x),r(x)\in\mathbb{P}_2$, and $a,b\in\mathbb{R}$ then:

$$p\bullet q=\int_0^1 p(x)q(x)dx=\int_0^1 q(x)p(x)dx=q\bullet p$$

$$(ap+bq)\bullet r$$

$$=\int_0^1 (ap(x)+bq(x))r(x)dx=a\int_0^1 p(x)r(x)dx+b\int_0^1 q(x)r(x)dx=a(p\bullet r)+b(q\bullet r)$$

$$p\bullet p=\int_0^1 p(x)p(x)dx=\int_0^1 (p(x))^2dx>0$$

This verifies that we have defined an inner product space.

(b) The length of $p(x)=ax^2+bx+c$ is defined by

$$\|p\|^2=p\bullet p=\int_0^1 (ax^2+bx+c)^2dx$$

$$=\int_0^1 a^2x^4+2abx^3+(2ac+b^2)x^2+2bcx+c^2dx$$

$$=\left[\frac{a^2}{5}x^5+\frac{ab}{2}x^4+\frac{2ac+b^2}{3}x^3+bcx^2+c^2x\right]_0^1$$

$$=\frac{a^2}{5}+\frac{3ab+4ac}{6}+\frac{b^2}{3}+bc+c^2$$

So we have $\boxed{\|ax^2+bx+c\| = \sqrt{\frac{a^2}{5}+\frac{3ab+4ac}{6}+\frac{b^2}{3}+bc+c^2}}$

(c) We can take any basis for V and use Gram-Schmidt to make it into an orthogonal basis. We start with $\mathcal{B} = \{1, x, x^2\}$:

We keep the first vector the same: $\vec{w}_1 = \vec{v}_1 = 1$

Then we use $\vec{w}_2 = \vec{v}_2 - \frac{\vec{v}_2 \bullet \vec{w}_1}{\vec{w}_1 \bullet \vec{w}_1}\vec{w}_1$, so we calculate:

$$\vec{v}_2 \bullet \vec{w}_1 = \int_0^1 x \cdot 1 dx = \frac{1}{2}$$

$$\vec{w}_1 \bullet \vec{w}_1 = \int_0^1 1 \cdot 1 dx = 1$$

Therefore, $\vec{w}_2 = \vec{v}_2 - \frac{1}{2}\vec{w}_1 = x - \frac{1}{2}$

Lastly we will use $\vec{w}_3 = \vec{v}_3 - \frac{\vec{v}_3 \bullet \vec{w}_1}{\vec{w}_1 \bullet \vec{w}_1}\vec{w}_1 - \frac{\vec{v}_3 \bullet \vec{w}_2}{\vec{w}_2 \bullet \vec{w}_2}\vec{w}_2$:

$$\vec{v}_3 \bullet \vec{w}_1 = \int_0^1 x^2 \cdot 1 dx = \frac{1}{3}$$

$$\vec{v}_3 \bullet \vec{w}_2 = \int_0^1 x^2 \cdot \left(x - \frac{1}{2}\right) dx = \frac{1}{12}$$

$$\vec{w}_2 \bullet \vec{w}_2 = \int_0^1 \left(x - \frac{1}{2}\right) \cdot \left(x - \frac{1}{2}\right) dx = \frac{1}{12}$$

Therefore, $\vec{w}_3 = \vec{v}_3 - \frac{1}{3}\vec{w}_1 - 1\vec{w}_2 = x^2 - x - \frac{1}{3}$

So our orthogonal basis with respect to this inner product is $\boxed{\left\{1, x - \frac{1}{2}, x^2 - x - \frac{1}{3}\right\}}$

Example 4.3.3.4: Another Inner Product on $\mathbb{P}_2$

Consider the vector space $V = \mathbb{P}_2$ and define the inner product $p \bullet q = p(-1)q(-1) + p(0)q(0) + p(1)q(1)$

(a) Find the length of the general vector $p(x) = ax^2 + bx + c$

(b) Find an orthogonal basis for V

Solution

(a) The length of $p(x) = ax^2 + bx + c$ is defined by

$$\begin{aligned}\|p\|^2 = p \bullet p &= p(-1)p(-1) + p(0)p(0) + p(1)p(1) \\ &= (a - b + c)^2 + (c)^2 + (a + b + c)^2 \\ &= 2a^2 + 4ac + 2b^2 + 3c^2\end{aligned}$$

So we have $\boxed{\|ax^2+bx+c\| = \sqrt{2a^2+4ac+2b^2+3c^2}}$

(b) We can take any basis for V and use Gram-Schmidt to make it into an orthogonal basis. We start with $\mathcal{B} = \{1, x, x^2\}$:

We keep the first vector the same: $\vec{w}_1 = \vec{v}_1 = 1$

Then we use $\vec{w}_2 = \vec{v}_2 - \frac{\vec{v}_2 \bullet \vec{w}_1}{\vec{w}_1 \bullet \vec{w}_1}\vec{w}_1$, so we calculate:

$$\vec{v}_2 \bullet \vec{w}_1 = (-1)(1) + (0)(1) + (1)(1) = 0$$

These are already orthogonal, so $\vec{w}_2 = \vec{v}_2 - 0\vec{w}_1 = x$

Lastly we will use $\vec{w}_3 = \vec{v}_3 - \frac{\vec{v}_3 \bullet \vec{w}_1}{\vec{w}_1 \bullet \vec{w}_1}\vec{w}_1 - \frac{\vec{v}_3 \bullet \vec{w}_2}{\vec{w}_2 \bullet \vec{w}_2}\vec{w}_2$:

$$\vec{v}_3 \bullet \vec{w}_1 = (1)(1) + (0)(1) + (1)(1) = 2$$

$$\vec{w}_1 \bullet \vec{w}_1 = (1)(1) + (1)(1) + (1)(1) = 3$$

$$\vec{v}_3 \bullet \vec{w}_2 = (1)(-1) + (0)(0) + (1)(1) = 0$$

Therefore, $\vec{w}_3 = \vec{v}_3 - \frac{2}{3}\vec{w}_1 - 0\vec{w}_2 = x^2 - \frac{2}{3}$

So our orthogonal basis with respect to this inner product is $\boxed{\left\{1, x, x^2 - \frac{2}{3}\right\}}$

Note

It should not be incredibly surprising that the concept of "how we measure an angle" is directly linked to "how we measure distance" between objects. This of course means that the concept of **orthogonality** is defined differently when we change the inner product.

Don't fall for the common misconception: neither of these necessarily mean that the graphs of these functions meet at right angles (or even meet at all).

4.3.3: Inner Product Spaces is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.3.4: Markov Chains

In probability theory, a **Markov Chain** is a process that describes a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. For example, a random walk on the integers where a coin is flipped after each step, if the coin is heads you take one step in the positive direction, and if it flips tails you take a step in the negative direction. In this case, where you will end up only depends on where you were last standing (instead of how many steps it took you to get there, or your last coin flip, etc.)

Illustration 4.3.4.4: The two state problem

One simple example is a two state system, where everyone in Group A flips a coin and everyone in Group B flips a coin, and everyone who got heads moves to Group A while everyone who flipped tails moves to Group B. Then we repeat the process.

Let's assume that people in Group A flip a coin which has probability p of coming up heads, while people in Group B flip a different coin which has probability q of coming up heads. We can draw a diagram of what is happening:

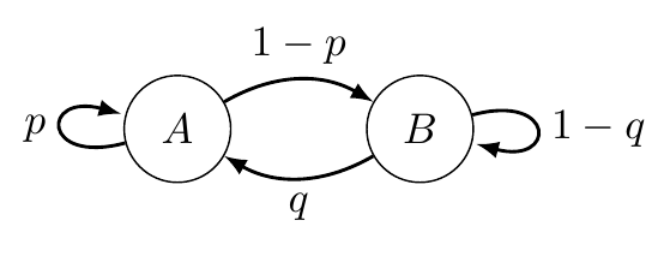


Figure 4.3.4.1

Let $\vec{v} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ represent that there are α people starting in Group A and β people starting in Group B. Of the α who begin in A, we expect $p\alpha$ will flip heads (and remain in A) while $(1-p)\alpha$ will move to B. Similarly, of the β who begin in B, we expect $q\beta$ people will flip heads (and move to A) while $(1-q)\beta$ will remain in B.

This can be represented by matrix multiplication by $M = \begin{bmatrix} p & q \\ 1-p & 1-q \end{bmatrix}$:

$$\begin{bmatrix} \text{new population percent in A} \\ \text{new population percent in B} \end{bmatrix} = \begin{bmatrix} p\alpha + q\beta \\ (1-p)\alpha + (1-q)\beta \end{bmatrix} = \begin{bmatrix} p & q \\ 1-p & 1-q \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = M\vec{v}$$

And if this coin flip is repeated n times, we would have our population vector given by $M^n\vec{v}$.

How do we calculate M^n? Diagonalization! First we find our characteristic polynomial:

$$f(\lambda) = \det(M - \lambda I) = \begin{vmatrix} p-\lambda & q \\ 1-p & 1-q-\lambda \end{vmatrix}$$
$$= (p-\lambda)(1-q-\lambda) - (1-p)q = (\lambda - 1)(\lambda - p + q)$$

So our eigenvalues are $\lambda_1 = 1, \lambda_2 = p - q$, and each has algebraic multiplicity 1, we are assured to be diagonalizable. We find the eigenspaces:

$$E_1 = \text{Null}(M - I) = \text{Null}\left(\begin{bmatrix} p-1 & q \\ 1-p & -q \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} p-1 & q \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} q \\ 1-p \end{bmatrix}\right\}$$

$$E_{p-q} = \text{Null}(M - (p-q)I) = \text{Null}\left(\begin{bmatrix} q & q \\ 1-p & 1-p \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$$

And therefore, $M = CDC^{-1}$ for $D = \begin{bmatrix} 1 & 0 \\ 0 & p-q \end{bmatrix}, C = \begin{bmatrix} q & -1 \\ 1-p & 1 \end{bmatrix}$, so we get

$$M^n = CD^nC^{-1} = \begin{bmatrix} q & -1 \\ 1-p & 1 \end{bmatrix} \begin{bmatrix} 1^n & 0 \\ 0 & (p-q)^n \end{bmatrix} \begin{bmatrix} \frac{1}{q-p+1} & \frac{1}{q-p+1} \\ \frac{p-1}{q-p+1} & \frac{q}{q-p+1} \end{bmatrix}$$
$$= \begin{bmatrix} q & -(p-q)^n \\ 1-p & (p-q)^n \end{bmatrix} \begin{bmatrix} \frac{1}{q-p+1} & \frac{1}{q-p+1} \\ \frac{p-1}{q-p+1} & \frac{q}{q-p+1} \end{bmatrix} = \begin{bmatrix} \frac{q-(p-1)(p-q)^n}{q-p+1} & \frac{q-q(p-q)^n}{q-p+1} \\ \frac{1-p-(1-p)(p-q)^n}{q-p+1} & \frac{1-p+q(p-q)^n}{q-p+1} \end{bmatrix}$$

So after n steps, if we started at $\vec{v} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ we would wind up with an expected distribution of

$$M^n\vec{v} = \begin{bmatrix} \frac{q-(p-1)(p-q)^n}{q-p+1} & \frac{q-q(p-q)^n}{q-p+1} \\ \frac{1-p-(1-p)(p-q)^n}{q-p+1} & \frac{1-p+q(p-q)^n}{q-p+1} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{q}{q-p+1}\alpha + \frac{q}{q-p+1}\beta + \frac{((1-p)\alpha-q\beta)}{q-p+1}(p-q)^n \\ \frac{1-p}{q-p+1}\alpha + \frac{1-p}{q-p+1}\beta + \frac{(q\beta-(1-p)\alpha)}{q-p+1}(p-q)^n \end{bmatrix}$$

If we take the limit of this as $n \to \infty$, since $|p-q| < 1$ the ugliest terms go to zero:

$$\lim_{n\to\infty} M^n\vec{v} = \begin{bmatrix} \frac{q}{q-p+1}\alpha + \frac{q}{q-p+1}\beta \\ \frac{1-p}{q-p+1}\alpha + \frac{1-p}{q-p+1}\beta \end{bmatrix} = (\alpha+\beta)\begin{bmatrix} \frac{q}{q-p+1} \\ \frac{1-p}{q-p+1} \end{bmatrix}$$

The vector $\vec{v}_s = \begin{bmatrix} \frac{q(\alpha+\beta)}{q-p+1} \\ \frac{(1-p)(\alpha+\beta)}{q-p+1} \end{bmatrix}$ that we got from this limit is the **steady-state** vector. That is, the vector that is unchanged by multiplication by the matrix M.

Of course, we might also recognize this as an element of the eigenspace E_1 (since $M\vec{v}_s = \vec{v}_s$), whose entries sum to 1. In fact, we can use this as a shortcut to avoid much of the ugly calculations above:

$$\vec{v}_s \in E_1 \implies \vec{v}_s = \begin{bmatrix} cq \\ c(1-p) \end{bmatrix} \text{ such that } cq + c(1-p) = \alpha+\beta \implies c = \frac{\alpha+\beta}{q-p+1}$$

Example 4.3.4.5

A spaceship is filled with 1000 sterile crew-members, and all of them are either healthy, sick, or dead at any given time.

If you are healthy today, you have a 99% chance of being healthy tomorrow, 1% chance of being sick tomorrow, and 0% chance of being dead tomorrow.
If you are sick today, you have a 50% chance of being healthy tomorrow, 49% chance of being sick tomorrow, and a 1% chance of being dead tomorrow.
If you are dead, you stay dead with 100% probability.

How many crew-members are expected to be dead after a "very long time" passes?

Solution

There are three states to track so we put them into a vector: $\vec{v} = \begin{bmatrix} \text{number of healthy persons} \\ \text{number of sick persons} \\ \text{number of dead persons} \end{bmatrix}$

We can see that tomorrow this distribution will be $M\vec{v}$ for the matrix $M = \begin{bmatrix} .99 & .50 & 0 \\ .01 & .49 & 0 \\ 0 & .01 & 1 \end{bmatrix}$

We skip the ugly calculations, and look for E_1:

$$E_1 = \text{Null}(M-I) = \text{Null}\left(\begin{bmatrix} -.01 & .50 & 0 \\ .01 & -.51 & 0 \\ 0 & .01 & 0 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$$

Therefore, our steady state vector $\vec{v}_s = \begin{bmatrix} 0 \\ 0 \\ c \end{bmatrix}$ with $0+0+c = 1000$.

Tragically, this means that all 1000 crew-members are expected to die after enough time passes.

4.3.4: Markov Chains is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

4.3.5: Page Rank

Internet searching in the 1990s was very inefficient. Yahoo or AltaVista would scan pages for your search text, and simply list the results with the most occurrences of those words. Not surprisingly, the more unsavory websites soon learned that by putting the words "Alanis Morissette" a million times in their pages, they could show up first every time an angsty teenager tried to find *Jagged Little Pill* on Napster.

