MAT 149 TOPICS IN FINITE MATHEMATICS



Daniel Holz
SUNY Schenectady County Community College

LibreTexts™

SUNY SCCC MAT 149 Topics In Finite Mathematics

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

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

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



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

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

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

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

This text was compiled on 06/10/2025

TABLE OF CONTENTS

Licensing

1: Logic

- 1.1: Statements and Truth Values
  - 1.1.1: Introduction
  - 1.1.2: Statements and Quantifiers
  - 1.1.3: Compound Statements
- 1.2: Truth Tables
  - 1.2.1: Constructing Truth Tables
  - 1.2.2: Truth Tables for the Conditional and Biconditional
- 1.3: Conditional Statements
  - 1.3.1: Equivalent Statements
- 1.4: Deductive Arguments
  - 1.4.1: Evaluating Deductive Arguments with Truth Tables
  - 1.4.2: Forms of Valid Arguments

2: Functions

- 2.1: Functions
  - 2.1.1: Function Notation
  - 2.1.2: Domain and Range of a Function
  - 2.1.3: Piecewise-defined Functions
  - 2.1.4: Graphing Equations by Hand
- 2.2: Linear Functions
  - 2.2.1: Slope of a Line
    - 2.2.1E: Exercises
  - 2.2.2: Find the Equation of a Line
- 2.3: Linear Models
  - 2.3.1: Linear Models
  - 2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know)
- 2.4: Graphing Linear Functions
  - 2.4.1: Graphing Linear Equations
- 2.5: Linear Regression

3: Systems and Matrices

- 3.1: Systems of Linear Equations
  - 3.1.1: Solving a System of Equations Algebraically
- 3.2: Matrices
  - 3.2.1: Gaussian Elimination
- 3.3: Solving Using Matrices
  - 3.3.1: Gaussian Elimination
- 3.4: Matrix Algebra
  - 3.4.1: Matrix Arithmetic

- 3.4.2: Multiplication of Matrices
- 3.4.3: The ijth Entry of a Product
- 3.5: Matrix Inverses
  - 3.5.1: The Identity and Inverses
  - 3.5.2: Finding the Inverse of a Matrix
  - 3.5.3: The Matrix Inverse
    - 3.5.3.1: Exercises 2.6
- 3.6: Linear Inequalities
  - 3.6.1: Solving Systems of Linear Inequalities (Two Variables)
- 3.7: Linear Programming
  - 3.7.1: Maximization Applications
    - 3.7.1.1: Maximization Applications (Exercises)
  - 3.7.2: Minimization Applications

4: Sets and Counting

- 4.1: Sets and Operations
  - 4.1.1: Basics of Sets
  - 4.1.2: Union, Intersection, and Complement
  - 4.1.3: Venn Diagrams
- 4.2: Number of Elements in a Set
  - 4.2.1: Cardinality
- 4.3: Multiplication Principle
  - 4.3.1: Tree Diagrams, Tables, and Outcomes
  - 4.3.2: The Multiplication Rule for Counting
- 4.4: Permutations and Combinations
  - 4.4.1: Permutations
  - 4.4.2: Combinations

5: Probability

- 5.1: Sample Spaces and Probability
  - 5.1.1: Basic Concepts
- 5.2: Probability Rules
  - 5.2.1: The Addition Rule for Probability
  - 5.2.2: Conditional Probability and the Multiplication Rule
- 5.3: Binomial Distributions
  - 5.3.1: The Binomial Distribution
- 5.4: Markov Chains
  - 5.4.1: Stochastic Matrices

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: Logic

1.1: Statements and Truth Values

1.1.1: Introduction

1.1.2: Statements and Quantifiers

1.1.3: Compound Statements

1.2: Truth Tables

1.2.1: Constructing Truth Tables

1.2.2: Truth Tables for the Conditional and Biconditional

1.3: Conditional Statements

1.3.1: Equivalent Statements

1.4: Deductive Arguments

1.4.1: Evaluating Deductive Arguments with Truth Tables

1.4.2: Forms of Valid Arguments

1: Logic is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

1.1: Statements and Truth Values

1.1: Statements and Truth Values is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

1.1.1: Introduction

What is logic? Logic is the study of reasoning, and it has applications in many fields, including philosophy, law, psychology, digital electronics, and computer science.

In law, constructing a well-reasoned, logical argument is extremely important. The main goal of arguments made by lawyers is to convince a judge and jury that their arguments are valid and well-supported by the facts of the case, so the case should be ruled in their favor. Think about Thurgood Marshall arguing for desegregation in front of the U.S. Supreme Court during the *Brown v. Board of Education of Topeka* lawsuit in 1954, or Ruth Bader Ginsburg arguing for equality in social security benefits for both men and women under the law during the mid-1970s. Both these great minds were known for the preparation and thoroughness of their logical legal arguments, which resulted in victories that advanced the causes they fought for. Thurgood Marshall and Ruth Bader Ginsburg would later become well respected justices on the U.S. Supreme Court themselves.



Figure 1.1.1.1: Logic is key to a well-reasoned argument, in both math and law. (credit: modification of work "Lady Justicia holding sword and scale bronze figurine with judge hammer on wooden table" by Jernej Furman/Flickr, CC BY 2.0)

In this chapter, we will explore how to construct well-reasoned logical arguments using varying structures. Your ability to form and comprehend logical arguments is a valuable tool in many areas of life, whether you're planning a dinner date, negotiating the purchase of a new car, or persuading your boss that you deserve a raise.

This page titled 1.1.1: Introduction is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **2.0: Introduction** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

1.1.2: Statements and Quantifiers



Figure 1.1.2.1: Construction of a logical argument, like that of a house, requires you to begin with the right parts. (credit: modification of work “Barn Raising” by Robert Stinnett/Flickr, CC BY 2.0)

Learning Objectives

After completing this section, you should be able to:

1. Identify logical statements.
2. Represent statements in symbolic form.
3. Negate statements in words.
4. Negate statements symbolically.
5. Translate negations between words and symbols.
6. Express statements with quantifiers of all, some, and none.
7. Negate statements containing quantifiers of all, some, and none.

Have you ever built a club house, tree house, or fort with your friends? If so, you and your friends likely started by gathering some tools and supplies to work with, such as hammers, saws, screwdrivers, wood, nails, and screws. Hopefully, at least one member of your group had some knowledge of how to use the tools correctly and helped to direct the construction project. After all, if your house isn't built on a strong foundation, it will be weak and could possibly fall apart during the next big storm. This same foundation is important in logic.

In this section, we will begin with the parts that make up all logical arguments. The building block of any logical argument is a logical statement, or simply a statement. A logical statement has the form of a complete sentence, and it must make a claim that can be identified as being true or false.

When making arguments, sometimes people make false claims. When evaluating the strength or validity of a logical argument, you must also consider the truth values, or the identification of true or false, of all the statements used to support the argument. While a false statement is still considered a logical statement, a strong logical argument starts with true statements.

Identifying Logical Statements



Figure 1.1.2.2: Not all roses are red. (credit: “assorted pink yellow white red roses macro” by ProFlowers/Flickr, CC BY 2.0)

An example of logical statement with a false truth value is, “All roses are red.” It is a logical statement because it has the form of a complete sentence and makes a claim that can be determined to be either true or false. It is a false statement because not all roses are red: some roses are red, but there are also roses that are pink, yellow, and white. Requests, questions, or directives may be complete sentences, but they are not logical statements because they cannot be determined to be true or false. For example, suppose someone said to you, “Please, sit down over there.” This request does not make a claim and it cannot be identified as true or false; therefore, it is not a logical statement.

? Exercise 1.1.2.1: Identifying Logical Statements

Determine whether each of the following sentences represents a logical statement. If it is a logical statement, determine whether it is true or false.

1. Tiger Woods won the Master’s golf championship at least five times.
2. Please, sit down over there.
3. All cats dislike dogs.

Answer

1. This is a logical statement because it is a complete sentence that makes a claim that can be identified as being true or false. As of 2021, this statement is true: Tiger Woods won the Master’s in 1997, 2001, 2002, 2005 and 2019.
2. This is not a logical statement because, although it is a complete sentence, this request does not make a claim that can be identified as being either true or false.
3. This is a logical statement because it is a complete sentence that makes a claim that can be identified as being true or false. This statement is false because some cats do like some dogs.

✔ Your Turn 1.1.2.1

Determine whether each of the following sentences represents a logical statement. If it is a logical statement, determine whether it is true or false.

1. The Buffalo Bills defeated the New York Giants in Super Bowl XXV
2. Michael Jackson’s album *Thriller* was released in 1982.
3. Would you like some coffee or tea?

Representing Statements in Symbolic Form

When analyzing logical arguments that are made of multiple logical statements, symbolic form is used to reduce the amount of writing involved. Symbolic form also helps visualize the relationship between the statements in a more concise way in order to determine the strength or validity of an argument. Each logical statement is represented symbolically as a single lowercase letter, usually starting with the letter p.

To begin, you will practice how to write a single logical statement in symbolic form. This skill will become more useful as you work with compound statements in later sections.

? Exercise 1.1.2.2: Representing Statements Using Symbolic Form

Write each of the following logical statements in symbolic form.

1. Barry Bonds holds the Major League Baseball record for total career home runs.
2. Some mammals live in the ocean.
3. Ruth Bader Ginsburg served on the U.S. Supreme Court from 1993 to 2020.

Answer

1. p : Barry Bonds holds the Major League Baseball record for total career home runs. The statement is labeled with a p. Once the statement is labeled, use p as a replacement for the full written statement to write and analyze the argument symbolically.
2. q : Some mammals live in the ocean. The letter q is used to distinguish this statement from the statement in question 1, but any lower-case letter may be used.
3. r : Ruth Bader Ginsburg served on the U.S. Supreme Court from 1993 to 2020. When multiple statements are present in later sections, you will want to be sure to use a different letter for each separate logical statement.

✔ Your Turn 1.1.2.2

Write each of the following logical statements in symbolic form.

1. The movie *Gandhi* won the Academy Award for Best Picture in 1982.
2. Soccer is the most popular sport in the world.
3. All oranges are citrus fruits.

Who Knew?

Mathematics is not the only language to use symbols to represent thoughts or ideas. The Chinese and Japanese languages use symbols known as Hanzi and Kanji, respectively, to represent words and phrases. At one point, the American musician Prince famously changed his name to a symbol representing love.

BBC Prince Symbol Article

Negating Statements

Consider the false statement introduced earlier, “All roses are red.” If someone said to you, “All roses are red,” you might respond with, “Some roses are not red.” You could then strengthen your argument by providing additional statements, such as, “There are also white roses, yellow roses, and pink roses, to name a few.”

The negation of a logical statement has the opposite truth value of the original statement. If the original statement is false, its negation is true, and if the original statement is true, its negation is false. Most logical statements can be negated by simply adding or removing the word *not*. For example, consider the statement, “Emma Stone has green eyes.” The negation of this statement would be, “Emma Stone does not have green eyes.” The table below gives some other examples.

Table 1.1.2.1

Logical Statement	Negation
Gordon Ramsey is a chef.	Gordon Ramsey is not a chef.

Logical Statement	Negation
Tony the Tiger does not have spots.	Tony the Tiger has spots.

The way you phrase your argument can impact its success. If someone presents you with a false statement, the ability to rebut that statement with its negation will provide you with the tools necessary to emphasize the correctness of your position.

? Exercise 1.1.2.3: Negating Logical Statements

Write the negation of each logical statement in words.

1. Michael Phelps was an Olympic swimmer.
2. Tom is a cat.
3. Jerry is not a mouse.

Answer

1. Add the word *not* to negate the affirmative statement: “Michael Phelps was not an Olympic swimmer.”
2. Add the word *not* to negate the affirmative statement: “Tom is not a cat.”
3. Remove the word *not* to negate the negative statement: “Jerry is a mouse.”

✓ Your Turn 1.1.2.3

Write the negation of each logical statement in words.

1. Ted Cruz was not born in Texas
2. Adele has a beautiful singing voice.
3. Leaves convert carbon dioxide to oxygen during the process of photosynthesis.

The symbol for negation, or not, in logic is the tilde, ~. So, not p is represented as $\sim p$. To negate a statement symbolically, remove or add a tilde. The negation of not (not p) is p. Symbolically, this equation is $\sim(\sim p) = p$.

? Exercise 1.1.2.4: Negating Logical Statements Symbolically

Write the negation of each logical statement symbolically.

1. p: Michael Phelps was an Olympic swimmer.
2. r: Tom is not a cat.
3. $\sim q$: Jerry is not a mouse.

Answer

1. To negate an affirmative logical statement symbolically, add a tilde: $\sim p$.
2. Because the symbol for this statement is r, its negation is $\sim r$.
3. The symbol for this statement is $\sim q$, so to negate it we simply remove the ~, because $\sim(\sim q) = q$. The answer is q.

✓ Your Turn 1.1.2.4

Write the negation of each logical statement symbolically.

1. $\sim p$: Ted Cruz was not born in Texas
2. q: Adele has a beautiful voice.
3. r: Leaves convert carbon dioxide to oxygen during the process of photosynthesis.

Translating Negations Between Words and Symbols

In order to analyze logical arguments, it is important to be able to translate between the symbolic and written forms of logical statements.

? Exercise 1.1.2.5

Given the statements:

r: Elmo is a red Muppet.

p: Ketchup is not a vegetable.

1. Write the symbolic form of the statement, “Elmo is not a red Muppet.”
2. Translate the statement $\sim p$ into words.

Answer

1. “Elmo is not a red muppet” is the negation of “Elmo is a red muppet,” which is represented as r. Thus, we would write “Elmo is not a red muppet” symbolically as $\sim r$.
2. Because $p$is the symbol representing the statement, “Ketchup is not a vegetable,” $\sim p$ is equivalent to the statement, “Ketchup is a vegetable.”

✔ Your Turn 1.1.2.5

Given the statements:
r: Woody and Buzz Lightyear are best friends.
$\sim p$: Wonder Woman is not stronger than Captain Marvel.

1. Write the symbolic form of the statement, "Wonder Woman is stronger than Captain Marvel."
2. Translate the statement $\sim r$ into words

Expressing Statements with Quantifiers of All, Some, or None

A quantifier is a term that expresses a numerical relationship between two sets or categories. For example, all squares are also rectangles, but only some rectangles are squares, and no squares are circles. In this example, *all*, *some*, and *none* are quantifiers. In a logical argument, the logical statements made to support the argument are called premises, and the judgment made based on the premises is called the conclusion. Logical arguments that begin with specific premises and attempt to draw more general conclusions are called inductive arguments.

Consider, for example, a parent walking with their three-year-old child. The child sees a cardinal fly by and points it out. As they continue on their walk, the child notices a robin land on top of a tree and a duck flying across to land on a pond. The child recognizes that cardinals, robins, and ducks are all birds, then excitedly declares, "All birds fly!" The child has just made an inductive argument. They noticed that three different specific types of birds all fly, then synthesized this information to draw the more general conclusion that all birds can fly. In this case, the child's conclusion is false.

The specific premises of the child's argument can be paraphrased by the following statements:

- Premise: Cardinals are birds that fly.
- Premise: Robins are birds that fly.
- Premise: Ducks are birds that fly.

The general conclusion is: “All birds fly!”

All inductive arguments should include at least three specific premises to establish a pattern that supports the general conclusion. To counter the conclusion of an inductive argument, it is necessary to provide a counter example. The parent can tell the child about penguins or emus to explain why that conclusion is false.

On the other hand, it is usually impossible to prove that an inductive argument is true. So, inductive arguments are considered either strong or weak. Deciding whether an inductive argument is strong or weak is highly subjective and often determined by the background knowledge of the person making the judgment. Most hypotheses put forth by scientists using what is called the “scientific method” to conduct experiments are based on inductive reasoning.

In the following example, we will use quantifiers to express the conclusion of a few inductive arguments.

Exercise 1.1.2.6: Drawing General Conclusions to Inductive Arguments Using Quantifiers

For each series of premises, draw a logical conclusion to the argument that includes one of the following quantifiers: all, some, or none.

1. Squares and rectangles have four sides. A square is a parallelogram, and a rectangle is a parallelogram. What conclusion can be drawn from these premises?

2. $1+2=3, 6+7=13$, and $23+24=47$. Of these, 1 and 2,6 and 7 , and 23 and 24 are consecutive integers; 3,13 , and 47 are odd numbers. What conclusion can be drawn from these premises?

3. Sea urchins live in the ocean, and they do not breathe air. Sharks live in the ocean, and they do not breathe air. Eels live in the ocean, and they do not breath air. What conclusion can be drawn from these premises?