Larry Page and Sergey Brin invented a way to rank pages by *importance*. They founded Google based on their algorithm. Here is roughly how it works.

Each web page has an associated importance, or ***rank***. This is a positive number. This rank is determined by the following rule.

Definition: The Importance Rule

If a page P links to n other pages $Q_1, Q_2, \ldots, Q_n$, then each page Q_i inherits $\frac{1}{n}$ of P's importance.

In practice, this means:

- If a very important page links to your page (and not to a zillion other ones as well), then your page is considered important.
- If a zillion unimportant pages link to your page, then your page is still important.
- If only one unknown page links to yours, your page is not important.

Alternatively, there is the *random surfer interpretation*. A "random surfer" just sits at his computer all day, randomly clicking on links. The pages he spends the most time on should be the most important. So, the important (high-ranked) pages are those where a random surfer will end up most often. This measure turns out to be equivalent to the rank.

Definition: The Importance Matrix

Consider an internet with n pages. The ***importance matrix*** is the $n \times n$ matrix A whose i, j-entry is the importance that page j passes to page i.

Illustration 4.3.5.5

Consider the following internet with only four pages. Links are indicated by arrows.

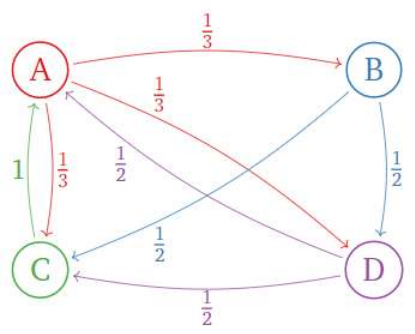


Figure 4.3.5.2

The importance rule says:

- Page A has 3 links, so it passes $\frac{1}{3}$ of its importance to pages B, C, D.
- Page B has 2 links, so it passes $\frac{1}{2}$ of its importance to pages C, D.
- Page C has one link, so it passes all of its importance to page A.
- Page D has 2 links, so it passes $\frac{1}{2}$ of its importance to pages A, C.

In terms of matrices, if $v = (a, b, c, d)$ is the vector containing the ranks a, b, c, d of the pages A, B, C, D, then

$$\begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} c + \frac{1}{2}d \\ \frac{1}{3}a \\ \frac{1}{3}a + \frac{1}{2}b + \frac{1}{2}d \\ \frac{1}{3}a + \frac{1}{2}b \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

The matrix on the left is the ***importance matrix***, and the final equality expresses the importance rule.

The above example illustrates the key observation.

Key Observation

The PageRank vector is an eigenvector of the importance matrix with eigenvalue 1.

To ensure 1 is an eigenvalue, we fix the importance matrix by replacing each zero column with a column of $1/n$s, where n is the number of pages. This ***modified importance matrix*** A' will always have nontrivial eigenspace E_1

Now we choose a number p in $(0, 1)$, called the ***damping factor***. (A typical value is $p = 0.15$.)

Definition: The Google Matrix

Let A be the importance matrix for an internet with n pages, and let A' be the modified importance matrix. The ***Google Matrix*** is the matrix

$$M = (1-p) \cdot A' + p \cdot B \quad \text{where} \quad B = \frac{1}{n} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

In the random surfer interpretation, this matrix M says: with probability p, our surfer will surf to a completely random page; otherwise, he'll click a random link on the current page, unless the current page has no links, in which case he'll surf to a completely random page in either case.

If we declare that the ranks of all of the pages must sum to 1, then we find:

Note: The 25 Billion Dollar Eigenvector

The PageRank vector is the steady state of the Google Matrix.

The hard part is calculating it: in real life, the Google Matrix has zillions of rows.

Example 4.3.5.6: PageRank Vector

What is the PageRank vector for the following internet? (Use the damping factor $p = 0.15$.)

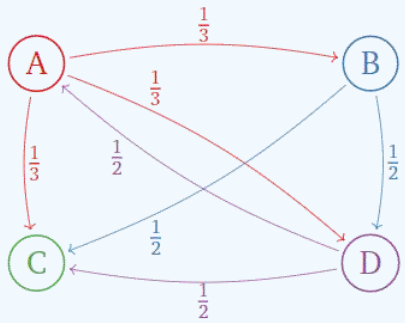


Figure 4.3.5.3

Which page is the most important? Which is the least important?

Solution

First, we compute the modified importance matrix:

$$A = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} \xrightarrow{\text{modify}} A' = \begin{bmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & 0 \end{bmatrix}$$

Choosing the damping factor $p = 0.15$, the Google Matrix is

$$M = 0.85 \cdot \begin{bmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & 0 \end{bmatrix} + 0.15 \cdot \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{bmatrix}$$

$$\approx \begin{bmatrix} 0.0375 & 0.0375 & 0.2500 & 0.4625 \\ 0.3208 & 0.0375 & 0.2500 & 0.0375 \\ 0.3208 & 0.4625 & 0.2500 & 0.4625 \\ 0.3208 & 0.4625 & 0.2500 & 0.0375 \end{bmatrix}$$

The PageRank vector is the steady state vector:

$$\vec{w} \approx \begin{bmatrix} 0.2192 \\ 0.1752 \\ 0.3558 \\ 0.2498 \end{bmatrix}$$

This is the PageRank:

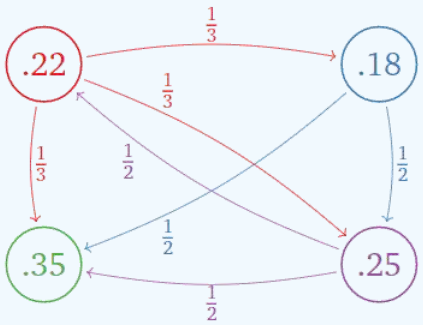


Figure 4.3.5.4

Page C is the most important, with a rank of 0.3558, and page B is the least important, with a rank of 0.1752

Linear algebra is everywhere. The more you study it, the more you can apply it to everyday life, or even better: solve more exciting math problems!

4.3.5: Page Rank is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

- **5.6: Stochastic Matrices** by Dan Margalit & Joseph Rabinoff is licensed GNU FDL. Original source: https://textbooks.math.gatech.edu/ila.

4.3: Practice Set

Exercise 4.3.1a

Consider the linear transformation $T : C''(\mathbb{R}) \to \mathbb{R}$ satisfying $T(f) = 2f''(x) - 3f'(x)$

(i) Find all values of a such that $e^{ax} \in \ker(T)$

(ii) Find all values of b such that $e^{bx} \in \ker(T + I)$

(iii) Find all values of c such that $e^{cx} \in \ker(T - 2I)$

(iv) Find a function $g(x)$ satisfying $2g''(x) - 3g'(x) = 0$ and $g(0) = 0, g'(0) = 6$

(v) Find a function $h(x)$ satisfying $2h''(x) - 3h'(x) + h(x) = \pi$ and $h(0) = \pi, h'(0) = 1$

Answer

(i) $T(e^{ax}) = 2a^2e^{ax} - 3ae^{ax} = (2a^2 - 3a)e^{ax} = a(2a - 3)e^{ax}$

So $e^{ax} \in \ker(T)$ when $a(2a - 3) = 0 \longrightarrow a = 0, \frac{3}{2}$

(ii) $(T + I)(e^{bx}) = 2b^2e^{bx} - 3be^{bx} + e^{bx} = (2b^2 - 3b + 1)e^{bx} = (2b - 1)(b - 1)e^{bx}$

So $e^{bx} \in \ker(T + I)$ when $(2b - 1)(b - 1) = 0 \longrightarrow b = 1, \frac{1}{2}$

(iii) $(T - 2I)(e^{cx}) = 2c^2e^{cx} - 3ce^{cx} - 2e^{cx} = (2c^2 - 3c - 2)e^{cx} = (2c + 1)(c - 2)e^{cx}$

So $e^{cx} \in \ker(T - 2I)$ when $(2c + 1)(c - 2) = 0 \longrightarrow c = 2, -\frac{1}{2}$

(iv) We are solving for some $g(x)$ in $\ker(T)$, which contains e^{0x} and $e^{\frac{3}{2}x}$

This means it must contain (at least) all vectors inside of $\text{Span}\left\{1, e^{\frac{3}{2}x}\right\}$

Letting $g(x) = A \cdot 1 + B \cdot e^{\frac{3}{2}x}$ we see that $g(0) = A + B, g'(0) = \frac{3}{2}B$

Solving $\begin{cases} A + B = 0 \\ \frac{3}{2}B = 6 \end{cases}$ gives $A = -4, B = 4$

Therefore, we know the function $g(x) = -4 + 4e^{\frac{3}{2}x}$ is at least one solution to this differential equation.

(v) We start by *guessing* a particular solution to $2h''(x) - 3h'(x) + h(x) = \pi$, like the constant function $h_p(x) = \pi$.

We now find the shape of the homogeneous solutions: $\ker(T + I)$, which contains e^{1x} and $e^{\frac{1}{2}x}$

This means $\ker(T + I)$ must contain (at least) all vectors inside of $\text{Span}\left\{e^x, e^{\frac{1}{2}x}\right\}$

Letting $h(x) = h_p(x) + A \cdot e^x + B \cdot e^{\frac{1}{2}x}$ we see that $h(0) = \pi + A + B, h'(0) = A + \frac{1}{2}B$

Solving $\begin{cases} \pi + A + B = \pi \\ A + \frac{1}{2}B = 1 \end{cases}$ gives $A = 2, B = -2$

Therefore, we know the function $h(x) = \pi + 2e^x - 2e^{\frac{1}{2}x}$ is at least one solution to this differential equation.

Exercise 4.3.1b

Consider the linear transformation $U : C''(\mathbb{R}) \to \mathbb{R}$ satisfying $U(f) = f''(x) - f(x)$

(i) Find all values of a such that $e^{ax} \in \ker(U)$

(ii) Find all values of b such that $e^{bx} \in \ker(U - 3I)$

(iii) Find all values of c such that $e^{cx} \in \ker(U + 2I)$

(iv) Find a function $g(x)$ satisfying $g''(x) - g(x) = 0$ and $g(0) = 12, g'(0) = 0$

(v) Find a function $h(x)$ satisfying $h''(x) - 4h(x) = 4$ and $h(0) = -5, h'(0) = 8$

Answer

(i) $U(e^{ax}) = a^2e^{ax} - e^{ax} = (a^2 - 1)e^{ax} = (a - 1)(a + 1)e^{ax}$

So $e^{ax} \in \ker(U)$ when $(a - 1)(a + 1) = 0 \longrightarrow a = 1, -1$

(ii) $(U - 3I)(e^{bx}) = b^2e^{bx} - e^{bx} - 3e^{bx} = (b^2 - 4)e^{bx} = (b - 2)(b + 2)e^{bx}$

So $e^{bx} \in \ker(U - 3I)$ when $(b - 2)(b + 2) = 0 \longrightarrow b = 2, -2$

(iii) $(U + 2I)(e^{cx}) = c^2e^{cx} - e^{cx} + 2e^{cx} = (c^2 + 1)e^{cx}$

So $e^{cx} \in \ker(U + 2I)$ when $c^2 + 1 = 0 \longrightarrow c = \pm i$

(iv) We are solving for some $g(x)$ in $\ker(U)$, which contains e^{1x} and e^{-1x}

This means it must contain (at least) all vectors inside of $\text{Span}\{e^x, e^{-x}\}$

Letting $g(x) = A \cdot e^x + B \cdot e^{-x}$ we see that $g(0) = A + B, g'(0) = A - B$

Solving $\begin{cases} A + B = 12 \\ A - B = 0 \end{cases}$ gives $A = 6, B = 6$

Therefore, we know the function $g(x) = 6e^x + 6e^{-x}$ is at least one solution to this differential equation.

(v) We start by *guessing* a particular solution to $h''(x) - 4h(x) = 4$, like the constant function $h_p(x) = -1$.

We now find the shape of the homogeneous solutions: $\ker(U - 3I)$, which contains e^{2x} and e^{-2x}

This means $\ker(U - 3I)$ must contain (at least) all vectors inside of $\text{Span}\{e^{2x}, e^{-2x}\}$

Letting $h(x) = h_p(x) + A \cdot e^{2x} + B \cdot e^{-2x}$ we see that $h(0) = -1 + A + B, h'(0) = 2A - 2B$

Solving $\begin{cases} -1 + A + B = -5 \\ 2A - 2B = 8 \end{cases}$ gives $A = 0, B = -4$

Therefore, we know the function $h(x) = -1 - 4e^{-2x}$ is at least one solution to this differential equation.

? Exercise 4.3.1c

Consider the linear transformation $T : C'(\mathbb{R}) \to \mathbb{R}$ satisfying $T(f) = 2xf'(x)$

(i) Find all values of a such that $x^a \in \ker(T)$

(ii) Find all values of b such that $x^b \in \ker(T - 4I)$

(iii) Find a function $g(x)$ satisfying $2xg'(x) - 4g(x) = 4$ and $g(1) = 2$

Answer

(i) $T(x^a) = 2x \cdot ax^{a-1} = (2a)x^a$

So $x^a \in \ker(T)$ when $2a = 0 \longrightarrow a = 0$

(ii) $(T - 4I)(x^b) = 2x \cdot bx^{b-1} - 4x^b = (2b - 4)x^b = 2(b - 2)x^b$

So $x^b \in \ker(T - 4I)$ when $2(b - 2) = 0 \longrightarrow b = 2$

(iii) We start by *guessing* a particular solution to $2xg'(x) - 4g(x) = 4$, like the constant function $g_p(x) = -1$.

We now find the shape of the homogeneous solutions: $\ker(T - 4I)$, which contains x^2

This means $\ker(T - 4I)$ must contain (at least) all vectors inside of $\text{Span}\{x^2\}$

Letting $g(x) = g_p(x) + A \cdot x^2$ we see that $g(1) = -1 + A$

Solving $-1 + A = 2$ gives $A = 3$

Therefore, we know the function $g(x) = -1 + 3x^2$ is at least one solution to this differential equation.

Exercise 4.3.1d

Consider the linear transformation $U : C''(\mathbb{R}) \to \mathbb{R}$ satisfying $U(f) = x^2 f''(x) + 2xf'(x)$

(i) Find all values of a such that $x^a \in \ker(U)$

(ii) Find all values of b such that $x^b \in \ker(U - 6I)$

(iii) Find a function $g(x)$ satisfying $x^2 g''(x) + 2xg'(x) - 6g(x) = 0$ and $g(1) = 15, g'(1) = 0$

(iv) Find a function $h(x)$ satisfying $x^2 h''(x) + 2xh'(x) - 6h(x) = 6$ and $h(1) = 4, h'(1) = 0$

Answer

(i) $U(x^a) = x^2 \cdot a(a-1)x^{a-2} + 2x \cdot ax^{a-1} = (a^2 + a)x^a = a(a+1)x^a$

So $x^a \in \ker(U)$ when $a(a+1) = 0 \longrightarrow a = 0, -1$

(ii) $(U - 6I)(x^b) = x^2 \cdot b(b-1)x^{b-2} + 2x \cdot bx^{b-1} - 6x^b = (b^2 + b - 6)x^b = (b+3)(b-2)x^b$

So $x^b \in \ker(U - 6I)$ when $(b+3)(b-2) = 0 \longrightarrow b = 2, -3$

(iii) We are solving for some $g(x)$ in $\ker(U - 6I)$, which contains x^2 and x^{-3}

This means it must contain (at least) all vectors inside of $\text{Span}\{x^2, x^{-3}\}$

Letting $g(x) = A \cdot x^2 + B \cdot x^{-3}$ we see that $g(1) = A + B, g'(1) = 2A - 3B$

Solving $\begin{cases} A + B = 15 \\ 2A - 3B = 0 \end{cases}$ gives $A = 9, B = 6$

Therefore, we know the function $g(x) = 9x^2 + 6x^{-3}$ is at least one solution to this differential equation.

(iv) We start by *guessing* a particular solution to $x^2 h''(x) + 2xh'(x) - 6h(x) = 6$, like the constant function $h_p(x) = -1$.

In part (c) we saw $\ker(U - 6I)$ contains (at least) all vectors inside of $\text{Span}\{x^2, x^{-3}\}$

Letting $h(x) = h_p(x) + A \cdot x^2 + B \cdot x^{-3}$ we see that $h(1) = -1 + A + B, h'(1) = 2A - 3B$

Solving $\begin{cases} -1 + A + B = 4 \\ 2A - 3B = 0 \end{cases}$ gives $A = 3, B = 2$

Therefore, we know the function $h(x) = -1 + 3x^2 + 2x^{-3}$ is at least one solution to this differential equation.

Exercise 4.3.2a

Prove that for $\mathcal{L}(f(t)) = \int_0^\infty f(t)e^{-st}dt$ we must have:

(i) $\mathcal{L}(1) = \dfrac{1}{s}$

(ii) $\mathcal{L}(e^{at}) = \dfrac{1}{s-a}$

(iii) $\mathcal{L}(t^n) = \dfrac{n!}{s^{n+1}}$

(iv) $\mathcal{L}(f'(t)) = s\mathcal{L}(f(t)) - f(0)$

(v) $\mathcal{L}(f''(t)) = s^2\mathcal{L}(f(t)) - f'(0) - sf(0)$

(vi) $\mathcal{L}(\sin(at)) = \dfrac{a}{s^2 + a^2}$

(vii) $\mathcal{L}(\cos(at)) = \dfrac{s}{s^2 + a^2}$

Answer

(i) $\mathcal{L}(1) = \int_0^\infty 1 \cdot e^{-st}dt = \int_0^\infty e^{-st}dt = \dfrac{e^{-st}}{-s}\Big]_0^\infty = 0 - \dfrac{1}{-s} = \dfrac{1}{s}$

(ii) $\mathcal{L}(e^{at}) = \int_0^\infty e^{at}\cdot e^{-st}dt = \int_0^\infty e^{(a-s)t}dt = \left.\frac{e^{(a-s)t}}{a-s}\right]_0^\infty = 0 - \frac{1}{a-s} = \frac{1}{s-a}$

(iii) $\mathcal{L}(t^n) = \int_0^\infty t^n\cdot e^{-st}dt = \left[t^n\cdot\frac{e^{-st}}{-s}\right]_0^\infty - \int_0^\infty nt^{n-1}\cdot\frac{e^{-st}}{-s}dt = 0 - 0 + \frac{n}{s}\int_0^\infty t^{n-1}\cdot e^{-st}dt = \frac{n}{s}\mathcal{L}(t^{n-1})$

Using this reduction formula repeatedly gives $\mathcal{L}(t^n) = \frac{n}{s}\mathcal{L}(t^{n-1}) = \frac{n(n-1)}{s^2}\mathcal{L}(t^{n-2}) = \cdots = \frac{n!}{s^n}\mathcal{L}(1) = \frac{n!}{s^{n+1}}$

(iv) $\mathcal{L}(f'(t)) = \int_0^\infty f'(t)\cdot e^{-st}dt = [e^{-st}\cdot f(t)]_0^\infty - \int_0^\infty f(t)\cdot(-se^{-st})\,dt = 0 - f(0) + s\int_0^\infty f(t)\cdot e^{-st}dt = s\mathcal{L}(f(t)) - f(0)$

(v) Use the result in (iv) twice: $\mathcal{L}(f''(t)) = s\mathcal{L}(f'(t)) - f'(0) = s\left(s\mathcal{L}(f(t)) - f(0)\right) - f'(0) = s^2\mathcal{L}(f(t)) - f'(0) - sf(0)$

(vi) Note that for $f(t) = \sin(at)$ we have $f''(t) = -a^2\sin(at)$ so using the result in (v) this says:

$\mathcal{L}(-a^2\sin(at)) = s^2\mathcal{L}(\sin(at)) - a\cos(a\cdot 0) - s\sin(a\cdot 0)$

This is equivalent to $-a^2\mathcal{L}(\sin(at)) = s^2\mathcal{L}(\sin(at)) - a \longrightarrow \mathcal{L}(\sin(at)) = \frac{a}{s^2+a^2}$

(vii) Note that for $f(t) = \cos(at)$ we have $f''(t) = -a^2\cos(at)$ so using the result in (v) this says:

$\mathcal{L}(-a^2\cos(at)) = s^2\mathcal{L}(\cos(at)) + a\sin(a\cdot 0) - s\cos(a\cdot 0)$

This is equivalent to $-a^2\mathcal{L}(\cos(at)) = s^2\mathcal{L}(\cos(at)) - s \longrightarrow \mathcal{L}(\cos(at)) = \frac{s}{s^2+a^2}$

? Exercise 4.3.2b

Use the results from the exercise above to evaluate each expression below:

(i) $\mathcal{L}^{-1}\left(\frac{2}{s-3}\right)$

(ii) $\mathcal{L}^{-1}\left(\frac{2}{5s}\right)$

(iii) $\mathcal{L}^{-1}\left(\frac{12}{s^5}\right)$

(iv) $\mathcal{L}^{-1}\left(\frac{3s^2-6s+8}{s^3-2s^2}\right)$

(v) $\mathcal{L}^{-1}\left(\frac{4s+2}{s^2+4}\right)$

(vi) $\mathcal{L}^{-1}\left(\frac{3}{s^4+s^2}\right)$

Answer

(i) $\mathcal{L}^{-1}\left(\frac{2}{s-3}\right) = 2\cdot\mathcal{L}^{-1}\left(\frac{1}{s-3}\right) = 2\cdot e^{3t} = 2e^{3t}$

(ii) $\mathcal{L}^{-1}\left(\frac{2}{5s}\right) = \frac{2}{5}\cdot\mathcal{L}^{-1}\left(\frac{1}{s}\right) = \frac{2}{5}\cdot 1 = \frac{2}{5}$

(iii) $\mathcal{L}^{-1}\left(\frac{12}{s^5}\right) = \frac{1}{2}\cdot\mathcal{L}^{-1}\left(\frac{4!}{s^5}\right) = \frac{1}{2}\cdot t^4 = \frac{t^4}{2}$

(iv) $\mathcal{L}^{-1}\left(\frac{3s^2-6s+8}{s^3-2s^2}\right) = \mathcal{L}^{-1}\left(\frac{1}{s} - \frac{4}{s^2} + \frac{2}{s-2}\right) = \mathcal{L}^{-1}\left(\frac{1}{s}\right) - 4\cdot\mathcal{L}^{-1}\left(\frac{1!}{s^2}\right) + 2\cdot\mathcal{L}^{-1}\left(\frac{1}{s-2}\right) = 1 - 4\cdot t^1 + 2\cdot e^{2t} = 1 - 4t + 2e^{2t}$

(v) $\mathcal{L}^{-1}\left(\frac{4s+2}{s^2+4}\right) = \mathcal{L}^{-1}\left(\frac{4s}{s^2+4} + \frac{2}{s^2+4}\right) = 4\cdot\mathcal{L}^{-1}\left(\frac{s}{s^2+2^2}\right) + \mathcal{L}^{-1}\left(\frac{2}{s^2+2^2}\right) = 4\cdot\cos(2t) + \sin(2t) = 4\cos(2t) + \sin(2t)$

(vi) $\mathcal{L}^{-1}\left(\frac{3}{s^4+s^2}\right) = \mathcal{L}^{-1}\left(\frac{3}{s^2} - \frac{3}{s^2+1}\right) = 3\cdot\mathcal{L}^{-1}\left(\frac{1!}{s^2}\right) - 3\cdot\mathcal{L}^{-1}\left(\frac{1}{s^2+1^2}\right) = 3\cdot t - 3\cdot\sin(1t) = 3t - 3\sin(t)$

? Exercise 4.3.2c

Use the results from the exercises above to solve each differential equation below:

(i) $y' = 4y$ with $y(0) = 2$

(ii) $y'' = 4y$ with $y(0) = 2, y'(0) = 0$

(iii) $y'' = -4y$ with $y(0) = 2, y'(0) = 0$

(iv) $y'' = -4y$ with $y(0) = 2, y'(0) = 2$

(v) $y'' = 6t$ with $y(0) = 1, y'(0) = 6$

(vi) $y'' - 2y' = 24t^2$ with $y(0) = 0, y'(0) = 0$

Answer

(i) Denote $\mathcal{L}(y(t)) = Y(s)$ and we Laplace both sides to get:

$\mathcal{L}(y'(t)) = s\mathcal{L}(y(t)) - y(0) = sY(s) - 2$ and $\mathcal{L}(4y(t)) = 4Y(s)$

$y' = 4y \longrightarrow sY(s) - 2 = 4Y(s) \longrightarrow Y(s) = \dfrac{2}{s-4}$

Therefore, $y(t) = \mathcal{L}^{-1}(Y(s)) = \mathcal{L}^{-1}\left(\dfrac{2}{s-4}\right) = 2 \cdot \mathcal{L}^{-1}\left(\dfrac{1}{s-4}\right) = 2e^{4t}$

(ii) Denote $\mathcal{L}(y(t)) = Y(s)$ and we Laplace both sides to get:

$\mathcal{L}(y''(t)) = s^2\mathcal{L}(y(t)) - sy(0) - y'(0) = s^2Y(s) - 2s$ and $\mathcal{L}(4y(t)) = 4Y(s)$

$y'' = 4y \longrightarrow s^2Y(s) - 2s = 4Y(s) \longrightarrow Y(s) = \dfrac{2s}{s^2-4}$

Therefore, $y(t) = \mathcal{L}^{-1}(Y(s)) = \mathcal{L}^{-1}\left(\dfrac{2s}{s^2-4}\right) = \mathcal{L}^{-1}\left(\dfrac{1}{s-2}\right) + \mathcal{L}^{-1}\left(\dfrac{1}{s+2}\right) = e^{2t} + e^{-2t}$

(iii) Denote $\mathcal{L}(y(t)) = Y(s)$ and we Laplace both sides to get:

$\mathcal{L}(y''(t)) = s^2\mathcal{L}(y(t)) - sy(0) - y'(0) = s^2Y(s) - 2s$ and $\mathcal{L}(-4y(t)) = -4Y(s)$

$y'' = -4y \longrightarrow s^2Y(s) - 2s = -4Y(s) \longrightarrow Y(s) = \dfrac{2s}{s^2+4}$

Therefore, $y(t) = \mathcal{L}^{-1}(Y(s)) = \mathcal{L}^{-1}\left(\dfrac{2s}{s^2+4}\right) = 2 \cdot \mathcal{L}^{-1}\left(\dfrac{s}{s^2+2^2}\right) = 2\cos(2t)$

(iv) Denote $\mathcal{L}(y(t)) = Y(s)$ and we Laplace both sides to get:

$\mathcal{L}(y''(t)) = s^2\mathcal{L}(y(t)) - sy(0) - y'(0) = s^2Y(s) - 2s - 2$ and $\mathcal{L}(-4y(t)) = -4Y(s)$

$y'' = -4y \longrightarrow s^2Y(s) - 2s - 2 = -4Y(s) \longrightarrow Y(s) = \dfrac{2s+2}{s^2+4}$

Therefore, $y(t) = \mathcal{L}^{-1}(Y(s)) = \mathcal{L}^{-1}\left(\dfrac{2s+2}{s^2+4}\right) = 2 \cdot \mathcal{L}^{-1}\left(\dfrac{s}{s^2+2^2}\right) + \mathcal{L}^{-1}\left(\dfrac{2}{s^2+2^2}\right) = 2\cos(2t) + \sin(2t)$

(v) Denote $\mathcal{L}(y(t)) = Y(s)$ and we Laplace both sides to get:

$\mathcal{L}(y''(t)) = s^2\mathcal{L}(y(t)) - sy(0) - y'(0) = s^2Y(s) - s - 6$ and $\mathcal{L}(6t) = \dfrac{6}{s^2}$

$y'' = 6t \longrightarrow s^2Y(s) - s - 6 = \dfrac{6}{s^2} \longrightarrow Y(s) = \dfrac{s^3 + 6s^2 + 6}{s^4}$

Therefore, $y(t) = \mathcal{L}^{-1}(Y(s)) = \mathcal{L}^{-1}\left(\dfrac{s^3+6s^2+6}{s^4}\right) = \mathcal{L}^{-1}\left(\dfrac{6}{s^4} + \dfrac{6}{s^2} + \dfrac{1}{s}\right) = \mathcal{L}^{-1}\left(\dfrac{3!}{s^4}\right) + 6 \cdot \mathcal{L}^{-1}\left(\dfrac{1!}{s^2}\right) + \mathcal{L}^{-1}\left(\dfrac{1}{s}\right) = t^3 + 6t + 1$

(vi) Denote $\mathcal{L}(y(t)) = Y(s)$ and we Laplace both sides to get:

$\mathcal{L}(y''(t) - 2y'(t)) = \mathcal{L}(y''(t)) - 2\mathcal{L}(y'(t)) = s^2\mathcal{L}(y(t)) - sy(0) - y'(0) - 2(\mathcal{L}(y(t)) - y(0)) = s^2Y(s) - 2sY(s)$

$\mathcal{L}(24t^2) = 24 \cdot \dfrac{2!}{s^3} = \dfrac{48}{s^3}$

$y'' - 2y' = 24t^2 \longrightarrow s^2Y(s) - 2sY(s) = \dfrac{48}{s^3} \longrightarrow Y(s) = \dfrac{48}{s^4(s-2)}$

Therefore, $y(t) = \mathcal{L}^{-1}(Y(s)) = \mathcal{L}^{-1}\left(\dfrac{48}{s^4(s-2)}\right) = \mathcal{L}^{-1}\left(-\dfrac{24}{s^4} - \dfrac{12}{s^3} - \dfrac{6}{s^2} - \dfrac{3}{s} + \dfrac{3}{s-2}\right)$