Answer

1. The conclusion you would likely come to here is “Some four-sided figures are parallelograms.” However, it would be incorrect to say that all four-sided figures are parallelograms because there are some four-sided figures, such as trapezoids, that are not parallelograms. This is a false conclusion.
2. From these premises, you may draw the conclusion “All sums of two consecutive counting numbers result in an odd number.” Most inductive arguments cannot be proven true, but several mathematical properties can be. If we let n n represent our first counting number, then $n + 1$ would be the next counting number and $n + (n + 1) = 2 n + 1$. Because $2n$ is divisible by two, it is an even number, and if you add one to any even number the result is always an odd number. Thus, the conclusion is true!
3. Based on the premises provided, with no additional knowledge about whales or dolphins, you might conclude “No creatures that live in the ocean breathe air.” Even though this conclusion is false, it still follows from the known premises and thus is a logical conclusion based on the evidence presented. Alternatively, you could conclude “Some creatures that live in the ocean do not breathe air.” The quantifier you choose to write your conclusion with may vary from another person’s based on how persuasive the argument is. There may be multiple acceptable answers.

Your Turn 1.1.2.6

For each series of premises, draw a logical conclusion to the argument that includes one of the following quantifiers: all, some, or none.

1. $1+2=3$, $5+6=11$, and $14+15=29$. Of these, 1 and 2 are consecutive integers, 5 and 6 are consecutive integers, and 14 and 15 are consecutive integers. Also, their sums, 3, 11, and 29 are all prime numbers. Prime numbers are positive integers greater than one that are only divisible by one and the number itself. What conclusion can you draw from these premises?
2. A robin is a bird that lays blue eggs. A chicken is a bird that typically lays white and brown eggs. An ostrich is a bird that lays exceptionally large eggs. If a bird lays eggs, then they do not give live birth to their young. What conclusion can you draw from these premises?
3. All parallelograms have four sides. All rectangles are parallelograms. All squares are rectangles. What additional conclusion can you make about squares from these premises?

Checkpoint

It is not possible to prove definitively that an inductive argument is true or false in most cases.

Negating Statements Containing Quantifiers

Recall that the negation of a statement will have the opposite truth value of the original statement. There are four basic forms that logical statements with quantifiers take on.

- All **A** are B.
- Some A are B.
- No A are B.

- Some A are not B.

The negation of logical statements that use the quantifiers *all*, *some*, or *none* is a little more complicated than just adding or removing the word *not*.

For example, consider the logical statement, “All oranges are citrus fruits.” This statement expresses as a subset relationship. The set of oranges is a subset of the set of citrus fruit. This means that there are no oranges that are outside the set of citrus fruit. The negation of this statement would have to break the subset relationship. To do this, you could say, “At least one orange is not a citrus fruit.” Or, more concisely, “Some oranges are not citrus fruit.” It is tempting to say "No oranges are citrus fruit," but that would be incorrect. Such a statement would go beyond breaking the subset relationship, to stating that the two sets have nothing in common. The negation of "A is a subset of B" would be to state that "A is not a subset of B," as depicted by the Venn diagram in Figure 1.1.2.3

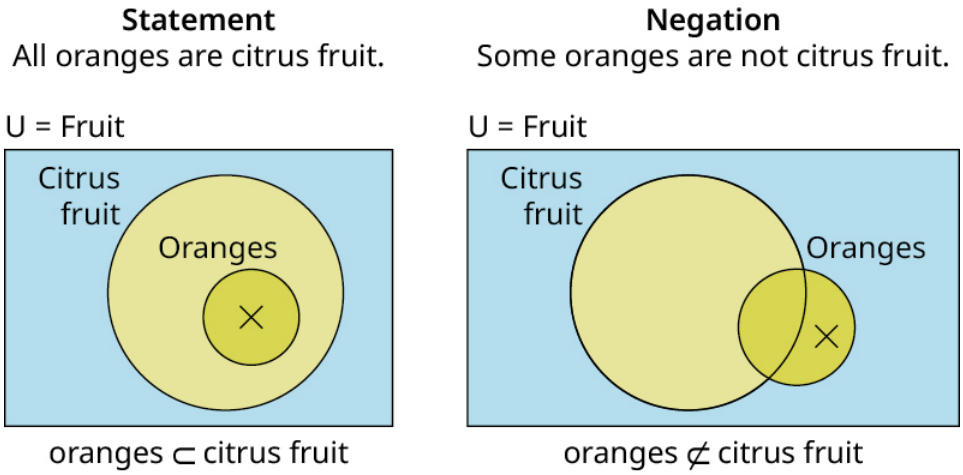


Figure 1.1.2.3: The statement, “All oranges are citrus fruit,” is true, so its negation, “Some oranges are not citrus fruit,” is false.

Now, consider the statement, “No apples are oranges.” This statement indicates that the set of apples is disjointed from the set of oranges. The negation must state that the two are not disjoint sets, or that the two sets have a least one member in common. Their intersection is not empty. TThe negation of the statement, “A intersection B is the empty set,” is the statement that " A intersection B is not empty," as depicted in the Venn diagram in Figure 1.1.2.4

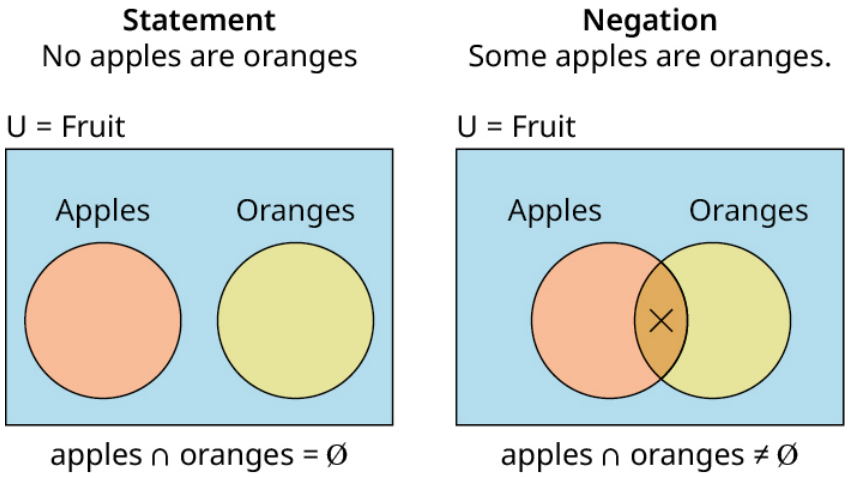


Figure 1.1.2.4:The negation of the true statement “No apples are oranges,” is the false statement, “Some apples are oranges.”

Table 1.1.2.2 summarizes the four different forms of logical statements involving quantifiers and the forms of their associated negations, as well as the meanings of the relationships between the two categories or sets A and B.

Table 1.1.2.2

Logical Statements with Quantifiers	Negation of Logical Statements w/Quantifiers

Logical Statements with Quantifiers	Negation of Logical Statements w/Quantifiers
Form: All A are B. Means: A is a subset of B, $A \subset B$. All zebras have stripes. (True)	Form: Some A are not B. Means: A is not a subset of B, $A \not\subset B$. Some zebras do not have stripes. (False)
Form: Some A are B. Means: A intersection B is not empty, $A \cap B \neq \emptyset$. Some fish are sharks. (True)	Form: No A are B. Means: A intersection B is empty, $A \cap B = \emptyset$. No fish are sharks. (False)
Form: No A are B. Means: A intersection B is empty, $A \cap B = \emptyset$. No trees are evergreens. (False)	Form: Some A are B. Means: A intersection B is not empty, $A \cap B \neq \emptyset$. Some trees are evergreens. (True)
Form: Some A are not B. Means: A is not a subset of B, $A \not\subset B$. Some horses are not mustangs. (True)	Form: All A are B. Means: A is a subset of B, $A \subset B$. All horses are mustangs. (False)

We covered sets in great detail in Chapter 1. To review, " A is a subset of B" means that every member of set A is also a member of set B. The intersection of two sets A and B is the set of all elements that they share in common. If A intersection B is the empty set, then sets A and B do not have any elements in common. The two sets do not overlap. They are disjoint.



? Exercise 1.1.2.7: Negating Statements Containing Quantifiers *All*, *Some*, or *None*

Given the statements:

p: All leopards have spots.
r: Some apples are red.
s: No lemons are sweet.

Write each of the following symbolic statements in words.

1. $\sim p$
2. $\sim r$
3. $\sim s$

Answer

1. The statement "All leopards have spots" is p and has the form "All A are B." According to Table 1.1.2.2 the negation will have the form "Some A are not B." The negation of p is the statement, "Some leopards do not have spots."
2. The statement "Some apples are red" has the form "Some A are B." This indicates that the categories A and B overlap or intersect. According to Table 1.1.2.2 the negation will have the form, "No A are B," indicating that A and B do not intersect. This results in the opposite truth value of the original statement, so the negation of "Some apples are red" is the statement: "No apples are red."

3. Because s is the statement: "No lemons are sweet," s is asserting that the set of lemons does not intersect with the set of sweet things. The negation of s, $\sim s$, must make the opposite claim. It must indicate that the set of lemons intersects with the set of sweet things. This means at least one lemon must be sweet. The statement, "Some lemons are sweet" is $\sim s$. The negation of the statement, "No A are B," is the statement, "Some A are B," as indicated in Table 1.1.2.2

Your Turn 1.1.2.7

Given the statements:
$\sim p$: Some apples are not sweet.
r : No triangles are squares.
s : Some vegetables are green. Write each of the following symbolic statements in words.

1. p
2. $\sim r$
3. $\sim s$

Check Your Understanding

1. A _____ _____ is a complete sentence that makes a claim that may be either true or false.
2. The _____ of a logical statement has the opposite truth value of the original statement.
3. If p represents the logical statement, “Marigolds are yellow flowers,” then _____ represents the statement, “Marigolds are not yellow flowers.”
4. The statement $\sim(\sim p)$ has the same truth value as the statement _____.
5. The logical statements used to support the conclusion of an argument are called _____.
6. _____ arguments attempt to draw a general conclusion from specific premises.
7. All, some, and none are examples of _____, words that assign a numerical relationship between two or more groups.
8. The negation of the statement, “All giraffes are tall,” is _____.

This page titled 1.1.2: Statements and Quantifiers is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **2.1: Statements and Quantifiers** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

1.1.3: Compound Statements



1.1.3.1:A person seeking their driver's license must pass two tests. A compound statement can be used to explain performance on both tests at once. (credit: modification of work “Drivers License -Teen driver” by State Farm/Flickr, CC BY 2.0)

Learning Objectives

After completing this section, you should be able to:

1. Translate compound statements into symbolic form.
2. Translate compound statements in symbolic form with parentheses into words.
3. Apply the dominance of connectives.

Suppose your friend is trying to get a license to drive. In most places, they will need to pass some form of written test proving their knowledge of the laws and rules for driving safely. After passing the written test, your friend must also pass a road test to demonstrate that they can perform the physical task of driving safely within the rules of the law.

Consider the statement: "My friend passed the written test, but they did not pass the road test." This is an example of a compound statement, a statement formed by using a connective to join two independent clauses or logical statements. The statement, “My friend passed the written test,” is an independent clause because it is a complete thought or idea that can stand on its own. The second independent clause in this compound statement is, “My friend did not pass the road test.” The word "but" is the connective used to join these two statements together, forming a compound statement. So, did your friend acquire their driving license?.

This section introduces common logical connectives and their symbols, and allows you to practice translating compound statements between words and symbols. It also explores the order of operations, or dominance of connectives, when a single compound statement uses multiple connectives.

Common Logical Connectives

Understanding the following logical connectives, along with their properties, symbols, and names, will be key to applying the topics presented in this chapter. The chapter will discuss each connective introduced here in more detail.

The joining of two logical statements with the word "and" or "but" forms a compound statement called a conjunction. In logic, for a conjunction to be true, all the independent logical statements that make it up must be true. The symbol for a conjunction is $\wedge$. Consider the compound statement, “Derek Jeter played professional baseball for the New York Yankees, and he was a shortstop.” If p represents the statement, “Derrick Jeter played professional baseball for the New York Yankees,” and if q represents the statement, “Derrick Jeter was a short stop,” then the conjunction will be written symbolically as $p \wedge q$.

The joining of two logical statements with the word “or” forms a compound statement called a disjunction. Unless otherwise specified, a disjunction is an inclusive *or* statement, which means the compound statement formed by joining two independent clauses with the word *or* will be true if a least one of the clauses is true. Consider the compound statement, "The office manager ordered cake for for an employee’s birthday or they ordered ice cream.” This is a disjunction because it combines the independent clause, “The office manager ordered cake for an employee’s birthday,” with the independent clause, “The office manager ordered ice cream,” using the connective, *or*. This disjunction is true if the office manager ordered only cake, only ice cream, or they ordered both cake and ice cream. Inclusive *or* means you can have one, or the other, or both!

Joining two logical statements with the word *implies*, or using the phrase "if *first statement*, then *second statement*," is called a conditional or *implication*. The clause associated with the "if" statement is also called the hypothesis or *antecedent*, while the clause following the "then" statement or the word *implies* is called the **conclusion** or *consequent*. The conditional statement is like a one-way contract or promise. The only time the conditional statement is false, is if the hypothesis is true and the conclusion is false. Consider the following conditional statement, "If Pedro does his homework, then he can play video games." The hypothesis/antecedent is the statement following the word *if*, which is "Pedro does/did his homework." The conclusion/consequent is the statement following the word *then*, which is "Pedro can play his video games."

Joining two logical statements with the connective phrase "if and only if" is called a biconditional. The connective phrase "if and only if" is represented by the symbol, $\leftrightarrow$. In the biconditional statement, $p \leftrightarrow q$, p is called the hypothesis or antecedent and q is called the conclusion or consequent. For a biconditional statement to be true, the truth values of p and q must match. They must both be true, or both be false.

The table below lists the most common connectives used in logic, along with their symbolic forms, and the common names used to describe each connective.

Table 1.1.3.1

Connective	Symbol	Name
and but	$\wedge$	conjunction
or	$\vee$	disjunction, inclusive or
not	~	negation
if . . ., then implies	$\rightarrow$	conditional, implication
if and only if	$\leftrightarrow$	biconditional

Checkpoint

These connectives, along with their names, symbols, and properties, will be used throughout this chapter and should be memorized.

Exercise 1.1.3.1

Associate Connectives with Symbols and Names

For each of the following connectives, write its name and associated symbol.

1. or
2. implies
3. but

Answer

A compound statement formed with the connective word *or* is called a disjunction, and it is represented by the $\vee$ symbol.

A compound statement formed with the connective word *implies* or phrase "*if …, then*" is called a conditional statement or implication and is represented by the $\rightarrow$ symbol.

A compound statement formed with the connective words *but* or *and* is called a conjunction, and it is represented by the $\wedge$ symbol.

Your Turn 1.1.3.1

For each connective write its name and associated symbol.

1. not

2. and
3. if and only if

Translating Compound Statements to Symbolic Form

To translate a compound statement into symbolic form, we take the following steps.

1. Identify and label all independent affirmative logical statements with a lower case letter, such as p, q, or r.
2. Identify and label any negative logical statements with a lowercase letter preceded by the negation symbol, such as $\sim p$, $\sim q$, or $\sim r$.
3. Replace the connective words with the symbols that represent them, such as $\wedge$, $\vee$, $\rightarrow$, or $\leftrightarrow$.

Consider the previous example of your friend trying to get their driver's license. Your friend passed the written test, but they did not pass the road test. Let p represent the statement, "My friend passed the written test." And, let $\sim q$ represent the statement, "My friend did not pass the road test." Because the connective *but* is logically equivalent to the word *and*, the symbol for *but* is the same as the symbol for *and*; replace *but* with the symbol $\wedge$. The compound statement is symbolically written as: $p \wedge \sim q$. My friend passed the written test, but they did not pass the road test.

Checkpoint

When translating English statements into symbolic form, do not assume that every connective "and" means you are dealing with a compound statement. The statement, "The stripes on the U.S. flag are red and white," is a simple statement. The word "white" doesn't express a complete thought, so it is not an independent clause and does not get its own symbol. This statement should be represented by a single letter, such as s: The stripes on the U.S. flag are red and white.

? Exercise 1.1.3.2: Translating Compound Statements into Symbolic Form

Let p represent the statement, "It is a warm sunny day," and let q represent the statement, "the family will go to the beach." Write the symbolic form of each of the following compound statements.

1. If it is a warm sunny day, then the family will go to the beach.
2. The family will go to the beach, and it is a warm sunny day.
3. The family will not go to the beach if and only if it is not a warm sunny day.
4. The family not go to the beach, or it is a warm sunny day.

Answer

1. Replace "it is a warm sunny day" with p. Replace "the family will go to the beach." with q. Next. Next, because the connective is if ..., then place the conditional symbol, $\rightarrow$, between p and q. The compound statement is written symbolically as: $p \rightarrow q$.
2. Replace "The family will go to the beach" with q. Replace "it is a warm sunny day." with p. Next, because the connective is and, place the $\wedge$ symbol between q and p. The compound statement is written symbolically as: $q \wedge p$.
3. Replace "The family will not go to the beach. with $\sim q$. Replace "it is not a warm sunny day" with $\sim p$. Next, because the connective is or, if and only if, place the biconditional symbol, $\leftrightarrow$ between $\sim q$ and $\sim p$. The compound statement is written symbolically as: $\sim q \leftrightarrow \sim p$.
4. Replace "The family will not go to the beach" with $\sim q$. Replace "it is a warm sunny day" with p. Next, because the connective is or, place the $\vee$ symbol between $\sim q$ and p. The compound statement is written symbolically as: $\sim q \vee p$.

✓ Your Turn 1.1.3.2

Let p represent the statement, "Last night it snowed," and let q represent the statement, "Today we will go skiing." Write the symbolic form of each of the following compound statements:

1. Today we will go skiing, but last night it did not snow.
2. Today we will go skiing if and only if it snowed last night.
3. Last night is snowed or today we will not go skiing.
4. If it snowed last night, then today we will go skiing.

Translating Compound Statements in Symbolic Form with Parentheses into Words

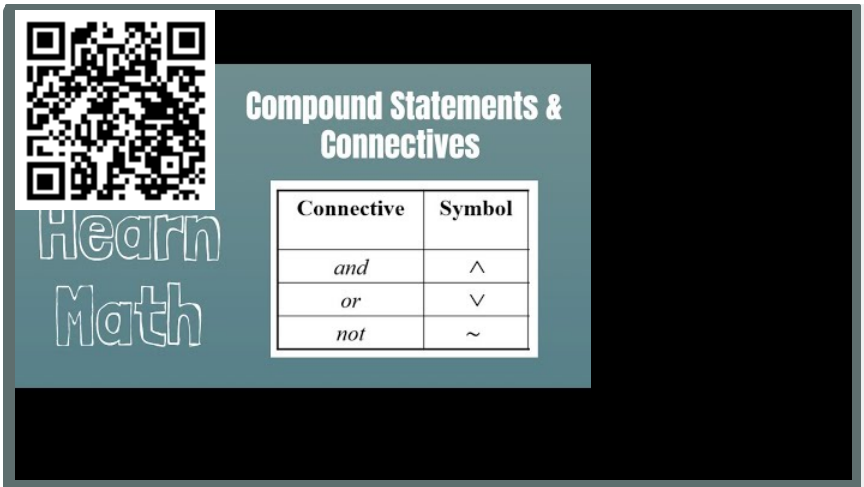
When parentheses are written in a logical argument, they group a compound statement together just like when calculating numerical or algebraic expressions. Any statement in parentheses should be treated as a single component of the expression. If multiple parentheses are present, work with the inner most parentheses first.

Consider your friend's struggles to get their license to drive. Let p represent the statement, "My friend passed the written test," let q represent the statement, "My friend passed the road test," and let r represent the statement, "My friend received a driver's license." The statement $(p \wedge q) \rightarrow r$ can be translated into words as follows: the statement $p \wedge q$ is grouped together to form the hypothesis of the conditional statement and r is the conclusion. The conditional statement has the form "if $p \wedge q$, then r." Therefore, the written form of this statement is: "If my friend passed the written test and they passed the road test, then my friend received a driver's license."

Sometimes a compound statement within parentheses may need to be negated as a group. To accomplish this, add the phrase, "it is not the case that" before the translation of the phrase in parentheses. For example, using p, q, and r of your friend obtaining a license, let's translate the statement $\sim (p \wedge q) \rightarrow \sim r$ into words.
In this case, the hypothesis of the conditional statement is $\sim (p \wedge q)$ and the conclusion is $\sim r$. To negate the hypothesis, add the phrase "it is not the case" before translating what is in parentheses. The translation of the hypothesis is the sentence, "It is not the case that my friend passed the written test and they passed the road test," and the translation of the conclusion is, "My friend did not receive a driver's license." So, a translation of the complete conditional statement, $\sim (p \wedge q) \rightarrow \sim r$ is: "If it is not the case that my friend passed the written test and the road test, then my friend did not receive a driver's license."

Checkpoint

It is acceptable to interchange proper names with pronouns and remove repeated phrases to make the written statement more readable, as long the meaning of the logical statement is not changed.



Logic Part 1B: Compound Statements, Connectives and Symbols

Exercise 1.1.3.3

Let p represent the statement, "My child finished their homework," let q represent the statement, "My child cleaned her room," let r represent the statement, "My child played video games," and let s represent the statement, "My child streamed a movie." Translate each of the following symbolic statements into words.

1. $\sim (p \wedge q)$
2. $(p \wedge q) \rightarrow (r \vee s)$
3. $\sim (r \vee s) \leftrightarrow \sim (p \wedge q)$

Answer

1. Replace with "It is not the case," and $\wedge$ with "and." One possible translation is: "It is not the case that my child finished their homework and cleaned their room."
2. The hypothesis of the conditional statement is, "My child finished their homework and cleaned their room." The conclusion of the conditional statement is, "My child played video games or streamed a movie." One possible translation of the entire statement is: "If my child finished their homework and cleaned their room, then they played video games or streamed a movie."
3. The hypothesis of the biconditional statement is $\sim (r \vee s)$ and is written in words as: "It is not the case that my child played video games or streamed a movie." The conclusion of the biconditional statement is $\sim (p \wedge q)$, which translates to: "It is not the case that my child finished their homework and cleaned their room." Because the biconditional, $\leftrightarrow$ translates to if and only if, one possible translation of the statement is: "It is not the case that my child played video games or streamed a movie if and only if it is not the case that my child finished their homework and cleaned their room."

Your Turn 1.1.3.3

Let p represent the statement, "My roommates ordered pizza," let q represent the statement, "I ordered wings," and let r be the statement, "Our friends came over to watch the game." Translate the following statements into words.

1. $\sim r \rightarrow (p \vee q)$
2. $(p \wedge q) \rightarrow r$
3. $\sim (p \vee r)$

The Dominance of Connectives

The order of operations for working with algebraic and arithmetic expressions provides a set of rules that allow consistent results. For example, if you were presented with the problem $1 + 3 \times 2$, and you were not familiar with the order of operation, you might assume that you calculate the problem from left to right. If you did so, you would add 1 and 3 to get 4, and then multiply this answer by 2 to get 8, resulting in an incorrect answer. Try inputting this expression into a scientific calculator. If you do, the calculator should return a value of 7, not 8.

Use the Scientific Calculator found under the Resources tab to the far left

The order of operations for algebraic and arithmetic operations states that all multiplication must be applied prior to any addition. Parentheses are used to indicate which operation—addition or multiplication—should be done first. Adding parentheses can change and/or clarify the order. The parentheses in the expression $1 + (3 \times 2)$ indicate that 3 should be multiplied by 2 to get 6, and then 1 should be added to 6 to get 7: $1 + (3 \times 2) = 7$.

As with algebraic expressions, there is a set of rules that must be applied to compound logical statements in order to evaluate them with consistent results. This set of rules is called the dominance of connectives. When evaluating compound logical statements, connectives are evaluated from least dominant to most dominant as follows:

1. Parentheses are the least dominant connective. So, any expression inside parentheses must be evaluated first. Add as many parentheses as needed to any statement to specify the order to evaluate each connective.
2. Next, we evaluate negations.
3. Then, we evaluate conjunctions and disjunctions from left to right, because they have equal dominance.
4. After evaluating all conjunctions and disjunctions, we evaluate conditionals.
5. Lastly, we evaluate the most dominant connective, the biconditional. If a statement includes multiple connectives of equal dominance, then we will evaluate them from left to right.

See Figure $1.1.3.2$for a visual breakdown of the dominance of connectives.

Dominance	Connective	Symbol	Evaluate
Least Dominant	Parentheses	( )	First
↓	Negation	~	↓
	Disjunction/Conjunction	∨, ∧	Left to right or add parentheses to specify order because or/and have equal dominance.
	Conditional	→	↓
Most Dominant	Biconditional	↔	Last

Figure 1.1.3.2

Let's revisit your friend's struggles to get their driver's license. Let p represent the statement, "My friend passed the written test," let q represent the statement, "My friend passed the road test," and let r represent the statement, "My friend received a driver's license." Let's use the dominance of connectives to determine how the compound statement $p \wedge \sim q \rightarrow r$ should be evaluated.

Step 1: There are no parentheses, which is least dominant of all connectives, so we can skip over that.

Step 2: Because negation is the next least dominant, we should evaluate $\sim q$ first. We could add parentheses to the statement to make it clearer which connecting needs to be evaluated first: $p \wedge \sim q \rightarrow r$ is equivalent to $p \wedge (\sim q) \rightarrow r$.

Step 3: Next, we evaluate the conjunction, $\wedge$. $p \wedge (\sim q) \rightarrow r$ is equivalent to $(p \wedge (\sim q)) \rightarrow r$.

Step 4: Finally, we evaluate the conditional, $\rightarrow$, as this is the most dominant connective in this compound statement.

Who Knew?

When using spreadsheet applications, like Microsoft Excel or Google Sheets, have you noticed that core functions, such as sum, average, or rate, have the same name and syntax for use? This is not a coincidence. Various standards organizations have developed requirements that software developers must implement to meet the conditions set by governments and large customers worldwide. The OpenDocument Format OASIS Standard enables transferring data between different office productivity applications and was approved by the International Standards Organization (ISO) and IEC on May 6, 2006.

Prior to adopting these standards, a government entity, business, or individual could lose access to their own data simply because it was saved in a format no longer supported by a proprietary software product—even data they were required by law to preserve, or data they simply wished to maintain access to.

Just as rules for applying the dominance of connectives help maintain understanding and consistency when writing and interpreting compound logical statements and arguments, the standards adopted for database software maintain global integrity and accessibility across platforms and user bases.

Exercise 1.1.3.4

For each of the following compound logical statements, add parentheses to indicate the order to evaluate the statement. Recall that parentheses are evaluated innermost first.

1. $p \wedge \sim q \vee r$
2. $q \rightarrow \sim p \wedge r$
3. $\sim (p \vee q) \leftrightarrow \sim p \wedge \sim q$

Answer

1. Because negation is the least dominant connective, we evaluate it first: $p \wedge (\sim q) \vee r$. Because conjunction and disjunction have the same dominance, we evaluate them left to right. So, we evaluate the conjunction next, as indicated by the additional set of parentheses: $(p \wedge (\sim q)) \vee r$. The only remaining connective is the disjunction, so it is

evaluated last, as indicated by the third set of parentheses. The complete solution is: $((p \wedge (\sim q)) \vee r)$.

2. Negation has the lowest dominance, so it is evaluated first: $q \rightarrow (\sim p) \wedge r$. The remaining connectives are the conditional and the conjunction. Because conjunction has a lower precedence than the conditional, it is evaluated next, as indicated by the second set of parentheses: $q \rightarrow ((\sim p) \wedge r)$. The last step is to evaluate the conditional, as indicated by the third set of parentheses: $(q \rightarrow ((\sim p) \wedge r))$.
3. This statement is known as De Morgan's Law for the negation of a disjunction. It is always true. Section 2.6 of this chapter will explore De Morgan's Laws in more detail.
   - First, we evaluate the negations on the right side of the biconditional prior to the conjunction.
   - Then, we evaluate the disjunction on the left side of the biconditional, followed by the negation of the disjunction on the left side.
   - Lastly, after completely evaluating each side of the biconditional, we evaluate the biconditional. It does not matter which side you begin with.
   The final solution is: $(\sim (p \vee q)) \leftrightarrow ((\sim p) \wedge (\sim q))$.

Example 1.1.3.4

For each of the following compound logical statements, add parentheses to indicate the order in which to evaluate the statement. Recall that parentheses are evaluated innermost first.

1. $p \vee q \wedge \sim r$
2. $\sim p \rightarrow q \vee r$
3. $\sim p \vee \sim q \leftrightarrow \sim (p \wedge q)$

WORK IT OUT: Logic Terms and Properties – Matching Game

Materials: For every group of four students, include 30 cards (game, trading, or playing cards), 30 individual clear plastic gaming card sleeves, and 30 card-size pieces of paper. Prepare a list of 60 questions and answers ahead of time related to definitions and problems in Statements and Quantifiers and Compound Statements. Provide each group of four students with 20 questions and their associated answers. Instruct each group to select 15 of the 20 questions, and then, for each problem selected, create one card with the question and one card with the answer. Instruct the groups to then place each problem and answer in a separate card sleeve. Once they create 15 problem cards and 15 answer cards, have students shuffle the set of cards.

To play the game, groups should exchange their card decks to ensure no team begins playing with the deck that they created. Split each four-person group into teams of two students. After shuffling the cards, one team lays cards face down on their desk in a five-by-six grid. The other team will go first.

Each player will have a turn to try matching the question with the correct answer by flipping two cards to the face up position. If a team makes a match, they get to flip another two cards; if they do not get a match, they flip the cards face down and it is the other team's turn to flip over two cards. The game continues in this manner until teams match all question cards with their corresponding answer cards. The team with the most set of matching cards wins.

In the first module of this chapter, we evaluated the truth value of negations. In the following modules, we will discuss how to evaluate conjunctions, disjunctions, conditionals, and biconditionals, and then evaluate compound logical statements using truth tables and our knowledge of the dominance of connectives.

Check Your Understanding

1. A _____ _____ is a logical statement formed by combining two or more statements with connecting words, such as and, or, but, not, and if …, then.
2. A _____ is a word or symbol used to join two or more logical statements together to form a compound statement.
3. The most dominant connective is the _____.
4. _____ are used to specify which logical connective should be evaluated first when evaluating a compound statement.
5. Both _____ and _____ have equal dominance and are evaluated from left to right when no parentheses are present in a compound logical statement.

This page titled 1.1.3: Compound Statements is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **2.2: Compound Statements** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

1.2: Truth Tables

1.2: Truth Tables is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

1.2.1: Constructing Truth Tables



Figure 1.2.1.1: Just like solving a puzzle, a computer programmer must consider all potential solutions in order to account for each possible outcome. (credit: modification of work “Deadline Xmas 2010” by Allan Henderson/Flickr, CC BY 2.0)

Learning Objectives

After completing this section, you should be able to:

1. Interpret and apply negations, conjunctions, and disjunctions.
2. Construct a truth table using negations, conjunctions, and disjunctions.
3. Construct a truth table for a compound statement and interpret its validity.

Are you familiar with the Choose Your Own Adventure book series written by Edward Packard? These gamebooks allow the reader to become one of the characters and make decisions that affect what happens next, resulting in different sequences of events in the story and endings based on the choices made. Writing a computer program is a little like what it must be like to write one of these books. The programmer must consider all the possible inputs that a user can put into the program and decide what will happen in each case, then write their program to account for each of these possible outcomes.

A truth table is a graphical tool used to analyze all the possible truth values of the component logical statements to determine the validity of a statement or argument along with all its possible outcomes. The rows of the table correspond to each possible outcome for the given logical statement identified at the top of each column. A single logical statement p has two possible truth values, true or false. In truth tables, a capital T will represent true values, and a capital F will represent false values.

In this section, you will use the knowledge built in Statements and Quantifiers and Compound Statements to analyze arguments and determine their truth value and validity. A logical argument is valid if its conclusion follows from its premises, regardless of whether those premises are true or false. You will then explore the truth tables for negation, conjunction, and disjunction, and use these truth tables to analyze compound logical statements containing these connectives.

Interpret and Apply Negations, Conjunctions, and Disjunctions

The negation of a statement will have the opposite truth value of the original statement. When p is true, $\sim p$ is false, and when p is false, $\sim p$ is true.

? Exercise 1.2.1.1

For each logical statement, determine the truth value of its negation.

1. p: $3+5=8$.
2. q: All horses are mustangs.
3. $\sim r$: New Delhi is not the capital of India.

Answer

p is true because 3 + 5 does equal 8; therefore, the negation of p, ~ p : 3 + 5 ≠ 8 ~ p : 3 + 5 ≠ 8 , is false.

q is false because there are other types of horses besides mustangs, such as Clydesdales or Arabians; therefore, the negation of q, $\sim q$, is true.

$\sim r$ is false because New Delhi is the capital of India; therefore, the negation of $\sim r$, r , is true.

Your Turn 1.2.1.1

For each logical statement, determine the truth value of its negation.

1. $\sim p$: $3 \times 5 = 14$.
2. $\sim q$: Some houses are built with bricks
3. r: Abuja is the capital of Nigeria.

A conjunction is a logical *and* statement. For a conjunction to be true, both statements that make up the conjunction must be true. If at least one of the statements is false, the *and* statement is false.

1. $p \wedge q$
2. $\sim q \wedge r$
3. $\sim p \wedge q$

Answer

1. p is true, and q is false. Because one statement is true, and the other statement is false, this makes the complete conjunction false.
2. q is false, so $\sim q$ is true, and r is true. Therefore, both statements are true, making the complete conjunction true.
3. $\sim p$ is false, and q is false. Because both statements are false, the complete conjunction is false.

Your Turn 1.2.1.2

Given p : Yellow is a primary color, q : Blue is a primary color, and r : Green is a primary color, determine the truth value of each conjunction.

1. $p \wedge q$
2. $q \wedge r$
3. $\sim r \wedge p$.