$= -4 \cdot \mathcal{L}^{-1}\left(\dfrac{3!}{s^4}\right) - 6 \cdot \mathcal{L}^{-1}\left(\dfrac{2!}{s^3}\right) - 6 \cdot \mathcal{L}^{-1}\left(\dfrac{1!}{s^2}\right) - 3 \cdot \mathcal{L}^{-1}\left(\dfrac{1}{s}\right) + 3 \cdot \mathcal{L}^{-1}\left(\dfrac{1}{s-2}\right) = -4t^3 - 6t^2 - 6t - 3 + 3e^{2t}$

Exercise 4.3.3a

Define the product $\begin{bmatrix} a \\ b \end{bmatrix} \odot \begin{bmatrix} x \\ y \end{bmatrix} = 2ax + 5by$

(i) Prove that $\odot$ is an inner product on $\mathbb{R}^2$

(ii) What is the length of the vector $\begin{bmatrix} 5 \\ -2 \end{bmatrix}$ with respect to $\odot$

(iii) Find the orthogonal complement of $W = \text{Span}\left\{\begin{bmatrix} 10 \\ 1 \end{bmatrix}\right\}$ with respect to $\odot$

Answer

(i) We check the three requirements of an inner product space:

$\begin{bmatrix} a \\ b \end{bmatrix} \odot \begin{bmatrix} x \\ y \end{bmatrix} = 2ax + 5by = 2xa + 5by = \begin{bmatrix} x \\ y \end{bmatrix} \odot \begin{bmatrix} a \\ b \end{bmatrix}$

$\left(r\begin{bmatrix} a \\ b \end{bmatrix} + s\begin{bmatrix} x \\ y \end{bmatrix}\right) \odot \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} ra+sx \\ rb+sy \end{bmatrix} \odot \begin{bmatrix} u \\ v \end{bmatrix} = 2(ra+sx)u + 5(rb+sy)v$

$= r(2au + 5bv) + s(2xu + 5yv) = r\left(\begin{bmatrix} a \\ b \end{bmatrix} \odot \begin{bmatrix} u \\ v \end{bmatrix}\right) + s\left(\begin{bmatrix} x \\ y \end{bmatrix} \odot \begin{bmatrix} u \\ v \end{bmatrix}\right)$

$\begin{bmatrix} a \\ b \end{bmatrix} \odot \begin{bmatrix} a \\ b \end{bmatrix} = 2a^2 + 5b^2 > 0$ when $ab \neq 0$

Therefore, $\odot$ is an inner product on $\mathbb{R}^2$

(ii) $\begin{bmatrix} 5 \\ -2 \end{bmatrix} \odot \begin{bmatrix} 5 \\ -2 \end{bmatrix} = 2 \cdot 5^2 + 5 \cdot (-2)^2 = 70$

(iii) $\begin{bmatrix} 10 \\ 1 \end{bmatrix} \odot \begin{bmatrix} x \\ y \end{bmatrix} = 20x + 5y = 0 \longrightarrow y = -4x$

Therefore, $W^{\perp} = \text{Span}\left\{\begin{bmatrix} 1 \\ -4 \end{bmatrix}\right\}$

Exercise 4.3.3b

Define the product $p(x) \odot q(x) = p(a)q(a) + p(b)q(b)$ for some fixed $a \neq b$ and $p, q \in \mathbb{P}_1(\mathbb{R})$

(i) Prove that $\odot$ is an inner product on $\mathbb{P}_1(\mathbb{R})$

(ii) What is the length of the vector $mx + c$ with respect to $\odot$

(iii) Find the orthogonal complement of $W = \text{Span}\{2x - 3\}$ with respect to $\odot$

Answer

(i) We check the three requirements of an inner product space:

$p(x) \odot q(x) = p(a)q(a) + p(b)q(b) = q(a)p(a) + q(b)p(b) = q(x) \odot p(x)$

$(rp + sq)(x) \odot w(x) = (rp+sq)(a)w(a) + (rp+sq)(b)w(b) = rp(a)w(a) + sq(a)w(a) + rp(b)w(b) + sq(b)w(b)$

$= r\left(p(a)w(a) + p(b)w(b)\right) + s\left(q(a)w(a) + q(b)w(b)\right) = rp(x) \odot w(x) + sq(x) \odot w(x)$

$p(x) \odot p(x) = p(a)p(a) + p(b)p(b) = [p(a)]^2 + [p(b)]^2 > 0$ since $p(a) = 0$ and $p(b) = 0$ cannot both hold when $a = b$

Therefore, $\odot$ is an inner product on $\mathbb{P}_1(\mathbb{R})$

(ii) $(mx + c) \odot (mx + c) = (ma + c)^2 + (mb + c)^2 = m^2(a^2 + b^2) + 2mc(a + b) + 2c^2$

(iii) $(2x - 3) \odot (mx + c) = (2a - 3)(ma + c) + (2b - 3)(mb + c) = m(2a^2 + 2b^2 - 3a - 3b) + c(2a + 2b - 6)$.

Setting this to zero gives us the orthogonal complement as $\text{Span}\left\{(6 - 2a - 2b)x + (2a^2 + 2b^2 - 3a - 3b)\right\}$

Exercise 4.3.3c

Define the product $f(x) \odot g(x) = \int_0^{2\pi} f(x)g(x)dx$ for $f, g \in C([0, 2\pi])$

(i) Prove that $\odot$ is an inner product on $C([0, 2\pi])$

(ii) What is the length of the vector $\sin(nx)$ with respect to $\odot$ (where n is some positive integer)

(iii) What is the length of the vector $\cos(nx)$ with respect to $\odot$ (where n is some positive integer)

(iv) Find an orthogonal basis for $W = \text{Span}\{\sin(x), \cos(x), x\}$

Answer

(i) We check the three requirements of an inner product space:

$f(x) \odot g(x) = \int_0^{2\pi} f(x)g(x)dx = \int_0^{2\pi} g(x)f(x)dx = g(x) \odot f(x)$

$(rf + sg)(x) \odot h(x) = \int_0^{2\pi} (rf + sg)(x)h(x)dx = r\int_0^{2\pi} f(x)h(x)dx + s\int_0^{2\pi} g(x)h(x)dx = rf(x) \odot h(x) + sg(x) \odot h(x)$

$f(x) \odot f(x) = \int_0^{2\pi} [f(x)]^2 dx > 0$ when $f(x) \neq$ the constant zero function.

Therefore, $\odot$ is an inner product on $C([0, 2\pi])$

(ii) $\sin(nx) \odot \sin(nx) = \int_0^{2\pi} [\sin(nx)]^2 dx = \int_0^{2\pi} \frac{1}{2} - \frac{1}{2}\cos(2nx)dx = \left[\frac{1}{2}x - \frac{1}{4n}\sin(2nx)\right]_0^{2\pi} = \pi$

(iii) $\cos(nx) \odot \cos(nx) = \int_0^{2\pi} [\cos(nx)]^2 dx = \int_0^{2\pi} \frac{1}{2} + \frac{1}{2}\cos(2nx)dx = \left[\frac{1}{2}x + \frac{1}{4n}\sin(2nx)\right]_0^{2\pi} = \pi$

(iv) We use Gram-Schmidt:

$\vec{w}_1 = \sin(x)$

$\vec{w}_2 = \cos(x) - \left(\frac{\cos(x) \odot \sin(x)}{\sin(x) \odot \sin(x)}\right)\sin(x)$

We calculate $\cos(x) \odot \sin(x) = \int_0^{2\pi} \cos(x)\sin(x)dx = \left[\frac{1}{2}\sin^2(x)\right]_0^{2\pi} = 0$

Therefore, $\vec{w}_2 = \cos(x) - 0\sin(x) = \cos(x)$

$\vec{w}_3 = x - \left(\frac{x \odot \sin(x)}{\sin(x) \odot \sin(x)}\right)\sin(x) - \left(\frac{x \odot \cos(x)}{\cos(x) \odot \cos(x)}\right)\cos(x)$

We calculate $x \odot \sin(x) = \int_0^{2\pi} x\sin(x)dx = [-x\cos(x) + \sin(x)]_0^{2\pi} = -2\pi.$

Similarly, $x \odot \cos(x) = \int_0^{2\pi} x\cos(x)dx = [x\sin(x) + \cos(x)]_0^{2\pi} = 2\pi$

Therefore, $\vec{w}_3 = x + \frac{2\pi}{\pi}\sin(x) - \frac{2\pi}{\pi}\cos(x) = x + \sin(x) - \cos(x)$

So W is spanned by the orthogonal set $\text{Span}\{\sin(x), \cos(x), x + \sin(x) - \cos(x)\}$

Exercise 4.3.3d

Define the product $(a_n) \odot (b_n) = \sum_{n=1}^{\infty} a_n b_n$ for $a_n, b_n \in \mathbb{S}$

(i) Prove that $\odot$ is an inner product on $\mathbb{S}$

(ii) What is the length of the vector $\left(\frac{1}{\sqrt{n!}}\right)$ with respect to $\odot$

(iii) Find an orthogonal basis for $W = \text{Span}\{2^{-n}, 3^{-n}, 4^{-n}\}$

Answer

(i) We check the three requirements of an inner product space:

$(a_n) \odot (b_n) = \sum_{n=1}^{\infty} a_n b_n = \sum_{n=1}^{\infty} b_n a_n = (b_n) \odot (a_n)$

$(r(a_n) + s(b_n)) \odot (c_n) = \sum_{n=1}^{\infty} (ra_n + sb_n)c_n = r\sum_{n=1}^{\infty} a_n c_n + s\sum_{n=1}^{\infty} b_n c_n = r(a_n) \odot (c_n) + s(b_n) \odot (c_n)$

$(a_n) \odot (a_n) = \sum_{n=1}^{\infty} a_n^2 > 0$ when $(a_n) \neq$ the constant zero sequence.

Therefore, $\odot$ is an inner product on $\mathbb{S}$

(ii) $\left(\frac{1}{\sqrt{n!}}\right) \odot \left(\frac{1}{\sqrt{n!}}\right) = \sum_{n=1}^{\infty} \left(\frac{1}{\sqrt{n!}}\right)^2 = \sum_{n=1}^{\infty} \frac{1^n}{n!} = e^1 - 1 = e - 1$

(above we used the Taylor series for $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$)

(iii) We use Gram-Schmidt:

$\vec{w}_1 = 2^{-n}$

$\vec{w}_2 = 3^{-n} - \left(\frac{3^{-n} \odot 2^{-n}}{2^{-n} \odot 2^{-n}}\right) 2^{-n}$

We calculate in general that $a^{-n} \odot b^{-n} = \sum_{n=1}^{\infty} a^{-n} \cdot b^{-n} = \sum_{n=1}^{\infty} (ab)^{-n} = \frac{(ab)^{-1}}{1-(ab)^{-1}} = \frac{1}{ab-1}$

This gives $3^{-n} \odot 2^{-n} = \frac{1}{5}$ and $2^{-n} \odot 2^{-n} = \frac{1}{3}$

Therefore, $\vec{w}_2 = 3^{-n} - \frac{3}{5} \cdot 2^{-n}$

$\vec{w}_3 = 4^{-n} - \left(\frac{4^{-n} \odot 2^{-n}}{2^{-n} \odot 2^{-n}}\right) 2^{-n} - \left(\frac{4^{-n} \odot 3^{-n}}{3^{-n} \odot 3^{-n}}\right) 3^{-n}$

We see that $4^{-n} \odot 2^{-n} = \frac{1}{7}$ and recall $2^{-n} \odot 2^{-n} = \frac{1}{3}$

Similarly, we have $4^{-n} \odot \left(3^{-n} - \frac{3}{5} \cdot 2^{-n}\right) = (4^{-n} \odot 3^{-n}) - \frac{3}{5}(4^{-n} \odot 2^{-n}) = \frac{1}{11} - \frac{3}{5} \cdot \frac{1}{7} = \frac{2}{385}$

and

$\left(3^{-n} - \frac{3}{5} \cdot 2^{-n}\right) \odot \left(3^{-n} - \frac{3}{5} \cdot 2^{-n}\right) = (3^{-n} \odot 3^{-n}) - \frac{6}{5}(3^{-n} \odot 2^{-n}) + \frac{9}{25}(2^{-n} \odot 2^{-n}) = \frac{1}{8} - \frac{6}{5} \cdot \frac{1}{5} + \frac{9}{25} \cdot \frac{1}{3} = \frac{1}{200}$

Therefore, $\vec{w}_3 = 4^{-n} - \frac{3}{7} \cdot 2^{-n} - \frac{80}{77}\left(3^{-n} - \frac{3}{5} \cdot 2^{-n}\right) = 4^{-n} - \frac{80}{77} \cdot 3^{-n} - \frac{81}{77} \cdot 2^{-n}$

So W is spanned by the orthogonal set $\text{Span}\left\{2^{-n}, 3^{-n} - \frac{3}{5} \cdot 2^{-n}, 4^{-n} - \frac{80}{77} \cdot 3^{-n} - \frac{81}{77} \cdot 2^{-n}\right\}$

? Exercise 4.3.4a

A person prepares their own lunch each day with one of two protein options: chicken or fish. The day after eating fish, this person has an equal chance of bringing chicken or fish for their lunch. On days after eating chicken, this person has a 25% chance of bringing fish and 75% chance of bringing chicken for their lunch.

(i) If this person ate chicken for lunch on Monday, what is the probability they will eat fish on Thursday? (assume they work all weekdays)

(ii) After many years, what is the probability of this person eating chicken on a given day?

Answer

We begin by defining the vector $\vec{v} = \begin{bmatrix} \text{chicken probability} \\ \text{fish probability} \end{bmatrix}$ and finding the matrix $A = \begin{bmatrix} .75 & .50 \\ .25 & .50 \end{bmatrix}$

(i) On Monday $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, so on Thursday we find $A^3\vec{v}$

$$A^3\vec{v} = \begin{bmatrix} .672 & .656 \\ .328 & .344 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} .672 \\ .328 \end{bmatrix}$$

Therefore, there is a 32.8% chance they will eat fish on Thursday.

(ii) After many years, the vector $\vec{v}$ will approach the steady state vector $A\vec{v}_s = \vec{v}_s$

$$E_1 = \text{Null}(A - I) = \text{Null}\left(\begin{bmatrix} -.25 & .50 \\ .25 & -.50 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right\}$$

Thus, $\vec{v}_s = \begin{bmatrix} 2c \\ c \end{bmatrix}$ with $2c + c = 1$

Clearly, this means that $c = \frac{1}{3}$ and that the probability of eating chicken is $\frac{2}{3}$

Exercise 4.3.4b

A student signs up for an early morning class despite their inability to set an alarm. Each day they either make it on time to class or oversleep and miss it. On days this student oversleeps, they also stay up later and then have a 70% chance of oversleeping class on the next day. On days the student makes it to class, they have a 40% chance of oversleeping class the next day.