Checkpoint

The only time a conjunction is true is if both statements that make up the conjunction are true.

A disjunction is a logical inclusive *or* statement, which means that a disjunction is true if one or both statements that form it are true. The only way a logical inclusive *or* statement is false is if both statements that form the disjunction are false.

1. $p \vee q$
2. $\sim q \vee r$
3. $\sim p \vee q$

Answer

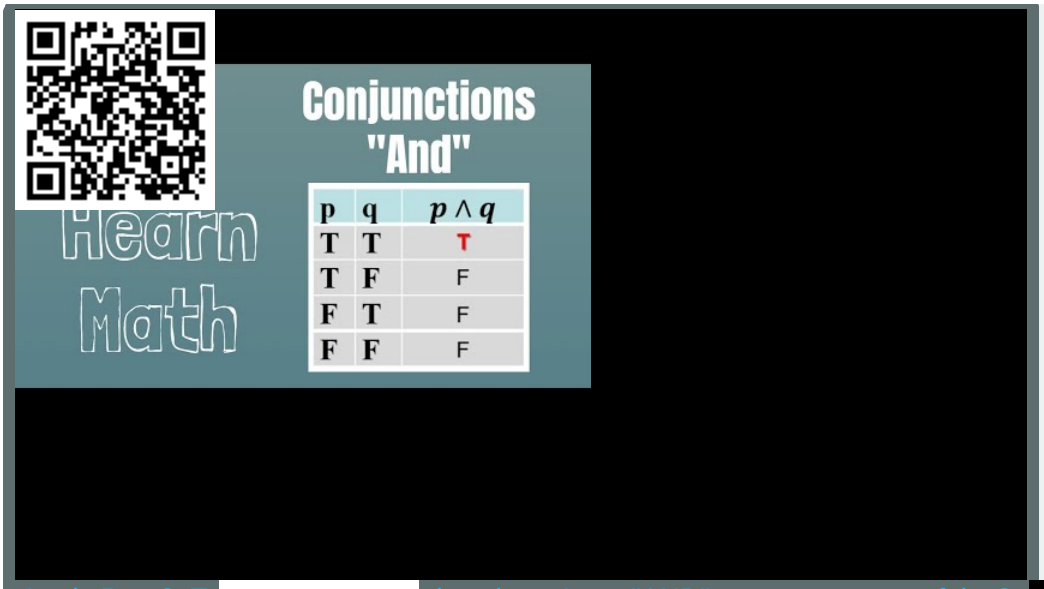
1. **p is true, and q is false. One statement is true, and one statement is false, which makes the complete disjunction true.**
2. **q is false, so $\sim q$ is true, and r is true. Therefore, one statement is true, and the other statement is true, which makes the complete disjunction true.**
3. **$\sim p$ is false, and q is false. When all of the component statements are false, the disjunction is false.**

Your Turn 1.2.1.3

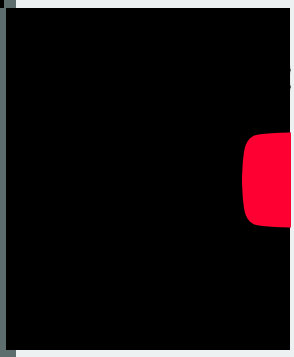
Given p : Yellow is a primary color, q : Blue is a primary color, and r : Green is a primary color, determine the truth value of each disjunction.

1. $p \vee q$
2. $\sim p \vee r$
3. $q \vee r$

In the next example, you will apply the dominance of connectives to find the truth values of compound statements containing negations, conjunctions, and disjunctions and use a table to record the results. When constructing a truth table to analyze an argument where you can determine the truth value of each component statement, the strategy is to create a table with two rows. The first row contains the symbols representing the components that make up the compound statement. The second row contains the truth values of each component below its symbol. Include as many columns as necessary to represent the value of each compound statement. The last column includes the complete compound statement with its truth value below it.



Logic Part 2: Tr njunctions: Is an "AND" statement true or false?



Logic Part 3: Truth Values of Disjunctions: Is an "OR" statement true or false?

1. $\sim p \wedge q \vee r$
2. $\sim p \vee q \wedge r$
3. $\sim (p \wedge r) \vee q$

Answer

1. **Step 1:** The statement " $\sim p \wedge q \vee r$ " contains three basic logical statements, p, q, and r, and three connectives, $\sim, \wedge, \vee$. When we place parentheses in the statement to indicate the dominance of connectives, the statement becomes $(((\sim p) \wedge q) \vee r)$.
Step 2: After we have applied the dominance of connectives, we create a two row table that includes a column for each basic statement that makes up the compound statement, and an additional column for the contents of each parentheses. Because we have three sets of parentheses, we include a column for $\sim p$, the innermost parentheses, a column for $(\sim p) \wedge q$, the next set of parentheses, and $((\sim p) \wedge q) \vee r$ in the last column for the third parentheses.
Step 3: Once the table is created, we determine the truth value of each statement starting from left to right. The truth values of p, q, and r are true, false, and true, respectively, so we place T, F, and T in the second row of the table. Because p is true, $\sim p$ is false.
Step 4: Next, evaluate $\sim p \wedge q$ from the table: $\sim p$ is false, and q is also false, so $\sim p \wedge q$ is false, because a conjunction is only true if both of the statements that make it are true. Place an F in the table below its heading.
Step 5: Finally, using the table, we understand that $(\sim p \wedge q)$ is false and r is true, so the complete statement $(\sim p \wedge q) \vee r$ is false or true, which is true (because a disjunction is true whenever at least one of the statements that make it is true). Place a T in the last column of the table. The complete statement $\sim p \wedge q \vee r$ is true.

2. **Step 1**: Applying the dominance of connectives to the original compound statement $\sim p \vee q \wedge r$, we get $(((\sim p) \vee q) \wedge r)$.
Step 2: The table needs columns for $p, q, r, \sim p, (\sim p) \vee q$, and $((\sim p) \vee q) \wedge r$.
Step 3:The truth values of p, q, r, and $\sim p$ are the same as in Question 1.
Step 4: Next, $\sim p \vee q$ is false or false, which is false, so we place an F below this statement in the table. This is the only time that a disjunction is false.
Step 5: Finally, $\sim p \vee q$ and r are the conjunction of the statements, $\sim p \vee q$ and r, and so the expressions evaluate to false and true, which is false. Recall that the only time an "and" statement is true is when both statements that form it are also true. The complete statement $\sim p \vee q \wedge r$ is false.

3. **Step 1**: Applying the dominance of connectives to the original statement, we have: $((\sim (p \wedge r)) \vee q)$.
Step 2: So, the table needs the following columns: $p, q, r, p \wedge r, \sim (p \wedge r)$, and $\sim (p \wedge r) \vee q$.
Step 3: The truth values of p, q, and r are the same as in Questions 1 and 2.
Step 4: From the table it can be seen that $p \wedge r$ is true and true, which is true. So the negation of p and r is false, because the negation of a statement always has the opposite truth value of the original statement.
Step 5: Finally, $\sim (p \wedge r) \vee q$ is the disjunction of $\sim (p \wedge r)$ with q, and so we have false or false, which makes the complete statement false.

T	F	T	T	F	F

✔ Your Turn ✔ 1.2.1.4

Given p : Yellow is a primary color, q : Blue is a primary color, and r : Green is a primary color, determine the truth value of each compound statement, by constructing a truth table.

1. $\sim q \wedge p \vee r$
2. $p \vee q \wedge \sim r$
3. $\sim (p \wedge r) \wedge q$

Construct Truth Tables to Analyze All Possible Outcomes

Recall from Statements and Questions that the negation of a statement will always have the opposite truth value of the original statement; if a statement p is false, then its negation $\sim p$ is true, and if a statement p is true, then its negation $\sim p$ is false. To create a truth table for the negation of statement p, add a column with a heading of $\sim p$, and for each row, input the truth value with the opposite value of the value listed in the column for p, as depicted in the table below.

Negation	
p	**~ p**
T	F
F	T

Conjunctions and disjunctions are compound statements formed when two logical statements combine with the connectives “and” and “or” respectively. How does that change the number of possible outcomes and thus determine the number of rows in our truth table? The multiplication principle, also known as the *fundamental counting principle*, states that the number of ways you can select an item from a group of n items and another item from a group with m items is equal to the product of m and n. Because each proposition p and q has two possible outcomes, true or false, the multiplication principle states that there will be $2 \times 2 = 4$ possible outcomes: {TT, TF, FT, FF}.

The tree diagram and table in Figure $1.2.1.2$demonstrate the four possible outcomes for two propositions p and q.

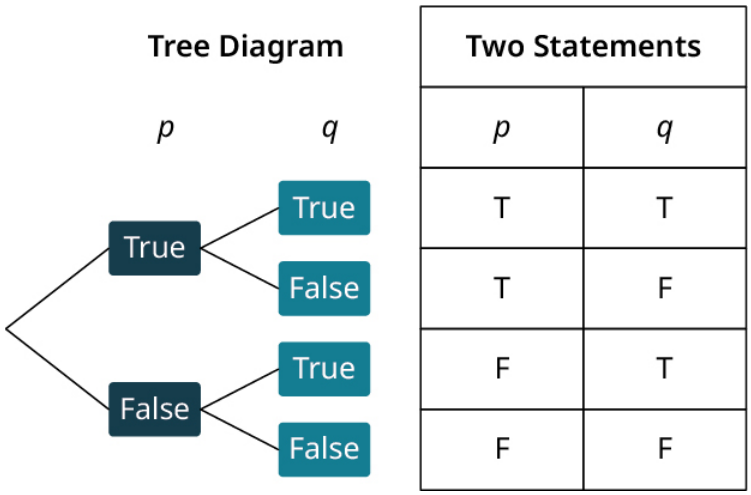


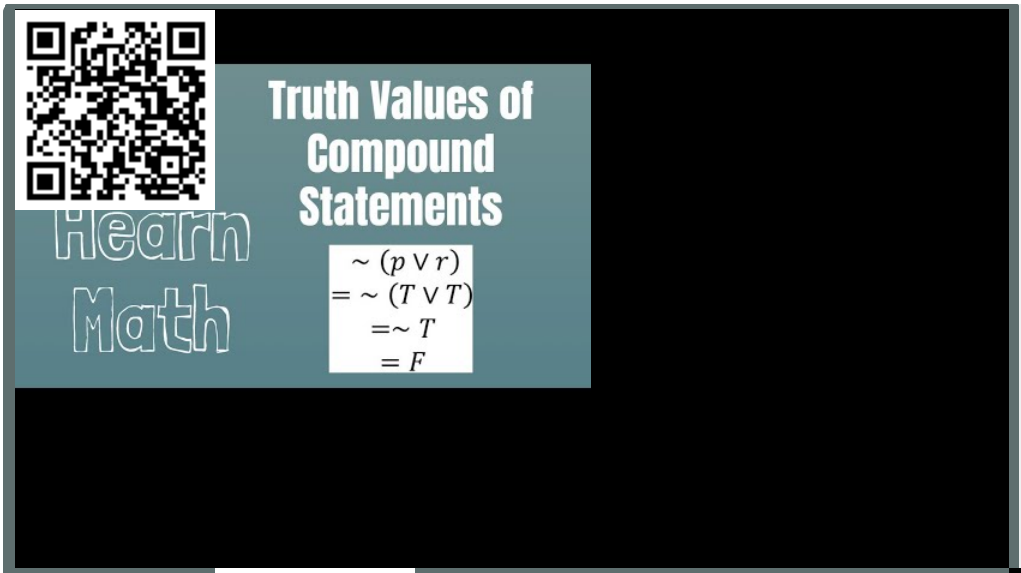
Figure $1.2.1.2$

A conjunction is a logical *and* statement. For a conjunction to be true, both the statements that make up the conjunction must be true. If at least one of the statements is false, the *and* statement is false.

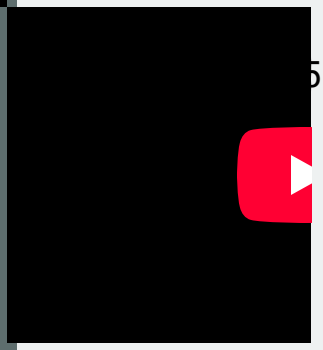
A disjunction is a logical inclusive *or* statement. Which means that a disjunction is true if one or both statements that make it are true. The only way a logical inclusive *or* statement is false is if both statements that make up the disjunction are false.

Conjunction (AND)	Disjunction (OR)

Conjunction (AND)				Disjunction (OR)		
p	q	$p \wedge q$		p	q	$p \vee q$
T	T	T		T	T	T
T	F	F		T	F	T
F	T	F		F	T	T
F	F	F		F	F	F



Logic Part 4: Tr ompound Statements with "and", "or", and "not"



Logic Part 5: What are truth tables? How do you set them up?

? Exercise 1.2.1.5: Constructing Truth Tables to Analyze Compound Statements

Construct a truth table to analyze all possible outcomes for each of the following arguments.

1. $p \vee \sim q$
2. $\sim (p \wedge q)$
3. $(p \vee \sim q) \wedge r$

Answer

1. Step 1: Because there are two basic statements, p and q, and each of these has two possible outcomes, we will have $2(2) = 4$ rows in our table to represent all possible outcomes: TT, TF, FT, and FF. The columns will include $p, q, \sim q$, and $p \vee \sim q$.
Step 2: Every value in column $\sim q$ will have the opposite truth value of the corresponding value in column q : F, T, F, and T.

Step 3: To complete the last column, evaluate each element in column p with the corresponding element in column $\sim q$ using the connective or.

p	q	$\sim q$	$p \vee \sim q$.
T	T	F	T
T	F	T	T
F	T	F	F
F	F	T	T

2. Step 1: The columns will include $p, q, p \wedge q$, and $\sim (p \wedge q)$. Because there are two basic statements, p and q, the table will have four rows to account for all possible outcomes.
Step 2: The $p \wedge q$ column will be completed by evaluating the corresponding elements in columns p and q respectively with the and connective.
Step 3: The final column, $\sim (p \wedge q)$, will be the negation of the $p \wedge q$ column.

p	q	$p \wedge q$	$\sim (p \wedge q)$
T	T	T	F
T	F	F	T
F	T	F	T
F	F	F	T

3. Step 1: This statement has three basic statements, p, q, and r. Because each basic statement has two possible truth values, true or false, the multiplication principal indicates there are $2(2)(2) = 8$ possible outcomes. So eight rows of outcomes are needed in the truth table to account for each possibility. Half of the eight possibilities must be true for the first statement, and half must be false.
Step 2: So, the first column for statement p, will have four T's followed by four F's. In the second column for statement q, when p is true, half the outcomes for q must be true and the other half must be false, and the same pattern will repeat for when p is false. So, column q will have TT, FF, FF, FF.

Step 3: The column for the third statement, r, must alternate between T and F. Once, the three basic propositions are listed, you will need a column for $\sim, p \vee \sim q$, and $(p \vee \sim q) \wedge r$.
Step 4: The column for the negation of $q, \sim q$, will have the opposite truth value of each value in column q.
Step 5: Next, fill in the truth values for the column containing the statement $p \vee \sim q$. The or statement is true if at least one of p or $\sim q$ is true, otherwise it is false.
Step 6: Finally, fill in the column containing the conjunction $(p \vee \sim q) \wedge r$. To evaluate this statement, combine column $p \vee \sim q$ and column r with the and connective. Recall, that only time "and" is true is when both values are true, otherwise the statement is false. The complete truth table is:

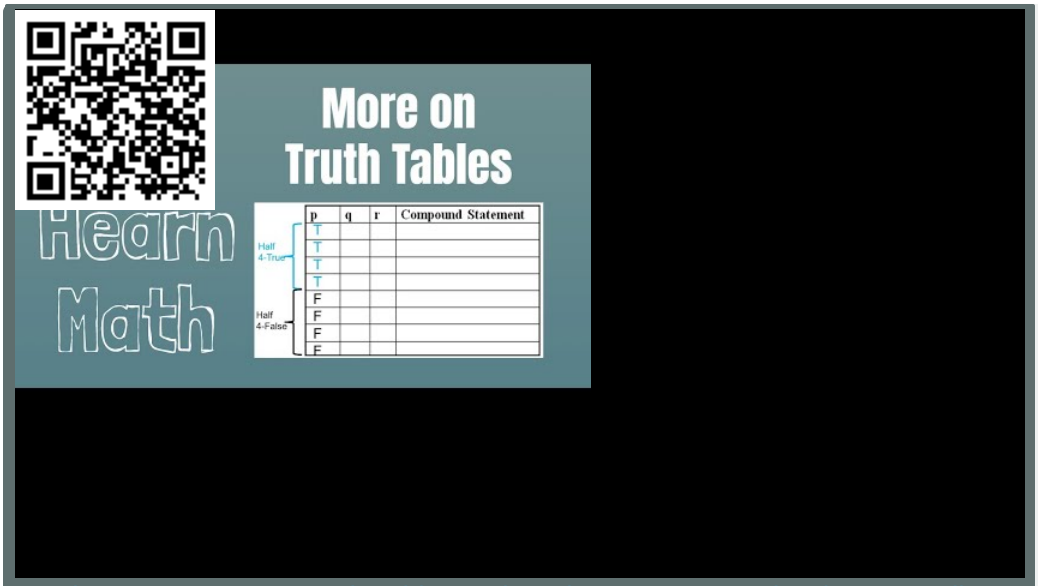
p	q	r	$\sim q$	$p \vee \sim q$	$(p \vee \sim q) \wedge r$
T	T	T	F	T	T
T	T	F	F	T	F
T	F	T	T	T	T
T	F	F	T	T	F
F	T	T	F	F	F
F	T	F	F	F	F
F	F	T	T	T	T

p	q	r	$\sim q$	$p \vee \sim q$	$(p \vee \sim q) \wedge r$
F	F	F	T	T	F

Your Turn 1.2.1.5

Construct a truth table to analyze all possible outcomes for each of the following arguments.