(i) If this person oversleeps class on Monday, what is the probability they will make it to class on time on Wednesday? (assume the class meets every day)

(ii) After many days of classes, what is the probability of this person oversleeping on the last day?

Answer

We begin by defining the vector $\vec{v} = \begin{bmatrix} \text{on time} \\ \text{oversleep} \end{bmatrix}$ and finding the matrix $A = \begin{bmatrix} .60 & .30 \\ .40 & .70 \end{bmatrix}$

(i) On Monday $\vec{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, so on Wednesday we find $A^2\vec{v}$

$$A^2\vec{v} = \begin{bmatrix} .48 & .39 \\ .52 & .61 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} .39 \\ .61 \end{bmatrix}$$

Therefore, there is a 39% chance they will be on time on Wednesday.

(ii) After many days, the vector $\vec{v}$ will approach the steady state vector $A\vec{v}_s = \vec{v}_s$

$$E_1 = \text{Null}(A - I) = \text{Null}\left(\begin{bmatrix} -.52 & .39 \\ .52 & -.39 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 4 & -3 \\ 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 3 \\ 4 \end{bmatrix}\right\}$$

Thus, $\vec{v}_s = \begin{bmatrix} 3c \\ 4c \end{bmatrix}$ with $3c + 4c = 1$

Clearly, this means that $c = \frac{1}{7}$ and that the probability of oversleeping is $\frac{4}{7}$

Exercise 4.3.4c

A teacher asks students questions during class for extra credit. Each question may be categorized as difficult, normal or easy. After a difficult question, the next question will be difficult with probability 0.2, will be normal with probability 0.5, and will be easy with probability 0.3. After normal and easy questions, the next question will be difficult with probability 0.5, will be normal with probability 0.25, and will be easy with probability 0.25. If you walk into class late and the teacher asks a question, what is the probability that it will be difficult?

Answer

We begin by defining the vector $\vec{v} = \begin{bmatrix} \text{difficult question} \\ \text{normal question} \\ \text{easy question} \end{bmatrix}$ and finding the matrix $A = \begin{bmatrix} .20 & .50 & .50 \\ .50 & .25 & .25 \\ .30 & .25 & .25 \end{bmatrix}$

After many questions, the vector $\vec{v}$ will approach the steady state vector $A\vec{v}_s = \vec{v}_s$

$$E_1 = \text{Null}(A - I) = \text{Null}\left(\begin{bmatrix} -.80 & .50 & .50 \\ .50 & -.75 & .25 \\ .30 & .25 & -.75 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -\frac{10}{7} \\ 0 & 1 & -\frac{9}{7} \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 10 \\ 9 \\ 7 \end{bmatrix}\right\}$$

Thus, $\vec{v}_s = \begin{bmatrix} 10c \\ 9c \\ 7c \end{bmatrix}$ with $10c + 9c + 7c = 1$

Clearly, this means that $c = \frac{1}{26}$ and that the probability of getting a difficult question is $\frac{10}{26} = \frac{5}{13}$

Exercise 4.3.4*d*

A meme stock moves up, down, or stays even every day for no apparent reason. The day after it goes up, it goes up again 30% of the time, down 50% of the time, and stays even 20% of the time. The day after it goes down, it goes up 50% of the time, down 20% of the time, and stays even 30% of the time. The day after it stays even, it goes up 50% of the time, and down 50% of the time.

(i) What is the probability that it goes up on a randomly selected day next year?

(ii) What is the probability that it goes down on a randomly selected day next year?

Answer

We begin by defining the vector $\vec{v} = \begin{bmatrix} \text{moves up} \\ \text{moves down} \\ \text{stays even} \end{bmatrix}$ and finding the matrix $A = \begin{bmatrix} .30 & .50 & .50 \\ .50 & .20 & .50 \\ .20 & .30 & 0 \end{bmatrix}$

After many days, the vector $\vec{v}$ will approach the steady state vector $A\vec{v}_s = \vec{v}_s$

$$E_1 = \text{Null}(A - I) = \text{Null}\left(\begin{bmatrix} -.70 & .50 & .50 \\ .50 & -.80 & .50 \\ .20 & .30 & -1 \end{bmatrix}\right) = \text{Null}\left(\begin{bmatrix} 1 & 0 & -\frac{65}{31} \\ 0 & 1 & -\frac{60}{31} \\ 0 & 0 & 0 \end{bmatrix}\right) = \text{Span}\left\{\begin{bmatrix} 65 \\ 60 \\ 31 \end{bmatrix}\right\}$$

Thus, $\vec{v}_s = \begin{bmatrix} 65c \\ 60c \\ 31c \end{bmatrix}$ with $65c + 60c + 31c = 156$

Clearly, this means that $c = \frac{1}{156}$ and that the probability of going up is $\frac{65}{156} = \frac{5}{12}$ and going down is $\frac{60}{156} = \frac{5}{13}$

4.3: Practice Set is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by LibreTexts.

Index

A

accessible
- InfoPage

addition
- 1.3.1: Vector Algebra and Geometry
- 1.6.1: Defining Subspaces
- 4.1.1: General Vector Spaces
- 4.1.3: Practice Set

adjugate formula
- 2.6.2: Cramer's Rule and Adjugates

algebra
- 1.1.2: Parametric Descriptions
- 1.1.3: Practice Set
- 1.2.1: Elimination and Augmented Matrices
- 1.3.1: Vector Algebra and Geometry
- 1.5.2: Understanding Linear Independence
- 2.2.1: Linear Transformations
- 2.2.3: Practice Set
- 2.4.3: Practice Set

algorithm
- 4.3.5: Page Rank

angle
- 3.4.1: Dot Products and Orthogonality

approximate solution
- 3.5.2: Least Squares Solutions

area
- 2.7.1: Determinants for Area and Volume
- 2.7.2: Practice Set

augmented matrix
- 1.2.1: Elimination and Augmented Matrices
- 1.2.2: Echelon Forms
- 1.3.2: Vector Equations and Spans
- 1.4: Matrix Equations
- 1.4.1: Matrix Equations, Spans, and Consistency

axioms
- 4.1.1: General Vector Spaces

B

bases
- 4.1.2: Calculations in Vector Spaces

basis
- 1.7.1: Defining Basis
- 1.7.4: Practice Set
- 2.4.2: Invertibility of Matrices
- 3.1.2: Eigenspaces
- 3.2.1: Coordinatization and Similar Matrices
- 3.2.2: Diagonalization
- 3.4.3: Practice Set
- 3.5.1: Orthogonal Decomposition
- 4.1.3: Practice Set
- 4.2.2: Coordinatization
- 4.3.3: Inner Product Spaces

basis theorem
- 1.7.2: Finding a Basis of a Subspace

basisvectors
- 3.6.1: Orthogonal Sets and Projection

block diagonalization
- 3.3: Geometry of Eigenvalues

C

calculation
- 3.5.3: Practice Set

calculations
- 2.4.3: Practice Set

calculus
- 4.3.6: Practice Set

characteristic polynomial
- 3.1.3: Characteristic Polynomial
- 3.1.4: Practice Set
- 3.2.3: Practice Set
- 3.3.1: Geometry of Diagonalizable Matrices
- 3.3.2: Complex Eigenvalues and Eigenvectors
- 3.3.4: Practice Set

characteristic_polynomial
- 3.2.2: Diagonalization

closest point
- 3.4: Dot Products and Orthogonal Complements

closure properties
- 1.6.1: Defining Subspaces

closure_properties
- 1.6.3: Practice Set

codomain
- 2.1.2: Matrix Transformations
- 2.3.2: Transformation Composition and Matrix Multiplication

coefficients
- 1.3.1: Vector Algebra and Geometry
- 1.5.3: Practice Set

cofactor
- 2.6: Cofactor Expansions
- 2.6.1: Cofactor Expansion

cofactor expansions
- 2.6: Cofactor Expansions

collaborative
- InfoPage

color_space
- 1.1.1: Lines, Planes, and Space

column space
- 1.6.2: Common Subspaces
- 1.7.3: Rank and Nullity
- 2.1.1: Matrices as Functions
- 2.1.2: Matrix Transformations
- 2.4.2: Invertibility of Matrices
- 3.4.2: Orthogonal Complements

column_space
- 3.4.3: Practice Set

columns
- 1.4.1: Matrix Equations, Spans, and Consistency

columnspace
- 4.2.2: Coordinatization

community
- InfoPage

complement
- 3.4.2: Orthogonal Complements

complex eigenvalues
- 3.3: Geometry of Eigenvalues

complex functions
- 4.3.2: Famous Transformations

complex numbers
- 3.3.2: Complex Eigenvalues and Eigenvectors

composition
- 2.3: Matrix Algebra

computational complexity
- 2.5.1: Definition of Determinant

consistency
- 1.3.3: Practice Set
- 1.4.3: Practice Set
- 2.4.2: Invertibility of Matrices

consistent_inconsistent
- 1.2.4: Practice Set

coordinate geometry
- 1.1.3: Practice Set

coordinate_systems
- 1.1.1: Lines, Planes, and Space

coordinates
- 1.1.2: Parametric Descriptions
- 3.2.1: Coordinatization and Similar Matrices

cosine_law
- 3.4.1: Dot Products and Orthogonality

Cramer's rule
- 2.6.2: Cramer's Rule and Adjugates

CramersRule
- 2.6.3: Practice Set

criteria
- 1.5.2: Understanding Linear Independence

curriculumDevelopment
- InfoPage

D

data modeling
- 3.5.2: Least Squares Solutions

data_representation
- 1.1.1: Lines, Planes, and Space

derivatives
- 4.3.1: Differential Equations

determinant
- 2.5: Determinants- Definition
- 2.5.1: Definition of Determinant
- 2.5.2: Properties of the Determinant
- 2.5.3: Practice Set
- 2.6.1: Cofactor Expansion
- 3.1.3: Characteristic Polynomial
- 3.1.4: Practice Set
- 3.3.2: Complex Eigenvalues and Eigenvectors

determinants
- 2.3.4: Practice Set
- 2.6.2: Cramer's Rule and Adjugates
- 2.6.3: Practice Set
- 2.7.1: Determinants for Area and Volume
- 2.7.2: Practice Set

diagonal
- 2.5: Determinants- Definition

diagonal matrices
- 3.3.1: Geometry of Diagonalizable Matrices

diagonal matrix
- 3.3.4: Practice Set

diagonal_matrix
- 3.2.2: Diagonalization

diagonalizable
- 3.2.2: Diagonalization
- 3.3.3: Geometry of Matrices with Complex Eigenvalues

diagonalization
- 3.3.1: Geometry of Diagonalizable Matrices
- 3.3.4: Practice Set

differential equations
- 4.3: Examples and Applications
- 4.3.2: Famous Transformations

differential_equations
- 4.3.1: Differential Equations

differentialequation
- 4.3.6: Practice Set

differentiation
- 4.3.6: Practice Set

dilation
- 2.1.1: Matrices as Functions

dimension
1.7.1: Defining Basis
1.7.2: Finding a Basis of a Subspace
1.7.4: Practice Set
3.4.3: Practice Set
4.2.3: Practice Set
dimensionality
4.1.2: Calculations in Vector Spaces
dimensions
1.1.1: Lines, Planes, and Space
1.7.3: Rank and Nullity
3.2.1: Coordinatization and Similar Matrices
4.2.2: Coordinatization
direction
3.4.1: Dot Products and Orthogonality
distance
3.5.1: Orthogonal Decomposition
4.3.3: Inner Product Spaces
distance minimization
3.5.2: Least Squares Solutions
domain
2.1.2: Matrix Transformations
2.3.2: Transformation Composition and Matrix Multiplication
dot product
3.4: Dot Products and Orthogonal Complements
4.3: Examples and Applications
dot_product
3.4.1: Dot Products and Orthogonality
dotproduct
4.3.3: Inner Product Spaces

E

echelon form
1.2.3: Free Variables
2.3.4: Practice Set
2.5.3: Practice Set
echelon_form
1.2.4: Practice Set
1.3.3: Practice Set
education
TitlePage
InfoPage
eigenspace
3.1: Eigenvalues and Eigenvectors
eigenspaces
3.1.2: Eigenspaces
3.1.4: Practice Set
3.2.3: Practice Set
4.2.2: Coordinatization
eigenvalue
3.1: Eigenvalues and Eigenvectors
3.1.2: Eigenspaces
eigenvalue theorem
3.1.3: Characteristic Polynomial
eigenvalues
3.1.1: Eigenvalues and Eigenvectors Definitions
3.1.3: Characteristic Polynomial
3.1.4: Practice Set
3.2.2: Diagonalization
3.2.3: Practice Set
3.3.1: Geometry of Diagonalizable Matrices
3.3.2: Complex Eigenvalues and Eigenvectors
3.3.3: Geometry of Matrices with Complex Eigenvalues
3.3.4: Practice Set
4.2.2: Coordinatization
4.2.3: Practice Set
4.3: Examples and Applications
4.3.2: Famous Transformations
eigenvector
3.1: Eigenvalues and Eigenvectors
eigenvectors
3.1.2: Eigenspaces
3.1.3: Characteristic Polynomial
3.1.4: Practice Set
3.2.2: Diagonalization
3.2.3: Practice Set
3.3.1: Geometry of Diagonalizable Matrices
3.3.2: Complex Eigenvalues and Eigenvectors
3.3.4: Practice Set
4.3.2: Famous Transformations
elementary matrices
2.5.2: Properties of the Determinant
elimination method
1.2: Row Reduction
1.2.1: Elimination and Augmented Matrices
equations
2.3.4: Practice Set
Euclidean_plane
1.1.1: Lines, Planes, and Space
examples of linear transformations
2.2.1: Linear Transformations
exercise
1.3.3: Practice Set
2.6.3: Practice Set
exercises
1.5.3: Practice Set
2.1.3: Practice Set
2.7.2: Practice Set
expansion
2.6.1: Cofactor Expansion
expectation
4.3.4: Markov Chains
exponentialfunction
4.3.6: Practice Set

F

Fourier transformation
4.3.2: Famous Transformations
free_variable
1.2.4: Practice Set
1.3.3: Practice Set
free_variables
1.4.2: Solutions to Matrix Equations
function
2.1.1: Matrices as Functions
functions
2.3.1: Sums of Transformations and Matrices
2.3.2: Transformation Composition and Matrix Multiplication
4.3.1: Differential Equations
fundamental theorem of algebra
3.3.2: Complex Eigenvalues and Eigenvectors