1. $p \wedge \sim q$
2. $\sim (p \vee q)$
3. $(p \wedge \sim q) \vee r$



Logic Part 6: More on Truth Tables and Setting Up Rows and Column Headings

Determine the Validity of a Truth Table for a Compound Statement

A logical statement is valid if it is always true regardless of the truth values of its component parts. To test the validity of a compound statement, construct a truth table to analyze all possible outcomes. If the last column, representing the complete statement, contains only true values, the statement is valid.

Exercise 1.2.1.6: Determining the Validity of Compound Statements

Construct a truth table to determine the validity of each of the following statements.

1. $\sim p \wedge q$
2. $\sim (p \wedge \sim p)$

Answer

1. Step 1: Because there are two statements, p and q, and each of these has two possible outcomes, there will be $2(2) = 4$ rows in our table to represent all possible outcomes: TT, TF, FT, and FF.
Step 2: The columns, will include $p, q, \sim p$ and $\sim p \wedge q$. Every value in column $\sim p$ will have the opposite truth value of the corresponding value in column p.
Step 3: To complete the last column, evaluate each element in column $\sim p$ with the corresponding element in column q using the connective and. The last column contains at least one false, therefore the statement $\sim p \wedge q$ is not valid.

p	q	$\sim p$	$\boldsymbol{\sim p \wedge q}$
T	T	F	F

p	q	$\sim p$	$\sim p \wedge q$
T	F	F	F
F	T	T	T
F	F	T	F

2. Step 1: Because the statement $\sim (p \wedge \sim p)$ only contains one basic proposition, the truth table will only contain two rows. Statement p may be either true or false.
Step 2: The columns will include $p, \sim p, p \wedge \sim p$, and $\sim (p \wedge \sim p)$. Evaluate column $p \wedge \sim p$ with the and connective, because the symbol $\wedge$ represents a conjunction or logical and statement. True and false is false, and false and true is also false.
Step 3: The final column is the negation of each entry in the third column, both of which are false, so the negation of false is true. Because all the truth values in the final column are true, the statement $\sim (p \wedge \sim p)$ is valid.

p	$\sim p$	$p \wedge \sim p$	$\sim (p \wedge \sim p)$
T	F	F	T
F	T	F	T

Your Turn1.2.1.6

Construct a truth table to determine the validity of each of the following statements.
1. $p \vee \sim p$
2. $\sim p \vee \sim q$

Check Your Understanding

1. A logical argument is _____ if its conclusion follows from its premises.
2. A logical statement is valid if it is always _____.
3. A _____ _____ is a tool used to analyze all the possible outcomes for a logical statement.
4. The truth table for the conjunction, $p \wedge q$, has _____ rows of truth values.
5. The truth table for the negation of a logical statement, $\sim p$, has _____ rows of truth values.

This page titled 1.2.1: Constructing Truth Tables is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **2.3: Constructing Truth Tables** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

1.2.2: Truth Tables for the Conditional and Biconditional

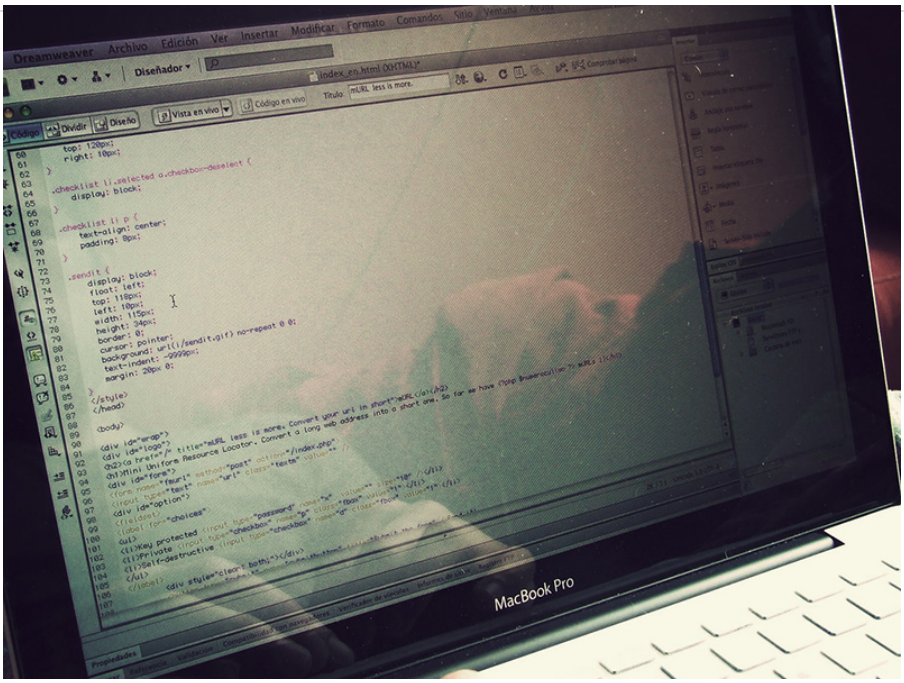


Figure 1.2.2.1: If-then statements use logic to execute directions. (credit: “Coding” by Carlos Varela/Flickr, CC BY 2.0)

Learning Objectives

After completing this section, you should be able to:

1. Use and apply the conditional to construct a truth table.
2. Use and apply the biconditional to construct a truth table.
3. Use truth tables to determine the validity of conditional and biconditional statements.

Computer languages use if-then or if-then-else statements as decision statements:

- *If* the hypothesis is true, *then* do something.
- Or, *if* the hypothesis is true, *then* do something; else do something else.

For example, the following representation of computer code creates an if-then-else decision statement:

Check value of variable i.

If $i < 1$, then print "Hello, World!" else print "Goodbye".

In this imaginary program, the if-then statement evaluates and acts on the value of the variable i. For instance, if $i = 0$, the program would consider the statement $i < 1$ as true and “Hello, World!” would appear on the computer screen. If instead, $i = 3$, the program would consider the statement $i < 1$ as false (because 3 is greater than 1), and print “Goodbye” on the screen.

In this section, we will apply similar reasoning without the use of computer programs.

People in Mathematics

The Countess of Lovelace, Ada Lovelace, is credited with writing the first computer program. She wrote an algorithm to work with Charles Babbage’s Analytical Engine that could compute the Bernoulli numbers in 1843. In doing so, she became the first person to write a program for a machine that would produce more than just a simple calculation. The computer programming language ADA is named after her.

Reference: Posamentier, Alfred and Spreitzer Christian, “Chapter 34 Ada Lovelace: English (1815-1852)” pp. 272-278, *Math Makers: The Lives and Works of 50 Famous Mathematicians*, Prometheus Books, 2019.

Use and Apply the Conditional to Construct a Truth Table

A conditional is a logical statement of the form if p, then q. The conditional statement in logic is a promise or contract. The only time the conditional, $p \rightarrow q$, is false is when the contract or promise is broken.

For example, consider the following scenario. A child's parent says, "If you do your homework, then you can play your video games." The child really wants to play their video games, so they get started right away, finish within an hour, and then show their parent the completed homework. The parent thanks the child for doing a great job on their homework and allows them to play video games. Both the parent and child are happy. The contract was satisfied; true implies true is true.

Now, suppose the child does not start their homework right away, and then struggles to complete it. They eventually finish and show it to their parent. The parent again thanks the child for completing their homework, but then informs the child that it is too late in the evening to play video games, and that they must begin to get ready for bed. Now, the child is really upset. They held up their part of the contract, but they did not receive the promised reward. The contract was broken; true implies false is false.

So, what happens if the child does not do their homework? In this case, the hypothesis is false. No contract has been entered, therefore, no contract can be broken. If the conclusion is false, the child does not get to play video games and might not be happy, but this outcome is expected because the child did not complete their end of the bargain. They did not complete their homework. False implies false is true. The last option is not as intuitive. If the parent lets the child play video games, even if they did not do their homework, neither parent nor child are going to be upset. False implies true is true.

The truth table for the conditional statement below summarizes these results.

p	q	$p \rightarrow q$
T	T	**T**
T	F	**F**
F	T	**T**
F	F	**T**

Checkpoint

Notice that the only time the conditional statement, $p \rightarrow q$, is false is when the hypothesis, p, is true and the conclusion, q, is false.

Video



Logic Part 8: The Conditional and Tautologies

? Exercise 1.2.2.1: Constructing Truth Tables for Conditional Statements

Assume both of the following statements are true: p: My sibling washed the dishes, and q: My parents paid them $5.00. Create a truth table to determine the truth value of each of the following conditional statements.

1. $p \rightarrow q$
2. $p \rightarrow \sim q$

3. $\sim p \rightarrow q$

Answer

Because p is true and q is true, the statement p → q is, “If my sibling washed the dishes, then my parents paid them $5.00.” My sibling did wash the dishes, since p p is true, and the parents did pay the sibling $5.00, so the contract was entered and completed. The conditional statement is true, as indicated by the truth table representing this case:
T → T = T.

p	q	$p \rightarrow q$
T	T	**T**

p → ~ q p

2. p → ~ q p → ~ q translates to the statement, “If my sibling washed the dishes, then my parents did not pay them $5.00.” p p is true, but ~ q ~ q is false. The sibling completed their end of the contract, but they did not get paid. The contract was broken by the parents. The conditional statement is false, as indicated by the truth table representing this case:

T → F = F.

p	q	$\sim q$	$p \rightarrow \sim q$
T	T	F	**F**

3. ~ p → q translates to the statement, “If my sibling did not wash the dishes, then my parents paid them $5.00.” ~ p is false, but q is true. The sibling did not do the dishes. No contract was entered, so it could not be broken. The parents decided to pay them $5.00 anyway. The conditional statement is true, as indicated by the truth table representing this case: F → T = T.

p	q	$\sim p$	$\sim p \rightarrow q$
T	T	F	**T**

Your Turn 1.2.2.1

Assume p is true and q is false. p: Kevin vacuumed the living room, and q: Kevin's parents did not let him borrow the car. Create a truth table to determine the truth value of each of the following conditional statements.

1. $p \rightarrow q$
2. $p \rightarrow \sim q$
3. $\sim p \rightarrow q$

Exercise 1.2.2.2: Determining Validity of Conditional Statements

Construct a truth table to analyze all possible outcomes for each of the following statements then determine whether they are valid.

1. $p \wedge q \rightarrow \sim q$
2. $p \rightarrow \sim p \vee q$

Answer

1. Applying the dominance of connectives, the statement $p \wedge q \rightarrow \sim q$ is equivalent to $(p \wedge q) \rightarrow (\sim q)$. So, the columns of the truth table will include p, q, $p \wedge q$, $\sim q$, and $p \wedge q \rightarrow \sim q$. Because there are only two basic propositions, p and q, the table will have 2 (2) = 4 2 (2) = 4 rows of truth values to account for all the possible outcomes. The statement is not valid because the last column is not all true.

p	q	$p \wedge q$	$\sim q$	$p \wedge q \rightarrow \sim q$
T	T	T	F	**F**

p	q	$p \wedge q$	$\sim q$	$p \wedge q \rightarrow \sim q$
T	F	F	T	**T**
F	T	F	F	**T**
F	F	F	T	**T**

Applying the dominance of connectives, the statement $p \rightarrow \sim p \vee q$ is equivalent to $(p) \rightarrow ((\sim p) \vee q)$. So, the columns of the truth table will include p , q , ~ p, ~ p ∨ q, and $p \rightarrow (\sim p \vee q)$. Because there are only two basic propositions, p and q , the table will have 2 (2) = 4 2 (2) = 4 rows of truth values to account for all the possible outcomes. The statement is not valid because the last column is not all true.

p	q	$\sim p$	$\sim p \vee q$	$p \rightarrow (\sim p \vee q)$
T	T	F	T	**T**
T	F	F	F	**F**
F	T	T	T	**T**
F	F	T	T	**T**

Your Turn 1.2.2.2

Construct a truth table to analyze all possible outcomes for each of the following statements, then determine whether they are valid.

1. $q \rightarrow \sim p \vee q$
2. $\sim p \rightarrow q \wedge p$

Use and Apply the Biconditional to Construct a Truth Table

The biconditional, $p \leftrightarrow q$, is a two way contract; it is equivalent to the statement $(p \rightarrow q) \wedge (q \rightarrow p)$. A biconditional statement, $p \leftrightarrow q$, is true whenever the truth value of the hypothesis matches the truth value of the conclusion, otherwise it is false.

The truth table for the biconditional is summarized below.

p	q	$p \leftrightarrow q$
T	T	**T**
T	F	**F**
F	T	**F**
F	F	**T**

Video



Exercise 1.2.2.3: Constructing Truth Tables for Biconditional Statements

Assume both of the following statements are true: p : The plumber fixed the leak, and q : The homeowner paid the plumber $150.00. Create a truth table to determine the truth value of each of the following biconditional statements.

1. $p \leftrightarrow q$
2. $p \leftrightarrow \sim q$
3. $\sim p \leftrightarrow \sim q$

Answer

1. Because p p is true and q q is true, the statement p ↔ q p ↔ q is "The plumber fixed the leak if and only if the homeowner paid them $150.00." Because both p p and q q are true, the leak was fixed and the plumber was paid, meaning both parties satisfied their end of the bargain. The biconditional statement is true, as indicated by the truth table representing this case: T ↔ T = T.

p	q	$p \leftrightarrow q$
T	T	**T**

2. p → ~ q p → ~ q translates to the statement, "The plumber fixed the leak if and only if the homeowner did not pay them $150." If the plumber fixed the leak and the homeowner did not pay them, the homeowner will have broken their end of the contract. The biconditional statement is false, as indicated by the truth table representing this case:
T ↔ F = F.

p	q	$\sim q$	$p \leftrightarrow \sim q$
T	T	F	**F**

3. ~ p ↔ ~ q ~ p ↔ ~ q translates to the statement, "The plumber did not fix the leak if and only if the homeowner did not pay them $150." In this case, neither party—the plumber nor the homeowner—entered into the contract. The leak was not repaired, and the plumber was not paid. No agreement was broken. The biconditional statement is true, as indicated by the truth table representing this case: F ↔ F = T.

p	q	$\sim p$	$\sim q$	$\sim p \leftrightarrow \sim q$
T	T	F	F	**T**

Your Turn 1.2.2.3

Assume p is true and q is false: p : The contractor fixed the broken window, and q : The homeowner paid the contractor $200. Create a truth table to determine the truth value of each of the following biconditional statements.

1. $p \leftrightarrow q$
2. $p \leftrightarrow \sim q$
3. $\sim p \leftrightarrow q$

Checkpoint

The biconditional, $p \leftrightarrow q$, is true whenever the truth values of p and q match, otherwise it is false.

Video



Logic Part 13: Truth Tables to Determine if Argument is Valid or Invalid

Exercise 1.2.2.1: Determining Validity of Biconditional Statements

Construct a truth table to analyze all possible outcomes for each of the following statements, then determine whether they are valid.

1. $p \wedge q \leftrightarrow p \wedge \sim q$
2. $p \vee q \leftrightarrow \sim p \vee q$
3. $p \rightarrow q \leftrightarrow \sim q \rightarrow \sim p$
4. $p \wedge q \rightarrow \sim r \leftrightarrow p \wedge q \wedge r$

Answer

1. Applying the dominance of connectives, the statement $p \wedge q \leftrightarrow p \wedge \sim q$ is equivalent to $(p \wedge q) \leftrightarrow (p \wedge (\sim q))$. So, the columns of the truth table will include $p, q, p \wedge q, \sim q, p \wedge \sim q$ and $(p \wedge q) \leftrightarrow (p \wedge \sim q)$. Because there are only two basic propositions, p and q, the table will have $2(2) = 4$ rows of truth values to account for all the possible outcomes. The statement is not valid because the last column is not all true.

p	q	$p \wedge q$	$\sim q$	$p \wedge \sim q$	$(p \wedge q) \leftrightarrow (p \wedge \sim q)$
T	T	T	F	F	**F**
T	F	F	T	T	**F**
F	T	F	F	F	**T**
F	F	F	T	F	**T**

2. Applying the dominance of connectives, the statement $p \vee q \leftrightarrow \sim p \vee q$ is equivalent to $(p \vee q) \leftrightarrow ((\sim p) \vee q)$. So, the columns of the truth table will include $p, q, p \vee q, \sim p, \sim p \vee q$, and $(p \vee q) \leftrightarrow (\sim p \vee q)$. Because there are only two basic propositions, p and q, the table will have $2(2) = 4$ rows of truth values to account for all the possible outcomes. The statement is not valid because the last column is not all true.

p	q	$p \vee q$	$\sim p$	$\sim p \vee q$	$(p \vee q) \leftrightarrow (\sim p \vee q)$
T	T	T	F	T	**T**
T	F	T	F	F	**F**
F	T	T	T	T	**T**
F	F	F	T	T	**F**

3. Applying the dominance of connectives, the statement $p \rightarrow q \leftrightarrow \sim q \rightarrow \sim p$ is equivalent to $(p \rightarrow q) \leftrightarrow ((\sim q) \rightarrow (\sim p))$. So, the columns of the truth table will include $p, q, p \rightarrow q, \sim q, \sim p, \sim q \rightarrow \sim p$, and $(p \rightarrow q) \leftrightarrow (\sim q \rightarrow \sim p)$. Because there are only two basic propositions, p and q the table will have $2(2) = 4$ rows of truth values to account for all the possible outcomes. The statement is valid because the last column is all true.

p	q	$p \to q$	$\sim q$	$\sim p$	$\sim q \to \sim p$	$(p \to q) \leftrightarrow (\sim q \to \sim p)$
T	T	T	F	F	T	**T**
T	F	F	T	F	F	**T**
F	T	T	F	T	T	**T**
F	F	T	T	T	T	**T**

4. Applying the dominance of connectives, the statement $p \wedge q \to \sim r \leftrightarrow p \wedge q \wedge r$ is equivalent to $((p \wedge q) \to (\sim r)) \leftrightarrow ((p \wedge q) \wedge r)$. So, the columns of the truth table will include $p, q, r, \sim r, p \wedge q, (p \wedge q) \wedge \sim r, (p \wedge q) \wedge r$, and $((p \wedge q) \to (\sim r)) \leftrightarrow ((p \wedge q) \wedge r)$. Because there are three basic propositions, p, q, and r, the table will have $2(2)(2) = 8$ rows of truth values to account for all the possible outcomes. The statement is not valid because the last column is not all true.

p	q	r	$\sim r$	$p \wedge q$	$(p \wedge q) \to \sim r$	$(p \wedge q) \wedge r$	$(p \wedge q \to \sim r) \leftrightarrow (p \wedge$
T	T	T	F	T	F	T	**T**
T	T	F	T	T	T	F	**T**
T	F	T	F	F	T	F	**F**
T	F	F	T	F	T	F	**F**
F	T	T	F	F	T	F	**F**
F	T	F	T	F	T	F	**F**
F	F	T	F	F	T	F	**F**
F	F	F	T	F	T	F	**F**

Your Turn 1.2.2.1

Construct a truth table to analyze all possible outcomes for each of the following statements, then determine whether they are valid.

1. $\sim (p \wedge q) \leftrightarrow (\sim p \vee \sim q)$
2. $\sim p \leftrightarrow q \wedge p$
3. $p \to q \leftrightarrow \sim p \vee q$
4. $p \wedge q \to r \leftrightarrow \sim p \vee \sim q \vee r$

Check Your Understanding

1. 1. In logic, a conditional statement can be thought of as a _____.
2. If the hypothesis, p, of a conditional statement is true, the _____, q, must also be true for the conditional statement \(p \to q to be true.
3. If the _____ of a conditional statement is false, the conditional statement is true.
4. The symbolic form of the _____ statement is $p \leftrightarrow q$
5. The _____ statement is equivalent to the statement $(p \to q) \wedge (q \to p)$.
6. p if and only if q is _____ whenever the truth value of p matches the truth value of q, otherwise it is false.

This page titled 1.2.2: Truth Tables for the Conditional and Biconditional is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **2.4: Truth Tables for the Conditional and Biconditional** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

1.3: Conditional Statements

1.3: Conditional Statements is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

1.3.1: Equivalent Statements



Figure 1.3.1.1: How your logical argument is stated affects the response, just like how you speak when holding a conversation can affect how your words are received. (credit: modification of work by Goelshivi/Flickr, Public Domain Mark 1.0)

Learning Objectives

After completing this section, you should be able to:

1. Determine whether two statements are logically equivalent using a truth table.
2. Compose the converse, inverse, and contrapositive of a conditional statement

Have you ever had a conversation with or sent a note to someone, only to have them misunderstand what you intended to convey? The way you choose to express your ideas can be as, or even more, important than what you are saying. If your goal is to convince someone that what you are saying is correct, you will not want to alienate them by choosing your words poorly.

Logical arguments can be stated in many different ways that still ultimately result in the same valid conclusion. Part of the art of constructing a persuasive argument is knowing how to arrange the facts and conclusion to elicit the desired response from the intended audience.

In this section, you will learn how to determine whether two statements are **logically equivalent** using truth tables, and then you will apply this knowledge to compose logically equivalent forms of the conditional statement. Developing this skill will provide the additional skills and knowledge needed to construct well-reasoned, persuasive arguments that can be customized to address specific audiences.

Checkpoint

An alternate way to think about logical equivalence is that the truth values have to match. That is, whenever p is true, q is also true, and whenever p is false, q is also false.

Determine Logical Equivalence

Two statements, p and q, are logically equivalent when $p \leftrightarrow q$ is a valid argument, or when the last column of the truth table consists of only true values. When a logical statement is always true, it is known as a tautology. To determine whether two statements p and q are logically equivalent, construct a truth table for $p \leftrightarrow q$ and determine whether it valid. If the last column is all true, the argument is a tautology, it is valid, and p is logically equivalent to q; otherwise, p is not logically equivalent to q.

Exercise 1.3.1.1: Determining Logical Equivalence with a Truth Table

Create a truth table to determine whether the following compound statements are logically equivalent.

1. $p \rightarrow q$; $\sim p \rightarrow \sim q$
2. $p \rightarrow q$; $\sim p \vee q$

Answer

1. Construct a truth table for the biconditional formed by using the first statement as the hypothesis and the second statement as the conclusion, $(p \rightarrow q) \leftrightarrow (\sim p \rightarrow \sim q)$.

p	q	$p \rightarrow q$	$\sim p$	$\sim q$	$\sim p \rightarrow \sim q$	$(p \rightarrow q) \leftrightarrow (\sim p \rightarrow \sim q$
T	T	T	F	F	T	**T**
T	F	F	F	T	T	**F**
F	T	T	T	F	F	**F**
F	F	T	T	T	T	**T**

Because the last column it not all true, the biconditional is not valid and the statement $p \rightarrow q$ is not logically equivalent to the statement $\sim p \rightarrow \sim q$.

2. Construct a truth table for the biconditional formed by using the first statement as the hypothesis and the second statement as the conclusion,
$(p \rightarrow q) \leftrightarrow (\sim p \vee q)$.

p	q	$p \rightarrow q$	$\sim p$	$\sim p \vee q$	$(p \rightarrow q) \leftrightarrow (\sim p \vee q)$
T	T	T	F	T	**T**
T	F	F	F	F	**T**
F	T	T	T	T	**T**
F	F	T	T	T	**T**

Because the last column is true for every entry, the biconditional is valid and the statement $p \rightarrow q$ is logically equivalent to the statement $\sim p \vee q$. Symbolically, $p \rightarrow q \equiv \sim p \vee q$.

Your Turn 1.3.1.1

Create a truth table to determine whether the following compound statements are logically equivalent.

1. $p \rightarrow q; q \rightarrow \sim p$
2. $p \rightarrow q; p \vee \sim q$

Compose the Converse, Inverse, and Contrapositive of a Conditional Statement

The converse, inverse, and contrapositive are variations of the conditional statement, $p \rightarrow q$.

- The converse is if q then p, and it is formed by interchanging the hypothesis and the conclusion. The converse is logically equivalent to the inverse.
- The inverse is if $\sim p$ then $\sim q$, and it is formed by negating both the hypothesis and the conclusion. The inverse is logically equivalent to the converse.
- The contrapositive is if $\sim q$ then $\sim p$, and it is formed by interchanging and negating both the hypothesis and the conclusion. The contrapositive is logically equivalent to the conditional.

The table below shows how these variations are presented symbolically.

				Conditional	Contrapositive	Converse	Inverse
p	q	$\sim p$	$\sim q$	$p \rightarrow q$	$\sim q \rightarrow \sim p$	$q \rightarrow p$	$\sim p \rightarrow \sim q$
T	T	F	F	T	T	T	T
T	F	F	T	F	F	T	T
F	T	T	F	T	T	F	F
F	F	T	T	T	T	T	T

Exercise 1.3.1.2: Writing the Converse, Inverse, and Contrapositive of a Conditional Statement

Use the statements, p: Harry is a wizard and q: Hermione is a witch, to write the following statements:

1. Write the conditional statement, $p \rightarrow q$, in words.
2. Write the converse statement, $q \rightarrow p$, in words.
3. Write the inverse statement, $\sim p \rightarrow \sim q$, in words.
4. Write the contrapositive statement, $\sim q \rightarrow \sim p$, in words.

Answer

1. The conditional statement takes the form, “if p p , then q q ,” so the conditional statement is: “If Harry is a wizard, then Hermione is a witch.” Remember the *if … then …* words are the connectives that form the conditional statement.
2. The converse swaps or interchanges the hypothesis, p p , with the conclusion, q q . It has the form, “if q q , then p p .” So, the converse is: "If Hermione is a witch, then Harry is a wizard."
3. To construct the inverse of a statement, negate both the hypothesis and the conclusion. The inverse has the form, “if ~ p ~ p , then ~ q ~ q ,” so the inverse is: "If Harry is not a wizard, then Hermione is not a witch."
4. The contrapositive is formed by negating and interchanging both the hypothesis and conclusion. It has the form, “if ~ q ~ q , then ~ p ~ p ," so the contrapositive statement is: "If Hermione is not a witch, then Harry is not a wizard."

Your Turn 1.3.1.2

Use the statements, p : Elvis Presley wore capes and q : Some superheroes wear capes, to write the following statements:

1. Write the conditional statement, $p \rightarrow q$, in words.
2. Write the converse statement, $q \rightarrow p$, in words.
3. Write the inverse statement, $\sim p \rightarrow \sim q$, in words.
4. Write the contrapositive statement, $\sim q \rightarrow \sim p$, in words.

Exercise 1.3.1.3

Use the conditional statement, “If all dogs bark, then Lassie likes to bark,” to identify the following.

1. Write the hypothesis of the conditional statement and label it with a p.
2. Write the conclusion of the conditional statement and label it with a q.
3. Identify the following statement as the converse, inverse, or contrapositive: “If Lassie likes to bark, then all dogs bark.”
4. Identify the following statement as the converse, inverse, or contrapositive: “If Lassie does not like to bark, then some dogs do not bark.”
5. Which statement is logically equivalent to the conditional statement?

Answer

1. The hypothesis is the phrase following the *if*. The answer is p: All dogs bark. Notice, the word *if* is not included as part of the hypothesis.
2. The conclusion of a conditional statement is the phrase following the *then*. The word *then* is not included when stating the conclusion. The answer is: q: Lassie likes to bark.

3. “Lassie likes to bark” is q and “All dogs bark” is p. So, “If Lassie likes to bark, then all dogs bark,” has the form “if q, then p,” which is the form of the converse. “Lassie does not like to bark” is $\sim q$ and “Some dogs do not bark” is $\sim p$. The statement, “If Lassie does not like to bark, then some dogs do not bark,” has the form “if $\sim q$, then $\sim p$,” which is the form of the contrapositive.
4. The contrapositive $\sim q \rightarrow \sim p$ is logically equivalent to the conditional statement $p \rightarrow q$.

Your Turn 1.3.1.3

Use the conditional statement, “If Dora is an explorer, then Boots is a monkey,” to identify the following:

1. Write the hypothesis of the conditional statement and label it with a p.

2. Write the conclusion of the conditional statement and label it with a q.

3. Identify the following statement as the converse, inverse, or contrapositive: “If Dora is not an explorer, then Boots is not a monkey.”

4. Identify the following statement as the converse, inverse, or contrapositive: “If Boots is a monkey, then Dora is an explorer.”

5. Which statement is logically equivalent to the inverse?

Exercise 1.3.1.4

Assume the conditional statement, p→q: “If Chadwick Boseman was an actor, then Chadwick Boseman did not star in the movie *Black Panther*” is false, and use it to answer the following questions.

1. Write the converse of the statement in words and determine its truth value.
2. Write the inverse of the statement in words and determine its truth value.
3. Write the contrapositive of the statement in words and determine its truth value.

Answer

The only way the conditional statement can be false is if the hypothesis, p: Chadwick Boseman was an actor, is true and the conclusion, q: Chadwick Boseman did not star in the movie *Black Panther*, is false. The converse is q → p, and it is written in words as: “If Chadwick Boseman did not star in the movie *Black Panther*, then Chadwick Boseman was an actor.” This statement is true, because false → true is true.

The inverse has the form “ $\sim p \rightarrow \sim q$.” The written form is: “If Chadwick Boseman was not an actor, then Chadwick Boseman starred in the movie *Black Panther*.” Because p is true, and q is false, $\sim p$ is false, and $\sim q$ is true. This means the inverse is false → true, which is true. Alternatively, from Question 1, the converse is true, and because the inverse is logically equivalent to the converse it must also be true.

The contrapositive is logically equivalent to the conditional. Because the conditional is false, the contrapositive is also false. This can be confirmed by looking at the truth values of the contrapositive statement. The contrapositive has the form “ $\sim q \rightarrow \sim p$.” Because q is false and p is true, ~ q is true and $\sim p$ is false. Therefore, $\sim q \rightarrow \sim p$ is true → false, which is false. The written form of the contrapositive is “If Chadwick Boseman starred in the movie *Black Panther*, then Chadwick Boseman was not an actor.”

Your Turn 1.3.1.4

Assume the conditional statement $p \rightarrow q$: “If my friend lives in San Francisco, then my friend does not live in California” is false, and use it to answer the following questions.

1. Write the converse of the statement in words and determine its truth value.

2. Write the inverse of the statement in words and determine its truth value.

3. Write the contrapositive of the statement in words and determine its truth value.

Check Your Understanding

1. Two statements p and q are logically equivalent to each other if the biconditional statement, $p \leftrightarrow q$ is _____.
2. The _____ statement has the form, "p then q."
3. The contrapositive is _____ _____ to the conditional statement, and has the form, "if $\sim q$, then $\sim p$."
4. The _____ of the conditional statement has the form, "if $\sim p$, then $\sim q$."
5. The _____ of the conditional statement is logically equivalent to the _____ and has the form, "if q then p."

This page titled 1.3.1: Equivalent Statements is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **2.5: Equivalent Statements** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

1.4: Deductive Arguments

1.4: Deductive Arguments is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

1.4.1: Evaluating Deductive Arguments with Truth Tables

Arguments can also be analyzed using truth tables, although this can be a lot of work.

Analyzing arguments using truth tables

To analyze an argument with a truth table:

1. Represent each of the premises symbolically
2. Create a conditional statement, joining all the premises to form the antecedent, and using the conclusion as the consequent.
3. Create a truth table for the statement. If it is always true, then the argument is valid.

Example 34

Consider the argument

Premise:	If you bought bread, then you went to the store.
Premise:	You bought bread.
Conclusion:	You went to the store.

Solution

While this example is fairly obviously a valid argument, we can analyze it using a truth table by representing each of the premises symbolically. We can then form a conditional statement showing that the premises together imply the conclusion. If the truth table is a tautology (always true), then the argument is valid.

We'll let b represent "you bought bread" and s represent "you went to the store". Then the argument becomes:

Premise:	$b \to s$
Premise:	b
Conclusion:	s

To test the validity, we look at whether the combination of both premises implies the conclusion; is it true that $[(b \to s) \wedge b] \to s$?

b	s	$b \to s$
T	T	T
T	F	F
F	T	T
F	F	T

b	s	$b \to s$	$(b \to s) \wedge b$
T	T	T	T
T	F	F	F
F	T	T	F
F	F	T	F

b	s	$b \to s$	$(b \to s) \wedge b$	$[(b \to s) \wedge b] \to s$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

Since the truth table for $[(b \to s) \wedge b] \to s$ is always true, this is a valid argument.

Try it Now 13

Determine whether the argument is valid:

Premise: If I have a shovel, I can dig a hole.
Premise: I dug a hole.
Conclusion: Therefore, I had a shovel.

Answer

Let $S =$ have a shovel, $D =$ dig a hole. The first premise is equivalent to $S \to D$. The second premise is D. The conclusion is S. We are testing $[(S \to D) \wedge D] \to S$

S	D	$S \to D$	$(S \to D) \wedge D$	$[(S \to D) \wedge D] \to S$
T	T	T	T	T
T	F	F	F	T
F	T	T	T	F
F	F	T	F	T

This is not a tautology, so this is an invalid argument.