G

Gaussian elimination
1.2.1: Elimination and Augmented Matrices
1.2.2: Echelon Forms
geometric argument
2.2.1: Linear Transformations
geometric interpretation
2.5.1: Definition of Determinant
3.5.2: Least Squares Solutions
geometric properties
3.3.1: Geometry of Diagonalizable Matrices
geometric representation
1.1.2: Parametric Descriptions
geometric shapes
1.1.3: Practice Set
geometric transformation
2.2.2: Matrix Representations
geometric_interpretation
1.4.2: Solutions to Matrix Equations
geometric_spaces
1.1.1: Lines, Planes, and Space
geometry
1.3.1: Vector Algebra and Geometry
1.3.2: Vector Equations and Spans
2.7.1: Determinants for Area and Volume
2.7.2: Practice Set
3.2.3: Practice Set
3.3.2: Complex Eigenvalues and Eigenvectors
Google
4.3.5: Page Rank
gram schmidt
3.6.2: Gram-Schmidt Process
GramSchmidt
4.3.3: Inner Product Spaces
graphing
1.1.3: Practice Set

H

h_value
1.2.4: Practice Set
homogeneous
1.4: Matrix Equations
homogeneous_equations
1.4.2: Solutions to Matrix Equations
homogeneous_solutions
1.4.3: Practice Set

I

identity
2.1.1: Matrices as Functions
identity matrix
2.2: Linear Transformations
2.3.3: Matrix Inverses
2.5.2: Properties of the Determinant
identity transformation
2.2.2: Matrix Representations
2.4.1: Invertibility of Linear Transformations
identitytransformation
4.2.1: Linear Transformations
importance
4.3.5: Page Rank
independence
3.5.1: Orthogonal Decomposition
infinite solutions
1.1.2: Parametric Descriptions
inhomogeneous
1.4: Matrix Equations
1.4.2: Solutions to Matrix Equations
initial_conditions
4.3.1: Differential Equations
injective
2.1.2: Matrix Transformations
inner product
4.3: Examples and Applications
innerproduct
4.3.3: Inner Product Spaces
input vector
2.3.1: Sums of Transformations and Matrices
2.3.2: Transformation Composition and Matrix Multiplication
integral calculus
4.3.2: Famous Transformations

integration
4.3.3: Inner Product Spaces
interactive
InfoPage
internet
4.3.5: Page Rank
intersection
1.1.3: Practice Set
intersections
1.1.2: Parametric Descriptions
introduction
TitlePage
inverse
2.3.4: Practice Set
2.4.3: Practice Set
inverse matrix
2.4: Invertibility
inverse transformation
2.4: Invertibility
invertibility
2.5.2: Properties of the Determinant
2.5.3: Practice Set
invertibility property
2.5: Determinants- Definition
invertible
2.4.2: Invertibility of Matrices
2.5.3: Practice Set
invertible matrices
2.3.3: Matrix Inverses
2.4: Invertibility
2.5.2: Properties of the Determinant
2.6.2: Cramer's Rule and Adjugates
3.2.1: Coordinatization and Similar Matrices
invertible matrix
2.4.1: Invertibility of Linear Transformations
3.3.4: Practice Set
invertible transformation
2.4.1: Invertibility of Linear Transformations

K

kernel
4.2.1: Linear Transformations
4.2.3: Practice Set
4.3.1: Differential Equations
4.3.6: Practice Set

L

Laplace Transformation
4.3.2: Famous Transformations
LaplaceTransform
4.3.6: Practice Set
LibreTexts
InfoPage
linear
3.5.1: Orthogonal Decomposition
linear algebra
1.2.2: Echelon Forms
1.2.3: Free Variables
1.4.1: Matrix Equations, Spans, and Consistency
1.6.1: Defining Subspaces
1.6.3: Practice Set
1.7.2: Finding a Basis of a Subspace
1.7.3: Rank and Nullity
1.7.4: Practice Set
2.1.1: Matrices as Functions
2.1.3: Practice Set
2.2.2: Matrix Representations
2.3.2: Transformation Composition and Matrix Multiplication
2.3.3: Matrix Inverses
2.3.4: Practice Set
2.4.3: Practice Set
2.5.1: Definition of Determinant
2.5.3: Practice Set
2.6.1: Cofactor Expansion
2.6.2: Cramer's Rule and Adjugates
2.7.1: Determinants for Area and Volume
3.1.1: Eigenvalues and Eigenvectors Definitions
3.1.2: Eigenspaces
3.1.3: Characteristic Polynomial
3.1.4: Practice Set
3.2.1: Coordinatization and Similar Matrices
3.2.3: Practice Set
3.3.1: Geometry of Diagonalizable Matrices
3.3.4: Practice Set
3.4.2: Orthogonal Complements
3.5.2: Least Squares Solutions
3.5.3: Practice Set
3.6.2: Gram-Schmidt Process
4.3: Examples and Applications
4.3.2: Famous Transformations
linear combination
1.3.1: Vector Algebra and Geometry
2.1.2: Matrix Transformations
linear combinations
1.3.2: Vector Equations and Spans
1.5.2: Understanding Linear Independence
linear dependence
1.5.2: Understanding Linear Independence
2.5.2: Properties of the Determinant
2.5.3: Practice Set
linear equations
1.1.2: Parametric Descriptions
1.1.3: Practice Set
1.2.1: Elimination and Augmented Matrices
1.2.2: Echelon Forms
1.3.2: Vector Equations and Spans
1.4: Matrix Equations
linear independence
1.5.2: Understanding Linear Independence
1.6.2: Common Subspaces
1.7.1: Defining Basis
1.7.2: Finding a Basis of a Subspace
1.7.4: Practice Set
2.4.2: Invertibility of Matrices
2.5.3: Practice Set
3.5.2: Least Squares Solutions
3.6.2: Gram-Schmidt Process
linear systems
1.2.3: Free Variables
linear transformation
2.1.2: Matrix Transformations
2.2: Linear Transformations
2.2.1: Linear Transformations
2.3.2: Transformation Composition and Matrix Multiplication
2.4.1: Invertibility of Linear Transformations
2.7.2: Practice Set
4.3.2: Famous Transformations
linear transformation matrix
2.2.2: Matrix Representations
linear transformations
2.2.2: Matrix Representations
2.3.1: Sums of Transformations and Matrices
2.3.4: Practice Set
4.3: Examples and Applications
linear_algebra
1.2.4: Practice Set
1.3.3: Practice Set
1.4.2: Solutions to Matrix Equations
1.4.3: Practice Set
1.6.2: Common Subspaces
2.5.2: Properties of the Determinant
3.2.2: Diagonalization
4.1.1: General Vector Spaces
4.1.3: Practice Set
4.2.1: Linear Transformations
linear_combination
1.3.3: Practice Set
3.2.1: Coordinatization and Similar Matrices
linear_homogeneous
4.3.1: Differential Equations
linear_independence
3.2.1: Coordinatization and Similar Matrices
3.4.3: Practice Set
linear_transformation
2.2.3: Practice Set
linear_transformations
4.3.1: Differential Equations
LinearAlgebra
TitlePage
1.5.1: Defining Linear Independence
1.5.3: Practice Set
1.6.3: Practice Set
2.4.1: Invertibility of Linear Transformations
2.6.3: Practice Set
3.3.3: Geometry of Matrices with Complex Eigenvalues
3.6.1: Orthogonal Sets and Projection
4.1.2: Calculations in Vector Spaces
4.2.2: Coordinatization
4.3.3: Inner Product Spaces
4.3.6: Practice Set
linearcombination
2.2.3: Practice Set
3.6.1: Orthogonal Sets and Projection
linearcombinations
4.1.2: Calculations in Vector Spaces
lineardependence
1.5.1: Defining Linear Independence
1.5.3: Practice Set
linearindependence
1.5.1: Defining Linear Independence
1.5.3: Practice Set
linearity
4.2.1: Linear Transformations
linearlyindependent
4.1.2: Calculations in Vector Spaces
lineartransformation
4.2.3: Practice Set
4.3.6: Practice Set
linearTransformations
4.2.1: Linear Transformations
4.2.2: Coordinatization
links
4.3.5: Page Rank

M

markovchain
4.3.4: Markov Chains
math
TitlePage
3.6.1: Orthogonal Sets and Projection

mathematical modeling
4.3.6: Practice Set
mathematical reasoning
1.1.3: Practice Set
1.2.1: Elimination and Augmented Matrices
mathematical_problems
1.2.4: Practice Set
mathematical_proofs
4.1.1: General Vector Spaces
mathematicalmodeling
4.3.4: Markov Chains
mathematicalProof
2.6.3: Practice Set
mathematics
TitlePage
1.6.1: Defining Subspaces
1.6.3: Practice Set
1.7.2: Finding a Basis of a Subspace
2.7.2: Practice Set
3.1.1: Eigenvalues and Eigenvectors Definitions
3.4.1: Dot Products and Orthogonality
3.6.1: Orthogonal Sets and Projection
matrices
2.1.3: Practice Set
2.2.3: Practice Set
2.4.3: Practice Set
2.6.3: Practice Set
3.3.3: Geometry of Matrices with Complex Eigenvalues
3.3.4: Practice Set
4.1.3: Practice Set
matrix
1.2.3: Free Variables
1.4: Matrix Equations
1.4.1: Matrix Equations, Spans, and Consistency
1.4.3: Practice Set
1.5.3: Practice Set
1.6.2: Common Subspaces
1.7.3: Rank and Nullity
1.7.4: Practice Set
2.1.1: Matrices as Functions
2.3.4: Practice Set
2.4.2: Invertibility of Matrices
2.4.3: Practice Set
2.5.1: Definition of Determinant
2.5.3: Practice Set
2.6.1: Cofactor Expansion
2.7.1: Determinants for Area and Volume
3.1.1: Eigenvalues and Eigenvectors Definitions
3.1.2: Eigenspaces
3.1.3: Characteristic Polynomial
3.1.4: Practice Set
3.2.1: Coordinatization and Similar Matrices
3.2.3: Practice Set
3.3.2: Complex Eigenvalues and Eigenvectors
3.3.3: Geometry of Matrices with Complex Eigenvalues
3.4.2: Orthogonal Complements
3.5.1: Orthogonal Decomposition
3.5.3: Practice Set
4.2.3: Practice Set
matrix addition
2.3.1: Sums of Transformations and Matrices
matrix equation
3.5.2: Least Squares Solutions
matrix equations
2.6.2: Cramer's Rule and Adjugates
3.1.3: Characteristic Polynomial
matrix inverse
2.3.3: Matrix Inverses
matrix inversion
2.6.2: Cramer's Rule and Adjugates
matrix multiplication
1.4.1: Matrix Equations, Spans, and Consistency
2.2.2: Matrix Representations
2.3.1: Sums of Transformations and Matrices
2.3.2: Transformation Composition and Matrix Multiplication
2.3.3: Matrix Inverses
2.4.1: Invertibility of Linear Transformations
4.3: Examples and Applications
matrix operations
1.2.1: Elimination and Augmented Matrices
2.3.3: Matrix Inverses
2.3.4: Practice Set
3.3.1: Geometry of Diagonalizable Matrices
3.3.4: Practice Set
matrix properties
3.1.4: Practice Set
matrix representation
2.4.1: Invertibility of Linear Transformations
3.3.1: Geometry of Diagonalizable Matrices
matrix theory
2.3.3: Matrix Inverses
matrix transformation
2.1.2: Matrix Transformations
2.2: Linear Transformations
2.2.1: Linear Transformations
2.3.2: Transformation Composition and Matrix Multiplication
matrix transformations
2.2.2: Matrix Representations
matrix_operations
1.4.2: Solutions to Matrix Equations
3.2.2: Diagonalization
matrix_representation
2.2.3: Practice Set
matrixequations
1.5.1: Defining Linear Independence
minor
2.6: Cofactor Expansions
multiplication
1.3.1: Vector Algebra and Geometry

N

non_homogeneous
4.3.1: Differential Equations
nontrivial_solution
1.3.3: Practice Set
norm of vector
3.4: Dot Products and Orthogonal Complements
null space
1.6.2: Common Subspaces
1.6.3: Practice Set
1.7.1: Defining Basis
1.7.3: Rank and Nullity
2.4.2: Invertibility of Matrices
3.1.2: Eigenspaces
3.1.4: Practice Set
3.2.3: Practice Set
3.3.4: Practice Set
3.4.2: Orthogonal Complements
null_space
3.4.3: Practice Set
nullity
1.7.3: Rank and Nullity
1.7.4: Practice Set
2.5.3: Practice Set
nullspace
4.2.2: Coordinatization
number_line
1.1.1: Lines, Planes, and Space

O

OER
InfoPage
one
2.1.2: Matrix Transformations
onto
2.1.3: Practice Set
4.2.3: Practice Set
operations
2.6.1: Cofactor Expansion
optimization
3.6.2: Gram-Schmidt Process
ordered_tuples
1.1.1: Lines, Planes, and Space
orthogonal
3.4: Dot Products and Orthogonal Complements
3.4.2: Orthogonal Complements
3.5.1: Orthogonal Decomposition
3.5.3: Practice Set
orthogonal basis
3.6.2: Gram-Schmidt Process
orthogonal decomposition
3.5: Orthogonal Projection and Least Squares
orthogonal projection
3.5: Orthogonal Projection and Least Squares
orthogonal set
3.6.2: Gram-Schmidt Process
orthogonalbasis
3.6.1: Orthogonal Sets and Projection
orthogonality
3.4.1: Dot Products and Orthogonality
3.4.2: Orthogonal Complements
3.4.3: Practice Set
4.3.3: Inner Product Spaces
orthonormal
3.6.1: Orthogonal Sets and Projection
output vector
2.3.2: Transformation Composition and Matrix Multiplication

P

PageRank
4.3.5: Page Rank
parallelepiped
2.7: Determinants and Volumes
parallelogram
2.7.2: Practice Set
Parallelogram law
1.3: Vectors
parameters
1.1.2: Parametric Descriptions
parametric form
1.2.3: Free Variables
1.4: Matrix Equations
3.1.2: Eigenspaces
parametric_solution
1.2.4: Practice Set
parametric_vector_form
1.4.2: Solutions to Matrix Equations
1.4.3: Practice Set
pivot
1.2.2: Echelon Forms
pivot column
1.2.3: Free Variables
pivot positions
1.7.1: Defining Basis

pivots
2.1.3: Practice Set
2.4.2: Invertibility of Matrices
planes
1.1.3: Practice Set
point
1.3: Vectors
points
1.1.3: Practice Set
polynomial
4.3.3: Inner Product Spaces
polynomial roots
3.1.3: Characteristic Polynomial
populationdistribution
4.3.4: Markov Chains
probabilisticmodels
4.3.4: Markov Chains
probability
4.3: Examples and Applications
probabilitytheory
4.3.4: Markov Chains
processanalysis
4.3.4: Markov Chains
projection
2.1.1: Matrices as Functions
2.2.2: Matrix Representations
3.5.1: Orthogonal Decomposition
3.5.3: Practice Set
projections
3.6.1: Orthogonal Sets and Projection
3.6.2: Gram-Schmidt Process
proof
1.5.1: Defining Linear Independence
1.5.2: Understanding Linear Independence
2.6.1: Cofactor Expansion
3.3.3: Geometry of Matrices with Complex Eigenvalues
properties
2.6.1: Cofactor Expansion
2.7.1: Determinants for Area and Volume
4.1.1: General Vector Spaces
4.1.3: Practice Set
properties of determinants
2.5.1: Definition of Determinant
2.5.2: Properties of the Determinant
properties of linear transformations
2.2.1: Linear Transformations
2.3.1: Sums of Transformations and Matrices
properties of matrix
1.4.1: Matrix Equations, Spans, and Consistency