Example 35

Premise: If I go to the mall, then I'll buy new jeans.
Premise: If I buy new jeans, I'll buy a shirt to go with it.
Conclusion: If I go to the mall, I'll buy a shirt.

Solution

Let $m =$ I go to the mall, $j =$ I buy jeans, and $s =$ I buy a shirt.

The premises and conclusion can be stated as:

Premise: $m \to j$
Premise: $j \to s$
Conclusion: $m \to s$

We can construct a truth table for $[(m \to j) \wedge (j \to s)] \to (m \to s)$. Try to recreate each step and see how the truth table was constructed.

m	j	s	$m \to j$	$j \to s$	$(m \to j) \wedge (j \to s)$	$m \to s$	$[(m \to j) \wedge (j \to s)] \to (m \to s)$
T	T	T	T	T	T	T	T
T	T	F	T	F	F	F	T
T	F	T	F	T	F	T	T
T	F	F	F	T	F	F	T
F	T	T	T	T	T	T	T
F	T	F	T	F	F	T	T
F	F	T	T	T	T	T	T
F	F	F	T	T	T	T	T

From the final column of the truth table, we can see this is a valid argument.

This page titled 1.4.1: Evaluating Deductive Arguments with Truth Tables is shared under a CC BY-SA 3.0 license and was authored, remixed, and/or curated by David Lippman (The OpenTextBookStore) via source content that was edited to the style and standards of the LibreTexts platform.

- **17.10: Evaluating Deductive Arguments with Truth Tables** by David Lippman is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.

1.4.2: Forms of Valid Arguments

Rather than making a truth table for every argument, we may be able to recognize certain common forms of arguments that are valid (or invalid). If we can determine that an argument fits one of the common forms, we can immediately state whether it is valid or invalid.

The Law of Detachment (*Modus Ponens*)

The law of detachment applies when a conditional and its antecedent are given as premises, and the consequent is the conclusion. The general form is:

Premise:	$p \to q$
Premise:	p
Conclusion:	q

The Latin name, *modus ponens*, translates to "mode that affirms".

Example 36

Recall this argument from an earlier example:

Premise:	If you bought bread, then you went to the store.
Premise:	You bought bread.
Conclusion:	You went to the store.

In symbolic form:

Premise:	$b \to s$
Premise:	b
Conclusion:	s

This argument has the structure described by the law of detachment. (The second premise and the conclusion are simply the two parts of the first premise detached from each other.) Instead of making a truth table, we can say that this argument is valid by stating that it satisfies the law of detachment.

The Law of Contraposition (*Modus Tollens*)

The law of contraposition applies when a conditional and the negation of its consequent are given as premises, and the negation of its antecedent is the conclusion. The general form is:

Premise:	$p \to q$
Premise:	$\sim q$
Conclusion:	$\sim p$

The Latin name, *modus tollens*, translates to "mode that denies".

Notice that the second premise and the conclusion look like the contrapositive of the first premise, $\sim q \to \sim p$, but they have been detached. You can think of the law of contraposition as a combination of the law of detachment and the fact that the contrapositive is logically equivalent to the original statement.

Example 37

Premise:	If I drop my phone into the swimming pool, my phone will be ruined.
Premise:	My phone isn't ruined.
Conclusion:	I didn't drop my phone into the swimming pool.

If we let $d =$ I drop the phone in the pool and $r =$ the phone is ruined, then we can represent the argument this way:

Premise: $d \to r$
Premise: $\sim r$
Conclusion: $\sim d$

The form of this argument matches what we need to invoke the law of contraposition, so it is a valid argument.

Try it Now 14

Is this argument valid?

Premise: If you brushed your teeth before bed, then your toothbrush will be wet.
Premise: Your toothbrush is dry.
Conclusion: You didn't brush your teeth before bed.

Answer

Let $b =$ brushed teeth and $w =$ toothbrush is wet.

Premise: $b \to w$
Premise: $\sim w$
Conclusion: $\sim b$

This argument is valid by the Law of Contraposition.

The Transitive Property (Hypothetical Syllogism)

The transitive property has as its premises a series of conditionals, where the consequent of one is the antecedent of the next. The conclusion is a conditional with the same antecedent as the first premise and the same consequent as the final premise. The general form is:

Premise: $p \to q$
Premise: $q \to r$
Conclusion: $p \to r$

The earlier example about buying a shirt at the mall is an example illustrating the transitive property. It describes a chain reaction: if the first thing happens, then the second thing happens, and if the second thing happens, then the third thing happens. Therefore, if we want to ignore the second thing, we can say that if the first thing happens, then we know the third thing will happen. We don't have to mention the part about buying jeans; we can simply say that the first event leads to the final event. We could even have more than two premises; as long as they form a chain reaction, the transitive property will give us a valid argument.

Example 38

Premise: If a soccer player commits a reckless foul, she will receive a yellow card.
Premise: If Hayley receives a yellow card, she will be suspended for the next match.
Conclusion: If Hayley commits a reckless foul, she will be suspended for the next match.

If we let $r =$ committing a reckless foul, $y =$ receiving a yellow card, and $s =$ being suspended, then our argument looks like this:

Premise: $r \to y$
Premise: $y \to s$
Conclusion: $r \to s$

This argument has the exact structure required to use the transitive property, so it is a valid argument.

Try it Now 15

Is this argument valid?

Premise:	If the old lady swallows a fly, she will swallow a spider.
Premise:	If the old lady swallows a spider, she will swallow a bird.
Premise:	If you brushed your teeth before bed, then your toothbrush will be wet.
Premise:	If you brushed your teeth before bed, then your toothbrush will be wet.
Premise:	If the old lady swallows a bird, she will swallow a cat.
Premise:	If the old lady swallows a cat, she will swallow a dog.
Premise:	If the old lady swallows a dog, she will swallow a goat.
Premise:	If the old lady swallows a goat, she will swallow a cow.
Premise:	If the old lady swallows a cow, she will swallow a horse.
Premise:	If the old lady swallows a horse, she will die, of course.
Conclusion:	If the old lady swallows a fly, she will die, of course.

Answer

This argument is valid by the Transitive Property, which can involve more than two premises, as long as they continue the chain reaction. The premises $f \to s, s \to b, b \to c, c \to d \quad d \to g, g \to w, w \to h, h \to x$ can be reduced to $f \to x$. (Because we had already used c and d we decided to use w for cow and x for death. If the old lady swallows the fly, she will eventually eat a horse and die.

Disjunctive Syllogism

In a disjunctive syllogism, the premises consist of an *or* statement and the negation of one of the options. The conclusion is the other option. The general form is:

Premise:	$p \vee q$
Premise:	$\sim p$
Conclusion:	q

The order of the two parts of the disjunction isn't important. In other words, we could have the premises $p \vee q$ and $\sim q$, and the conclusion p

Example 39

Premise:	I can either drive or take the train.
Premise:	I refuse to drive.
Conclusion:	I will take the train.

If we let $d = I$ drive and $t = I$ take the train, then the symbolic representation of the argument is:

Premise:	$d \vee t$
Premise:	$\sim d$
Conclusion:	t

This argument is valid because it has the form of a disjunctive syllogism. I have two choices, and one of them is not going to happen, so the other one must happen.

Try it Now 16

Is this argument valid?

Premise:	Alison was required to write a 10-page paper or give a 5-minute speech.
Premise:	Alison did not give a 5-minute speech.
Conclusion:	Alison wrote a 10-page paper.

Answer

Let $p =$ wrote a paper and $s =$ gave a speech.

Premise:	$p \vee s$
Premise:	$-s$
Conclusion:	p

This argument is valid by Disjunctive Syllogism. Alison had to do one or the other; she didn't choose the speech, so she must have chosen the paper.

Keep in mind that, when you are determining the validity of an argument, you must assume that the premises are true. If you don't agree with one of the premises, you need to keep your personal opinion out of it. Your job is to pretend that the premises are true and then determine whether they force you to accept the conclusion. You may attack the premises in a court of law or a political discussion, of course, but here we are focusing on the structure of the arguments, not the truth of what they actually say.

We have just looked at four forms of valid arguments; there are two common forms that represent *invalid* arguments, which are also called *fallacies*.

The Fallacy of the Converse

The fallacy of the converse arises when a conditional and its consequent are given as premises, and the antecedent is the conclusion. The general form is:

Premise:	$p \to q$
Premise:	q
Conclusion:	p

Notice that the second premise and the conclusion look like the converse of the first premise, $q \to p$, but they have been detached. The fallacy of the converse incorrectly tries to assert that the converse of a statement is equivalent to that statement.

Example 40

Premise:	If I drink coffee after noon, then I have a hard time falling asleep that night.
Premise:	I had a hard time falling asleep last night.
Conclusion:	I drank coffee after noon yesterday.

If we let $c =$ I drink coffee after noon and $h =$ I have a hard time falling asleep, then our argument looks like this:

Premise:	$c \to h$
Premise:	h
Conclusion:	c

This argument uses converse reasoning, so it is an invalid argument. There could be plenty of other reasons why I couldn't fall asleep: I could be worried about money, my neighbors might have been setting off fireworks, …

Try it Now 17

Is this argument valid?

Premise:	If you pull that fire alarm, you will get in big trouble.
Premise:	You got in big trouble.
Conclusion:	You must have pulled the fire alarm.

Answer

Let $f =$ pulled fire alarm and $t =$ got in big trouble.

Premise:	$f \to t$
Premise:	t
Conclusion:	f

The Fallacy of the Inverse

The fallacy of the inverse occurs when a conditional and the negation of its antecedent are given as premises, and the negation of the consequent is the conclusion. The general form is:

Premise:	$p \to q$
Premise:	$\sim p$
Conclusion:	$\sim q$

Again, notice that the second premise and the conclusion look like the inverse of the first premise, $\sim p \to \sim q$, but they have been detached. The fallacy of the inverse incorrectly tries to assert that the inverse of a statement is equivalent to that statement.

Example 41

Premise:	If you listen to the Grateful Dead, then you are a hippie.
Premise:	Sky doesn't listen to the Grateful Dead.
Conclusion:	Sky is not a hippie.

If we let $g =$ listen to the Grateful Dead and $h =$ is a hippie, then this is the argument:

Premise:	$g \to h$
Premise:	$\sim g$
Conclusion:	$\sim h$

This argument is invalid because it uses inverse reasoning. The first premise does not imply that all hippies listen to the Grateful Dead; there could be some hippies who listen to Phish instead.

Try it Now 18

Is this argument valid?

Premise:	If a hockey player trips an opponent, he will be assessed a 2-minute penalty.
Premise:	Alexei did not trip an opponent.
Conclusion:	Alexei will not be assessed a 2-minute penalty.

Answer

Let $t =$ tripped and $p =$ got a penalty.

Premise:	$t \to p$
Premise:	$\sim t$
Conclusion:	$\sim p$

This argument is invalid because it has the form of the Fallacy of the Inverse. Alexei may have gotten a penalty for an infraction other than tripping.

Of course, arguments are not limited to these six basic forms; some arguments have more premises, or premises that need to be rearranged before you can see what is really happening. There are plenty of other forms of arguments that are invalid. If an argument doesn't seem to fit the pattern of any of these common forms, though, you may want to use a Venn diagram or a truth table instead.

Lewis Carroll, author of *Alice's Adventures in Wonderland*, was a math and logic teacher, and wrote two books on logic. In them, he would propose premises as a puzzle, to be connected using syllogisms. The following example is one such puzzle.

Example 42

Solve the puzzle. In other words, find a logical conclusion from these premises.

All babies are illogical.

Nobody is despised who can manage a crocodile.

Illogical persons are despised.

Let $b =$ is a baby, $d =$ is despised, $i =$ is illogical, and $m =$ can manage a crocodile.

Then we can write the premises as:

$b \to i$

$m \to \sim d$

$i \to d$

Writing the second premise correctly can be a challenge; it can be rephrased as "If you can manage a crocodile, then you are not despised."

Using the transitive property with the first and third premises, we can conclude that $b \to d$, that all babies are despised. Using the contrapositive of the second premise, $d \to \sim m$, we can then use the transitive property with $b \to d$ to conclude that $b \to \sim m$, that babies cannot manage crocodiles. While it is silly, this is a logical conclusion from the given premises.

Example 43

Premise:	If I work hard, I'll get a raise.
Premise:	If I get a raise, I'll buy a boat.
Conclusion:	If I don't buy a boat, I must not have worked hard.

If we let $h =$ working hard, $r =$ getting a raise, and $b =$ buying a boat, then we can represent our argument symbolically:

Premise:	$h \to r$
Premise:	$r \to b$
Conclusion:	$\sim b \to \sim h$

Using the transitive property with the two premises, we can conclude that $h \to b$, if I work hard, then I will buy a boat. When we learned about the contrapositive, we saw that the conditional statement $h \to b$ is equivalent to $\sim b \to \sim h$. Therefore, the conclusion is indeed a logical syllogism derived from the premises.

Try it Now 19

Is this argument valid?

Premise:	If I go to the party, I'll be really tired tomorrow.
Premise:	If I go to the party, I'll get to see friends.
Conclusion:	If I don't see friends, I won't be tired tomorrow.

Answer

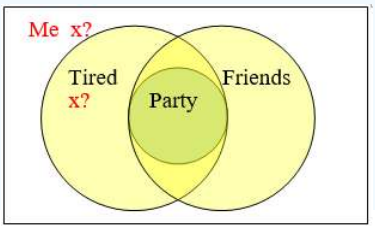
Let $p =$ go to party, $t =$ be tired, and $f =$ see friends.

Premise:	$p \to t$
Premise:	$p \to f$
Conclusion:	$-f \to \sim t$

We could try to rewrite the second premise using the contrapositive to state $\sim f \to \sim p$, but that does not allow us to form a syllogism. If I don't see friends, then I didn't go the party, but that is not sufficient to claim I won't be tired tomorrow.

Maybe I stayed up all night watching movies.

A Venn diagram can help, if we set it up correctly. The “party” circle must be completely contained within the intersection of the other circles. We know that I am somewhere outside the “friends” circle, but we cannot determine whether I am in the “tired” circle. All we really know for sure is that I didn’t go to the party.



This page titled 1.4.2: Forms of Valid Arguments is shared under a CC BY-SA 3.0 license and was authored, remixed, and/or curated by David Lippman (The OpenTextBookStore) via source content that was edited to the style and standards of the LibreTexts platform.

- **17.11: Forms of Valid Arguments** by David Lippman is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.

CHAPTER OVERVIEW

2: Functions

2.1: Functions

2.1.1: Function Notation

2.1.2: Domain and Range of a Function

2.1.3: Piecewise-defined Functions

2.1.4: Graphing Equations by Hand

2.2: Linear Functions

2.2.1: Slope of a Line

2.2.1E: Exercises

2.2.2: Find the Equation of a Line

2.3: Linear Models

2.3.1: Linear Models

2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know)

2.4: Graphing Linear Functions

2.4.1: Graphing Linear Equations

2.5: Linear Regression

2: Functions is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.1: Functions

2.1: Functions is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.1.1: Function Notation

The notation for a function is generally the $f(x)$ notation. In learning about graphing in algebra we typically use the x and y notation, that is: $y = 6x - 1$ In function notation, the dependent variable y is replaced by the notation $f(x)$:

$$f(x) = 6x - 1 \tag{2.1.1.1}$$

Function values for particular values of the independent variable x can be found by substituting the appropriate x value into the formula.

$$\begin{aligned} f(9) &= 6(9) - 1 = 53 \\ f(9) &= 53 \end{aligned} \tag{2.1.1.2}$$

Find each of the following values for the given functions:

$f(0) \quad f(-1) \quad f(3) \quad f\left(\frac{1}{2}\right) \quad f(x+2) \quad f(x+h)$

1) $f(x) = 2x^2 - 3x + 1$
2) $f(x) = 5x^2 + x - 7$
3) $f(x) = \frac{x}{x^2-1}$
4) $f(x) = \frac{x+3}{x^2+1}$

This page titled 2.1.1: Function Notation is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Richard W. Beveridge.

2.1.2: Domain and Range of a Function

The analysis of the behavior of functions addresses questions of when the function is increasing or decreasing, whether and where it has maximum or minimum values, where it crosses the x or y axis, and which values of x and y are to be included in the function.

The set of values available for the x, or independent variable is called the Domain of the function. The set of corresponding y values is called the Range of the function.

The linear function mentioned above $f(x) = 6x - 1$ has a domain of all real numbers and a range of all real numbers, $x \in \mathbb{R}$ and $y \in \mathbb{R}$. On the other hand, the function $f(x) = x^2$ has a domain of all real numbers, $x \in \mathbb{R}$, but its range is limited to the positive real numbers, $y \geq 0$

Considerations of the domain of a function typically refer to restrictions on which x values will generate real number values for y. The most common restrictions occur with the use of square root functions or rational functions.