Q

QRCode
TitlePage

R

random surfer
4.3.5: Page Rank
randomwalk
4.3.4: Markov Chains
range
2.1.2: Matrix Transformations
4.2.1: Linear Transformations
4.2.3: Practice Set
rank
1.7.3: Rank and Nullity
1.7.4: Practice Set
real number
1.3.1: Vector Algebra and Geometry
real_numbers
1.1.1: Lines, Planes, and Space
4.1.1: General Vector Spaces
4.1.3: Practice Set
realmatrix
3.3.3: Geometry of Matrices with Complex Eigenvalues
reduced row echelon form
1.2.2: Echelon Forms
3.1.2: Eigenspaces
reduced_row_echelon_form
1.4.2: Solutions to Matrix Equations
reflection
2.1.1: Matrices as Functions
2.2.2: Matrix Representations
results
4.3.5: Page Rank
Rn
1.7.1: Defining Basis
row echelon form
1.2.2: Echelon Forms
1.7.2: Finding a Basis of a Subspace
2.3.3: Matrix Inverses
2.4.2: Invertibility of Matrices
2.5.1: Definition of Determinant
row operations
1.2.1: Elimination and Augmented Matrices
2.5.1: Definition of Determinant
2.5.2: Properties of the Determinant
2.6.2: Cramer's Rule and Adjugates
row reduction
1.2.2: Echelon Forms
1.2.3: Free Variables
1.3.2: Vector Equations and Spans
2.3.4: Practice Set
2.4.3: Practice Set
2.5.1: Definition of Determinant
3.5.2: Least Squares Solutions
row_reduction
1.4.2: Solutions to Matrix Equations
3.4.3: Practice Set
row_space
3.4.3: Practice Set
rowreduction
1.4.3: Practice Set
1.5.1: Defining Linear Independence
rows
1.4.1: Matrix Equations, Spans, and Consistency
RREF
1.4.3: Practice Set
1.5.1: Defining Linear Independence
1.5.3: Practice Set
1.7.1: Defining Basis
2.2.3: Practice Set
2.4.3: Practice Set
2.5.3: Practice Set
3.5.3: Practice Set
4.2.3: Practice Set

S

scalar
1.3.1: Vector Algebra and Geometry
scalar multiplication
1.6.1: Defining Subspaces
2.2.1: Linear Transformations
2.3.1: Sums of Transformations and Matrices
2.4.1: Invertibility of Linear Transformations
scalar_multiplication
4.1.1: General Vector Spaces
4.1.3: Practice Set
scalarmultiplication
1.6.3: Practice Set
scalars
4.1.2: Calculations in Vector Spaces
scalartransformation
4.2.1: Linear Transformations
scaling rows
2.6.2: Cramer's Rule and Adjugates
scaling transformations
3.3.1: Geometry of Diagonalizable Matrices
searching
4.3.5: Page Rank
set theory
1.6.1: Defining Subspaces
sets
1.6.3: Practice Set
signed volume
2.7.1: Determinants for Area and Volume
similarity
3.2.2: Diagonalization
solution
2.6.3: Practice Set
3.5.1: Orthogonal Decomposition
3.5.3: Practice Set
4.3.6: Practice Set
solution set
1.1.2: Parametric Descriptions
1.4: Matrix Equations
solution_set
1.4.2: Solutions to Matrix Equations
solution_types
1.2.4: Practice Set
solutions
1.2.3: Free Variables
1.4.1: Matrix Equations, Spans, and Consistency
1.4.3: Practice Set
1.5.1: Defining Linear Independence
1.5.3: Practice Set
1.7.3: Rank and Nullity
4.1.2: Calculations in Vector Spaces
4.3.1: Differential Equations
solve systems
1.2.1: Elimination and Augmented Matrices
solving systems
1.2.2: Echelon Forms
3.1.3: Characteristic Polynomial
span
1.3.2: Vector Equations and Spans
1.3.3: Practice Set
1.4: Matrix Equations
1.5.2: Understanding Linear Independence
1.5.3: Practice Set
1.7.1: Defining Basis
1.7.4: Practice Set
3.1.4: Practice Set
3.2.3: Practice Set
3.4.2: Orthogonal Complements
3.5.3: Practice Set
spanning set
1.7.2: Finding a Basis of a Subspace
spans
1.6.2: Common Subspaces
4.1.2: Calculations in Vector Spaces
square matrix
2.5.1: Definition of Determinant
square_matrices
2.5.2: Properties of the Determinant
standard matrices
2.3.1: Sums of Transformations and Matrices

standard matrix
- 2.2: Linear Transformations
- 2.2.2: Matrix Representations

standard_basis
- 3.2.1: Coordinatization and Similar Matrices

steadystate
- 4.3.4: Markov Chains

subset
- 1.6.1: Defining Subspaces

subspace
- 1.6.1: Defining Subspaces
- 1.7.1: Defining Basis
- 1.7.2: Finding a Basis of a Subspace
- 2.2.3: Practice Set
- 3.4.2: Orthogonal Complements
- 3.4.3: Practice Set
- 3.5.1: Orthogonal Decomposition
- 3.5.3: Practice Set
- 3.6.1: Orthogonal Sets and Projection
- 4.1.3: Practice Set
- 4.2.3: Practice Set
- 4.3.1: Differential Equations

subspaces
- 1.6.2: Common Subspaces
- 1.6.3: Practice Set
- 3.4.1: Dot Products and Orthogonality
- 4.2.1: Linear Transformations

superposition principle
- 2.2: Linear Transformations

swapping rows
- 2.6.2: Cramer's Rule and Adjugates

system of equations
- 1.2.3: Free Variables

system of linear equations
- 1.4.1: Matrix Equations, Spans, and Consistency

systems of equations
- 1.1.2: Parametric Descriptions
- 1.1.3: Practice Set
- 1.2.1: Elimination and Augmented Matrices
- 1.3.2: Vector Equations and Spans
- 1.5.1: Defining Linear Independence

systems of linear equations
- 2.3.3: Matrix Inverses

systems_of_equations
- 1.2.4: Practice Set
- 1.4.3: Practice Set

systemsOfEquations
- 2.6.3: Practice Set

T

technology
- InfoPage

theorem
- 1.6.2: Common Subspaces
- 2.4.2: Invertibility of Matrices
- 3.3.3: Geometry of Matrices with Complex Eigenvalues
- 4.1.1: General Vector Spaces

theorems
- 1.5.2: Understanding Linear Independence

theory
- 1.7.2: Finding a Basis of a Subspace

transformation
- 2.1.1: Matrices as Functions
- 2.2.1: Linear Transformations
- 2.3.2: Transformation Composition and Matrix Multiplication
- 2.7.1: Determinants for Area and Volume
- 2.7.2: Practice Set
- 3.3.2: Complex Eigenvalues and Eigenvectors
- 4.2.2: Coordinatization

transformations
- 2.1.3: Practice Set
- 2.2.3: Practice Set
- 2.3.1: Sums of Transformations and Matrices
- 2.4.3: Practice Set
- 2.6.1: Cofactor Expansion
- 3.2.3: Practice Set
- 3.3.3: Geometry of Matrices with Complex Eigenvalues

transitionmatrix
- 4.3.4: Markov Chains

translation
- 2.2: Linear Transformations
- 2.2.1: Linear Transformations

transpose
- 2.5: Determinants- Definition

triangle
- 2.7.2: Practice Set

trivial solution
- 1.4: Matrix Equations

true_false_statements
- 1.3.3: Practice Set

U

unique solution
- 1.2.2: Echelon Forms
- 1.2.3: Free Variables
- 2.4.1: Invertibility of Linear Transformations

unit vector
- 3.4: Dot Products and Orthogonal Complements

unit_vector
- 3.4.1: Dot Products and Orthogonality

V

variables
- 1.2.4: Practice Set

vector
- 1.3.1: Vector Algebra and Geometry
- 1.4: Matrix Equations
- 1.4.1: Matrix Equations, Spans, and Consistency
- 1.7.3: Rank and Nullity
- 1.7.4: Practice Set
- 2.1.1: Matrices as Functions
- 2.2.3: Practice Set
- 2.6.1: Cofactor Expansion
- 3.2.1: Coordinatization and Similar Matrices
- 3.3.3: Geometry of Matrices with Complex Eigenvalues
- 3.4.1: Dot Products and Orthogonality
- 3.4.2: Orthogonal Complements
- 3.5.1: Orthogonal Decomposition
- 3.5.3: Practice Set
- 3.6.1: Orthogonal Sets and Projection
- 4.1.2: Calculations in Vector Spaces

vector addition
- 2.4.1: Invertibility of Linear Transformations

vector equations
- 1.3.2: Vector Equations and Spans

vector form
- 3.1.2: Eigenspaces

vector space
- 1.6.1: Defining Subspaces
- 1.6.2: Common Subspaces
- 1.7.2: Finding a Basis of a Subspace
- 1.7.4: Practice Set
- 2.1.2: Matrix Transformations
- 4.1: General Vector Spaces and Subspaces
- 4.3: Examples and Applications
- 4.3.2: Famous Transformations

vector spaces
- 1.3.2: Vector Equations and Spans
- 1.5.2: Understanding Linear Independence
- 2.2.1: Linear Transformations
- 2.3.1: Sums of Transformations and Matrices
- 2.3.3: Matrix Inverses
- 3.6.2: Gram-Schmidt Process

vector_space
- 1.3.3: Practice Set
- 2.2.3: Practice Set
- 3.4.3: Practice Set
- 4.1.1: General Vector Spaces
- 4.1.3: Practice Set

vector_spaces
- 3.2.2: Diagonalization

vectoraddition
- 1.6.3: Practice Set

vectors
- 1.3: Vectors
- 1.3.1: Vector Algebra and Geometry
- 1.3.2: Vector Equations and Spans
- 1.4.3: Practice Set
- 1.5.1: Defining Linear Independence
- 1.5.2: Understanding Linear Independence
- 1.5.3: Practice Set
- 1.6.2: Common Subspaces
- 1.7.1: Defining Basis
- 2.3.4: Practice Set
- 2.4.3: Practice Set
- 2.7.1: Determinants for Area and Volume
- 4.1.1: General Vector Spaces
- 4.1.3: Practice Set

vectorspace
- 4.3.3: Inner Product Spaces

vectorSpaces
- 4.1.2: Calculations in Vector Spaces
- 4.2.1: Linear Transformations
- 4.2.2: Coordinatization

volume
- 2.7.2: Practice Set

volumes
- 2.7.1: Determinants for Area and Volume

W

web pages
- 4.3.5: Page Rank

Z

zero vector
- 1.6.1: Defining Subspaces

zero_vector
- 1.6.3: Practice Set
- 3.4.1: Dot Products and Orthogonality

zerotransformation
- 4.2.1: Linear Transformations

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Math 26: Introduction to Linear Algebra

Webpages: 114

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC-SA 4.0: 98.2% (112 pages)
- Undeclared: 1.8% (2 pages)

By Page

- Math 26: Introduction to Linear Algebra - *CC BY-NC-SA 4.0*
  - Front Matter - *CC BY-NC-SA 4.0*
    - TitlePage - *CC BY-NC-SA 4.0*
    - InfoPage - *CC BY-NC-SA 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *CC BY-NC-SA 4.0*
  - 1: Systems of Linear Equations - *CC BY-NC-SA 4.0*
    - 1.1: Systems of Linear Equations - *CC BY-NC-SA 4.0*
      - 1.1.1: Lines, Planes, and Space - *CC BY-NC-SA 4.0*
      - 1.1.2: Parametric Descriptions - *CC BY-NC-SA 4.0*
      - 1.1.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.2: Row Reduction - *CC BY-NC-SA 4.0*
      - 1.2.1: Elimination and Augmented Matrices - *CC BY-NC-SA 4.0*
      - 1.2.2: Echelon Forms - *CC BY-NC-SA 4.0*
      - 1.2.3: Free Variables - *CC BY-NC-SA 4.0*
      - 1.2.4: Practice Set - *CC BY-NC-SA 4.0*
    - 1.3: Vectors - *CC BY-NC-SA 4.0*
      - 1.3.1: Vector Algebra and Geometry - *CC BY-NC-SA 4.0*
      - 1.3.2: Vector Equations and Spans - *CC BY-NC-SA 4.0*
      - 1.3.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.4: Matrix Equations - *CC BY-NC-SA 4.0*
      - 1.4.1: Matrix Equations, Spans, and Consistency - *CC BY-NC-SA 4.0*
      - 1.4.2: Solutions to Matrix Equations - *CC BY-NC-SA 4.0*
      - 1.4.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.5: Linear Independence - *CC BY-NC-SA 4.0*
      - 1.5.1: Defining Linear Independence - *CC BY-NC-SA 4.0*
      - 1.5.2: Understanding Linear Independence - *CC BY-NC-SA 4.0*
      - 1.5.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.6: Subspaces - *CC BY-NC-SA 4.0*
      - 1.6.1: Defining Subspaces - *CC BY-NC-SA 4.0*
      - 1.6.2: Common Subspaces - *CC BY-NC-SA 4.0*
      - 1.6.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.7: Basis and Dimension - *CC BY-NC-SA 4.0*
      - 1.7.1: Defining Basis - *CC BY-NC-SA 4.0*
      - 1.7.2: Finding a Basis of a Subspace - *CC BY-NC-SA 4.0*
      - 1.7.3: Rank and Nullity - *CC BY-NC-SA 4.0*
      - 1.7.4: Practice Set - *CC BY-NC-SA 4.0*
  - 2: Linear Transformations and Matrix Algebra - *CC BY-NC-SA 4.0*
    - 2.1: Matrix Transformations - *CC BY-NC-SA 4.0*
      - 2.1.1: Matrices as Functions - *CC BY-NC-SA 4.0*
      - 2.1.2: Matrix Transformations - *CC BY-NC-SA 4.0*
      - 2.1.3: Practice Set - *CC BY-NC-SA 4.0*
    - 2.2: Linear Transformations - *CC BY-NC-SA 4.0*
      - 2.2.1: Linear Transformations - *CC BY-NC-SA 4.0*
      - 2.2.2: Matrix Representations - *CC BY-NC-SA 4.0*
      - 2.2.3: Practice Set - *CC BY-NC-SA 4.0*
    - 2.3: Matrix Algebra - *CC BY-NC-SA 4.0*
      - 2.3.1: Sums of Transformations and Matrices - *CC BY-NC-SA 4.0*
      - 2.3.2: Transformation Composition and Matrix Multiplication - *CC BY-NC-SA 4.0*
      - 2.3.3: Matrix Inverses - *CC BY-NC-SA 4.0*
      - 2.3.4: Practice Set - *CC BY-NC-SA 4.0*
    - 2.4: Invertibility - *CC BY-NC-SA 4.0*
      - 2.4.1: Invertibility of Linear Transformations - *CC BY-NC-SA 4.0*
      - 2.4.2: Invertibility of Matrices - *CC BY-NC-SA 4.0*

- 2.4.3: Practice Set - *CC BY-NC-SA 4.0*
- 2.5: Determinants- Definition - *CC BY-NC-SA 4.0*
  - 2.5.1: Definition of Determinant - *CC BY-NC-SA 4.0*
  - 2.5.2: Properties of the Determinant - *CC BY-NC-SA 4.0*
  - 2.5.3: Practice Set - *CC BY-NC-SA 4.0*
- 2.6: Cofactor Expansions - *CC BY-NC-SA 4.0*
  - 2.6.1: Cofactor Expansion - *CC BY-NC-SA 4.0*
  - 2.6.2: Cramer's Rule and Adjugates - *CC BY-NC-SA 4.0*
  - 2.6.3: Practice Set - *CC BY-NC-SA 4.0*
- 2.7: Determinants and Volumes - *CC BY-NC-SA 4.0*
  - 2.7.1: Determinants for Area and Volume - *CC BY-NC-SA 4.0*
  - 2.7.2: Practice Set - *CC BY-NC-SA 4.0*

3: Eigenvalues and Eigenvectors - *CC BY-NC-SA 4.0*

- 3.1: Eigenvalues and Eigenvectors - *CC BY-NC-SA 4.0*
  - 3.1.1: Eigenvalues and Eigenvectors Definitions - *CC BY-NC-SA 4.0*
  - 3.1.2: Eigenspaces - *CC BY-NC-SA 4.0*
  - 3.1.3: Characteristic Polynomial - *CC BY-NC-SA 4.0*
  - 3.1.4: Practice Set - *CC BY-NC-SA 4.0*
- 3.2: Diagonalization and Geometry of Eigenvalues - *CC BY-NC-SA 4.0*
  - 3.2.1: Coordinatization and Similar Matrices - *CC BY-NC-SA 4.0*
  - 3.2.2: Diagonalization - *CC BY-NC-SA 4.0*
  - 3.2.3: Practice Set - *CC BY-NC-SA 4.0*
- 3.3: Geometry of Eigenvalues - *CC BY-NC-SA 4.0*
  - 3.3.1: Geometry of Diagonalizable Matrices - *CC BY-NC-SA 4.0*
  - 3.3.2: Complex Eigenvalues and Eigenvectors - *CC BY-NC-SA 4.0*
  - 3.3.3: Geometry of Matrices with Complex Eigenvalues - *CC BY-NC-SA 4.0*
  - 3.3.4: Practice Set - *CC BY-NC-SA 4.0*
- 3.4: Dot Products and Orthogonal Complements - *CC BY-NC-SA 4.0*
  - 3.4.1: Dot Products and Orthogonality - *CC BY-NC-SA 4.0*
  - 3.4.2: Orthogonal Complements - *CC BY-NC-SA 4.0*
  - 3.4.3: Practice Set - *CC BY-NC-SA 4.0*
- 3.5: Orthogonal Projection and Least Squares - *CC BY-NC-SA 4.0*
  - 3.5.1: Orthogonal Decomposition - *CC BY-NC-SA 4.0*
  - 3.5.2: Least Squares Solutions - *CC BY-NC-SA 4.0*
  - 3.5.3: Practice Set - *CC BY-NC-SA 4.0*
- 3.6: Orthogonal Sets - *CC BY-NC-SA 4.0*
  - 3.6.1: Orthogonal Sets and Projection - *CC BY-NC-SA 4.0*
  - 3.6.2: Gram-Schmidt Process - *CC BY-NC-SA 4.0*
  - 3.6.3: Practice Set - *CC BY-NC-SA 4.0*

4: Vector Spaces - *CC BY-NC-SA 4.0*

- 4.1: General Vector Spaces and Subspaces - *CC BY-NC-SA 4.0*
  - 4.1.1: General Vector Spaces - *CC BY-NC-SA 4.0*
  - 4.1.2: Calculations in Vector Spaces - *CC BY-NC-SA 4.0*
  - 4.1.3: Practice Set - *CC BY-NC-SA 4.0*
- 4.2: Linear Transformations in General Vector Spaces - *CC BY-NC-SA 4.0*
  - 4.2.1: Linear Transformations - *CC BY-NC-SA 4.0*
  - 4.2.2: Coordinatization - *CC BY-NC-SA 4.0*
  - 4.2.3: Practice Set - *CC BY-NC-SA 4.0*
- 4.3: Examples and Applications - *CC BY-NC-SA 4.0*
  - 4.3.1: Differential Equations - *CC BY-NC-SA 4.0*
  - 4.3.2: Famous Transformations - *CC BY-NC-SA 4.0*
  - 4.3.3: Inner Product Spaces - *CC BY-NC-SA 4.0*
  - 4.3.4: Markov Chains - *CC BY-NC-SA 4.0*
  - 4.3.5: Page Rank - *CC BY-NC-SA 4.0*
  - 4.3.6: Practice Set - *CC BY-NC-SA 4.0*

Back Matter - *CC BY-NC-SA 4.0*

- Index - *CC BY-NC-SA 4.0*
- Glossary - *CC BY-NC-SA 4.0*
- Detailed Licensing - *CC BY-NC-SA 4.0*
- Detailed Licensing - *Undeclared*

Detailed Licensing

Overview

Title: Math 26: Introduction to Linear Algebra

Webpages: 114

Applicable Restrictions: Noncommercial

All licenses found:

- CC BY-NC-SA 4.0: 98.2% (112 pages)
- Undeclared: 1.8% (2 pages)

By Page

- Math 26: Introduction to Linear Algebra - *CC BY-NC-SA 4.0*
  - Front Matter - *CC BY-NC-SA 4.0*
    - TitlePage - *CC BY-NC-SA 4.0*
    - InfoPage - *CC BY-NC-SA 4.0*
    - Table of Contents - *Undeclared*
    - Licensing - *CC BY-NC-SA 4.0*
  - 1: Systems of Linear Equations - *CC BY-NC-SA 4.0*
    - 1.1: Systems of Linear Equations - *CC BY-NC-SA 4.0*
      - 1.1.1: Lines, Planes, and Space - *CC BY-NC-SA 4.0*
      - 1.1.2: Parametric Descriptions - *CC BY-NC-SA 4.0*
      - 1.1.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.2: Row Reduction - *CC BY-NC-SA 4.0*
      - 1.2.1: Elimination and Augmented Matrices - *CC BY-NC-SA 4.0*
      - 1.2.2: Echelon Forms - *CC BY-NC-SA 4.0*
      - 1.2.3: Free Variables - *CC BY-NC-SA 4.0*
      - 1.2.4: Practice Set - *CC BY-NC-SA 4.0*
    - 1.3: Vectors - *CC BY-NC-SA 4.0*
      - 1.3.1: Vector Algebra and Geometry - *CC BY-NC-SA 4.0*
      - 1.3.2: Vector Equations and Spans - *CC BY-NC-SA 4.0*
      - 1.3.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.4: Matrix Equations - *CC BY-NC-SA 4.0*
      - 1.4.1: Matrix Equations, Spans, and Consistency - *CC BY-NC-SA 4.0*
      - 1.4.2: Solutions to Matrix Equations - *CC BY-NC-SA 4.0*
      - 1.4.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.5: Linear Independence - *CC BY-NC-SA 4.0*
      - 1.5.1: Defining Linear Independence - *CC BY-NC-SA 4.0*
      - 1.5.2: Understanding Linear Independence - *CC BY-NC-SA 4.0*
      - 1.5.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.6: Subspaces - *CC BY-NC-SA 4.0*
      - 1.6.1: Defining Subspaces - *CC BY-NC-SA 4.0*
      - 1.6.2: Common Subspaces - *CC BY-NC-SA 4.0*
      - 1.6.3: Practice Set - *CC BY-NC-SA 4.0*
    - 1.7: Basis and Dimension - *CC BY-NC-SA 4.0*
      - 1.7.1: Defining Basis - *CC BY-NC-SA 4.0*
      - 1.7.2: Finding a Basis of a Subspace - *CC BY-NC-SA 4.0*
      - 1.7.3: Rank and Nullity - *CC BY-NC-SA 4.0*
      - 1.7.4: Practice Set - *CC BY-NC-SA 4.0*
  - 2: Linear Transformations and Matrix Algebra - *CC BY-NC-SA 4.0*
    - 2.1: Matrix Transformations - *CC BY-NC-SA 4.0*
      - 2.1.1: Matrices as Functions - *CC BY-NC-SA 4.0*
      - 2.1.2: Matrix Transformations - *CC BY-NC-SA 4.0*
      - 2.1.3: Practice Set - *CC BY-NC-SA 4.0*
    - 2.2: Linear Transformations - *CC BY-NC-SA 4.0*
      - 2.2.1: Linear Transformations - *CC BY-NC-SA 4.0*
      - 2.2.2: Matrix Representations - *CC BY-NC-SA 4.0*
      - 2.2.3: Practice Set - *CC BY-NC-SA 4.0*
    - 2.3: Matrix Algebra - *CC BY-NC-SA 4.0*
      - 2.3.1: Sums of Transformations and Matrices - *CC BY-NC-SA 4.0*
      - 2.3.2: Transformation Composition and Matrix Multiplication - *CC BY-NC-SA 4.0*
      - 2.3.3: Matrix Inverses - *CC BY-NC-SA 4.0*
      - 2.3.4: Practice Set - *CC BY-NC-SA 4.0*
    - 2.4: Invertibility - *CC BY-NC-SA 4.0*
      - 2.4.1: Invertibility of Linear Transformations - *CC BY-NC-SA 4.0*
      - 2.4.2: Invertibility of Matrices - *CC BY-NC-SA 4.0*

- 2.4.3: Practice Set - *CC BY-NC-SA 4.0*
- 2.5: Determinants- Definition - *CC BY-NC-SA 4.0*
  - 2.5.1: Definition of Determinant - *CC BY-NC-SA 4.0*
  - 2.5.2: Properties of the Determinant - *CC BY-NC-SA 4.0*
  - 2.5.3: Practice Set - *CC BY-NC-SA 4.0*
- 2.6: Cofactor Expansions - *CC BY-NC-SA 4.0*
  - 2.6.1: Cofactor Expansion - *CC BY-NC-SA 4.0*
  - 2.6.2: Cramer's Rule and Adjugates - *CC BY-NC-SA 4.0*
  - 2.6.3: Practice Set - *CC BY-NC-SA 4.0*
- 2.7: Determinants and Volumes - *CC BY-NC-SA 4.0*
  - 2.7.1: Determinants for Area and Volume - *CC BY-NC-SA 4.0*
  - 2.7.2: Practice Set - *CC BY-NC-SA 4.0*

3: Eigenvalues and Eigenvectors - *CC BY-NC-SA 4.0*

- 3.1: Eigenvalues and Eigenvectors - *CC BY-NC-SA 4.0*
  - 3.1.1: Eigenvalues and Eigenvectors Definitions - *CC BY-NC-SA 4.0*
  - 3.1.2: Eigenspaces - *CC BY-NC-SA 4.0*
  - 3.1.3: Characteristic Polynomial - *CC BY-NC-SA 4.0*
  - 3.1.4: Practice Set - *CC BY-NC-SA 4.0*
- 3.2: Diagonalization and Geometry of Eigenvalues - *CC BY-NC-SA 4.0*
  - 3.2.1: Coordinatization and Similar Matrices - *CC BY-NC-SA 4.0*
  - 3.2.2: Diagonalization - *CC BY-NC-SA 4.0*
  - 3.2.3: Practice Set - *CC BY-NC-SA 4.0*
- 3.3: Geometry of Eigenvalues - *CC BY-NC-SA 4.0*
  - 3.3.1: Geometry of Diagonalizable Matrices - *CC BY-NC-SA 4.0*
  - 3.3.2: Complex Eigenvalues and Eigenvectors - *CC BY-NC-SA 4.0*
  - 3.3.3: Geometry of Matrices with Complex Eigenvalues - *CC BY-NC-SA 4.0*
  - 3.3.4: Practice Set - *CC BY-NC-SA 4.0*
- 3.4: Dot Products and Orthogonal Complements - *CC BY-NC-SA 4.0*
  - 3.4.1: Dot Products and Orthogonality - *CC BY-NC-SA 4.0*
  - 3.4.2: Orthogonal Complements - *CC BY-NC-SA 4.0*
  - 3.4.3: Practice Set - *CC BY-NC-SA 4.0*
- 3.5: Orthogonal Projection and Least Squares - *CC BY-NC-SA 4.0*
  - 3.5.1: Orthogonal Decomposition - *CC BY-NC-SA 4.0*
  - 3.5.2: Least Squares Solutions - *CC BY-NC-SA 4.0*
  - 3.5.3: Practice Set - *CC BY-NC-SA 4.0*
- 3.6: Orthogonal Sets - *CC BY-NC-SA 4.0*
  - 3.6.1: Orthogonal Sets and Projection - *CC BY-NC-SA 4.0*
  - 3.6.2: Gram-Schmidt Process - *CC BY-NC-SA 4.0*
  - 3.6.3: Practice Set - *CC BY-NC-SA 4.0*

4: Vector Spaces - *CC BY-NC-SA 4.0*

- 4.1: General Vector Spaces and Subspaces - *CC BY-NC-SA 4.0*
  - 4.1.1: General Vector Spaces - *CC BY-NC-SA 4.0*
  - 4.1.2: Calculations in Vector Spaces - *CC BY-NC-SA 4.0*
  - 4.1.3: Practice Set - *CC BY-NC-SA 4.0*
- 4.2: Linear Transformations in General Vector Spaces - *CC BY-NC-SA 4.0*
  - 4.2.1: Linear Transformations - *CC BY-NC-SA 4.0*
  - 4.2.2: Coordinatization - *CC BY-NC-SA 4.0*
  - 4.2.3: Practice Set - *CC BY-NC-SA 4.0*
- 4.3: Examples and Applications - *CC BY-NC-SA 4.0*
  - 4.3.1: Differential Equations - *CC BY-NC-SA 4.0*
  - 4.3.2: Famous Transformations - *CC BY-NC-SA 4.0*
  - 4.3.3: Inner Product Spaces - *CC BY-NC-SA 4.0*
  - 4.3.4: Markov Chains - *CC BY-NC-SA 4.0*
  - 4.3.5: Page Rank - *CC BY-NC-SA 4.0*
  - 4.3.6: Practice Set - *CC BY-NC-SA 4.0*

Back Matter - *CC BY-NC-SA 4.0*

- Index - *CC BY-NC-SA 4.0*
- Glossary - *CC BY-NC-SA 4.0*
- Detailed Licensing - *CC BY-NC-SA 4.0*
- Detailed Licensing - *Undeclared*