The domain of the function $f(x) = \sqrt{x-7}$ would be the set of $x \geq 7$, so that no negative values are permitted under the square root. This secures the necessary real values for y. The range for this function is $y \geq 0$

The domain of the function $f(x) = \frac{x}{2x-3}$ would be $x \in \mathbb{R}$ (all real numbers), but $x \neq \frac{3}{2}$, thus avoiding a zero denominator which is an undefined value. The range for this function would be $y \in \mathbb{R}$ (all real numbers), but $y \neq \frac{1}{2}$, due to the horizontal asymptote at $y = \frac{1}{2}$

The questions of domain and range become more interesting when considered in relation to functions defined by graphs, or in applications. In an application involving perimeter in which the perimeter of a rectangle is given as 50 feet, we know that

$$2\ell + 2w = 50 \tag{2.1.2.1}$$

Rewriting this as a function of w, we can say that

$$\ell = f(w) = 25 - w \tag{2.1.2.2}$$

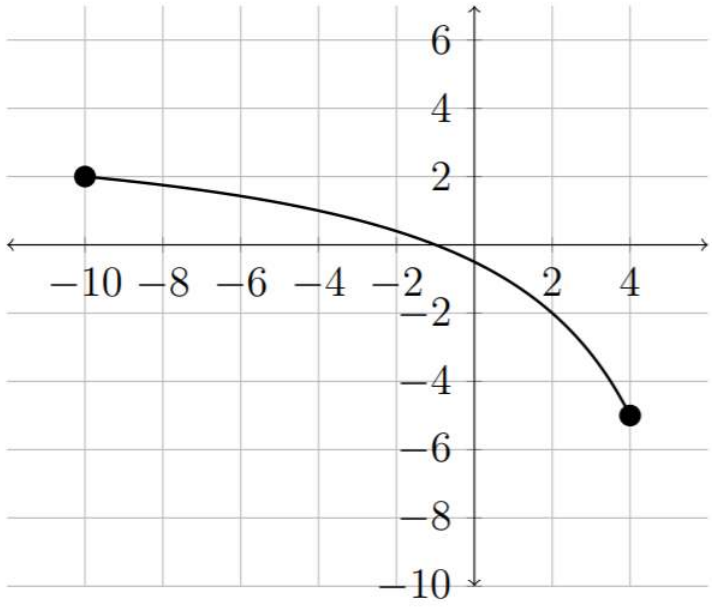
In this function relating the length and width based on a given perimeter, we can say the domain of the function is $0 < w < 25$. The width must be greater than 0 but less than 25, otherwise there would not be a rectangle. The same is true for the range or possible set of values for the length $0 < \ell < 25$ Exercises 4.2

Find the domain and range for each of the following functions:

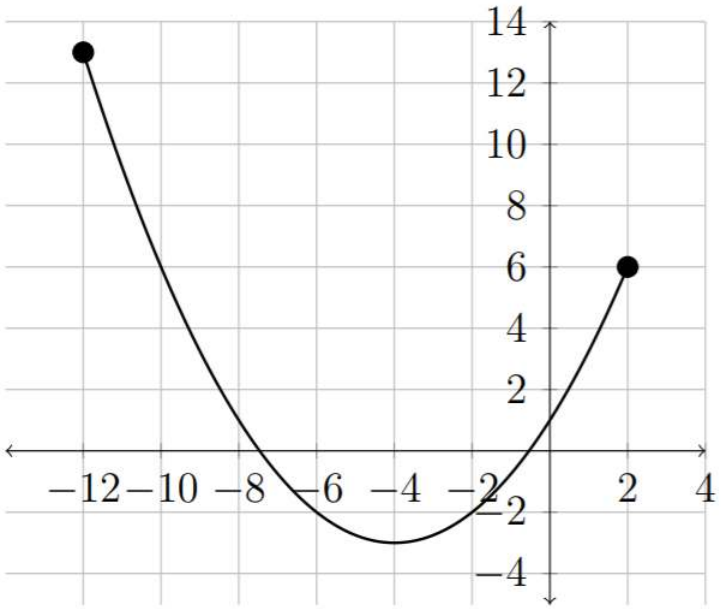
1) $f(x) = \sqrt{2x+1}$
2) $f(x) = \sqrt{3x-5}$
3) $\frac{x}{x+9}$
4) $f(x) = \frac{x+2}{2x-3}$

For the following graphs, assume that the endpoints for the domain and range are whole numbers.

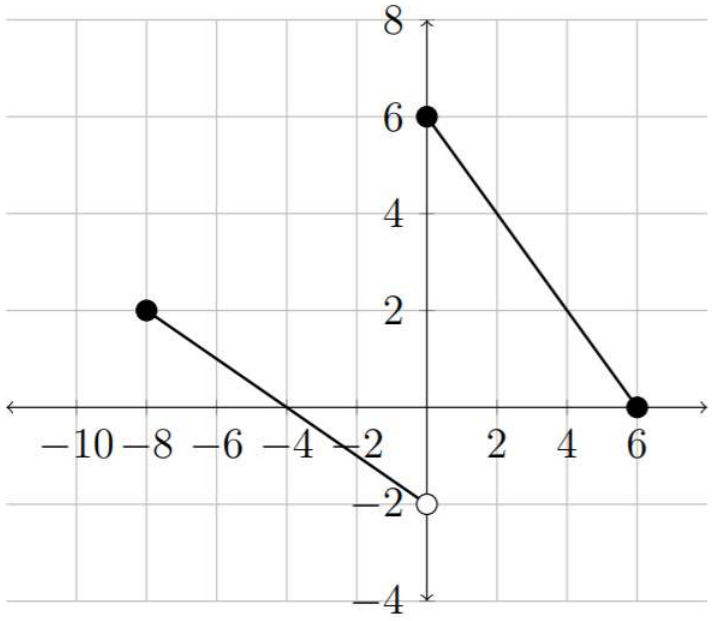
5)



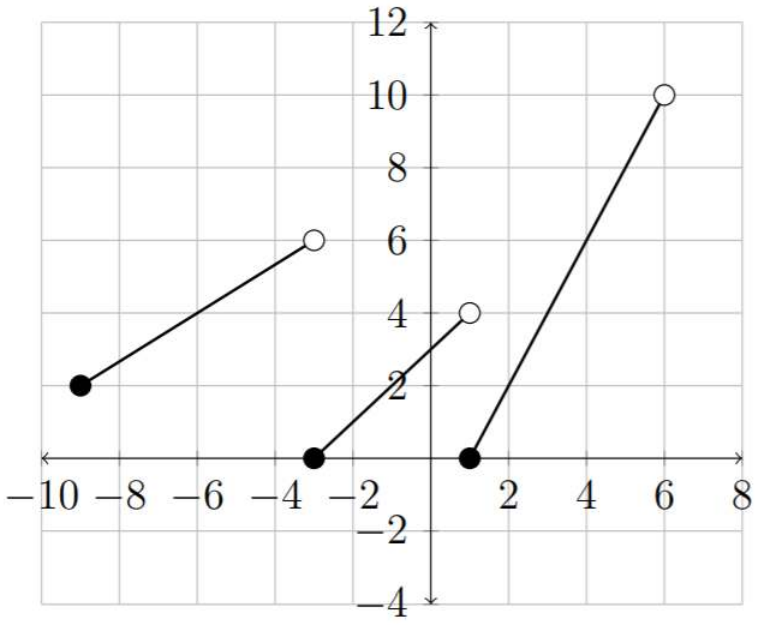
6)



7)



8)



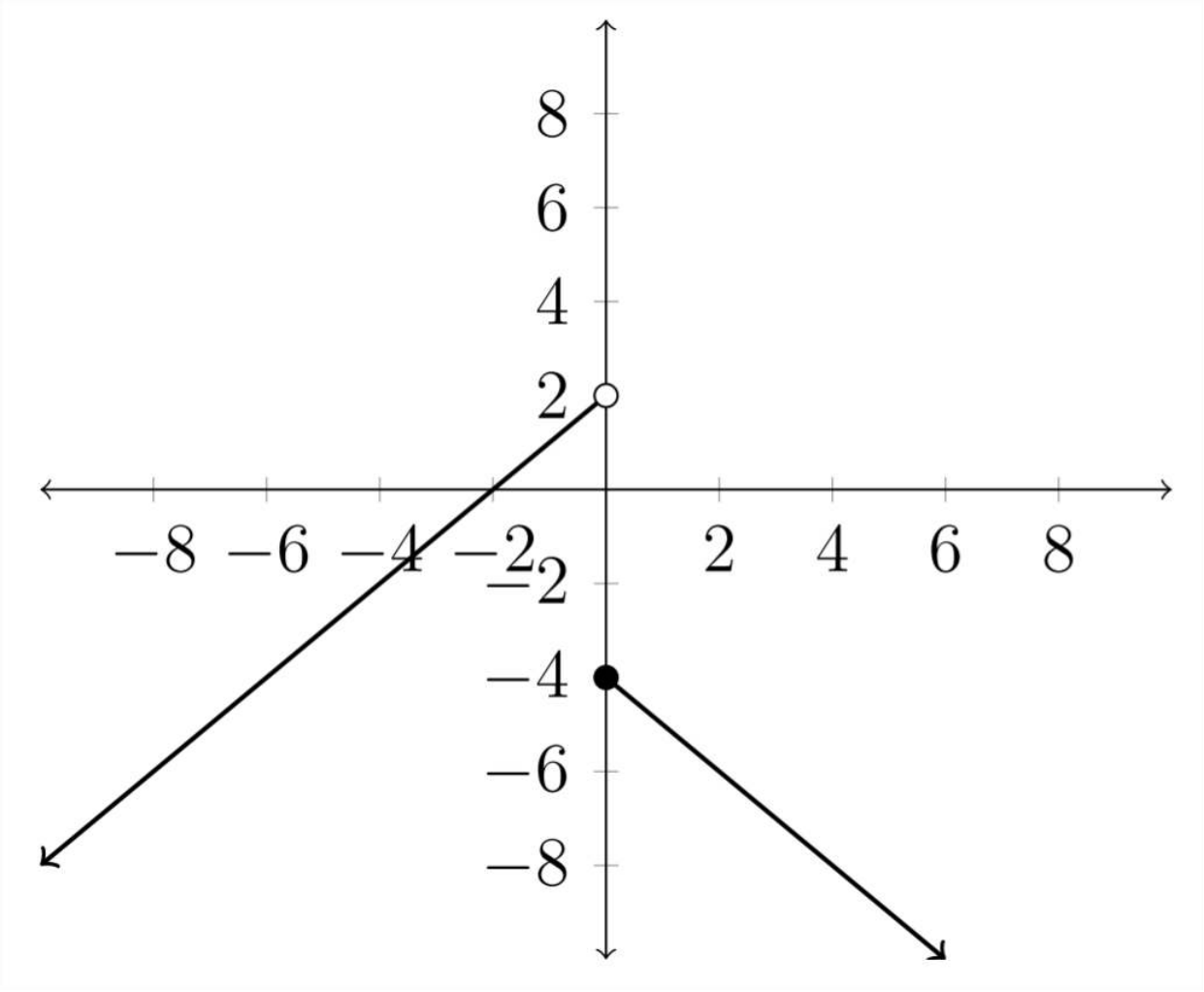
This page titled 2.1.2: Domain and Range of a Function is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Richard W. Beveridge.

2.1.3: Piecewise-defined Functions

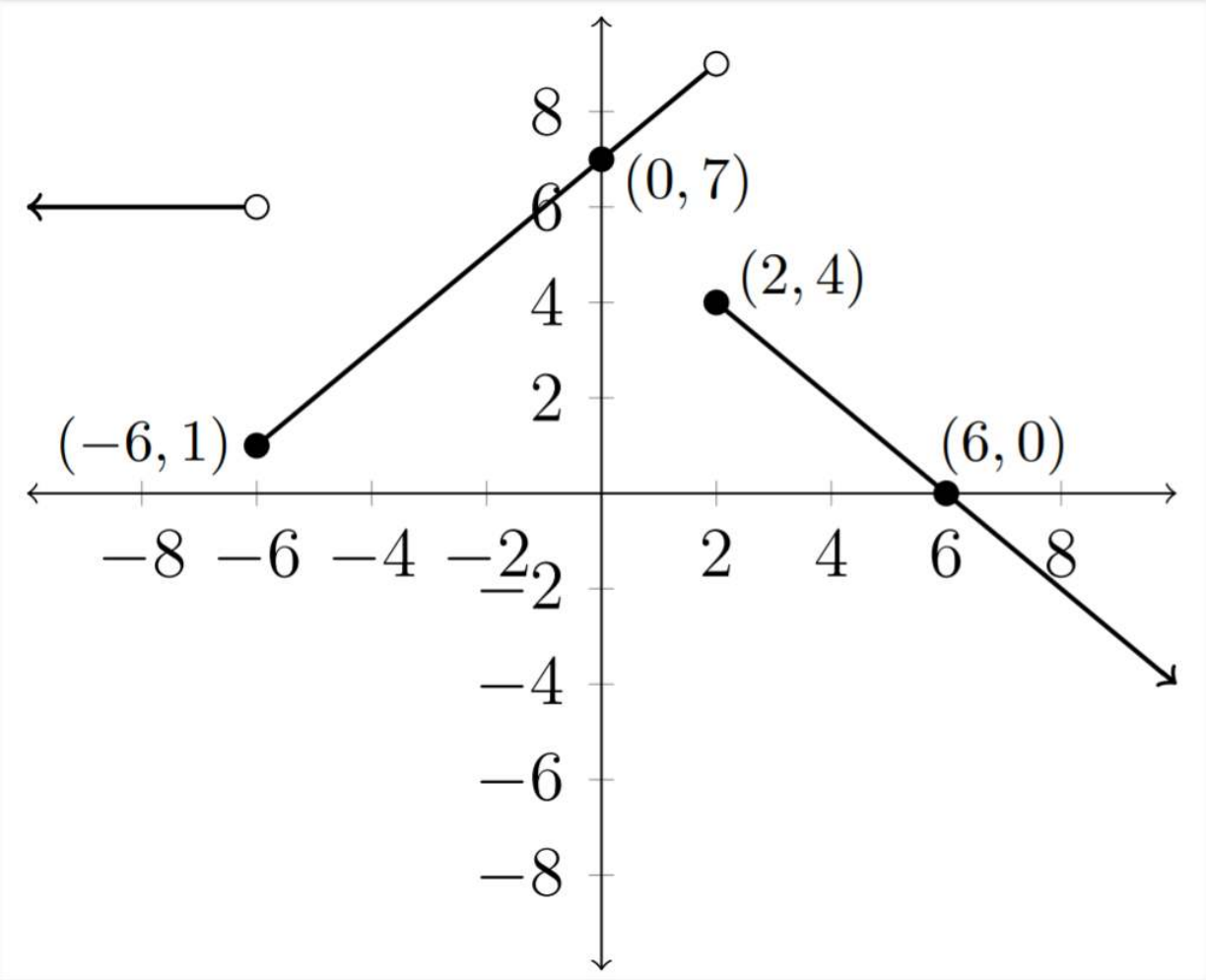
In certain situations a numerical relationship may follow one pattern of behavior for a while and then exhibit a different kind of behavior. In a situation such as this, it is helpful to use what is known as a piecewise defined function - a function that is defined in pieces.

$$f(x) = \begin{cases} x + 2 & x < 0 \\ -x - 4 & x \geq 0 \end{cases}$$

In the above example of a piecewise defined function, we see that the y values for the negative values of x are defined differently than the y values for the positive values of x



Sometimes we are given a graph and need to write a piecewise description of the function it describes.



The piecewise function pictured above could be described as follows:

$$f(x) = \begin{cases} 6 & x < -6 \\ x + 7 & -6 \leq x < 2 \\ -x + 6 & x \geq 2 \end{cases}$$

Exercises 4.6

Sketch a graph for each of the piecewise functions described below.

1) $$f(x) = \begin{cases} 1 & x < 2 \\ 2 & x \geq 2 \end{cases}$$

2) $$f(x) = \begin{cases} 0 & x \leq 1 \\ x - 1 & x > 1 \end{cases}$$

3) $$f(x) = \begin{cases} 4 & x < 2 \\ x - 2 & x \geq 2 \end{cases}$$

4) $$f(x) = \begin{cases} 2 - x & x < -2 \\ 4 & x \geq -2 \end{cases}$$

5) $$f(x) = \begin{cases} x & x \leq 1 \\ x+1 & x > 1 \end{cases}$$

6) $$f(x) = \begin{cases} 2x+3 & x < -2 \\ 3-x & x \geq -2 \end{cases}$$

7) $$f(x) = \begin{cases} -1 & x < -1 \\ 1 & -1 \leq x \leq 1 \\ -1 & x > 1 \end{cases}$$

8) $$f(x) = \begin{cases} -1 & x < -1 \\ x & -1 \leq x \leq 1 \\ 1 & x > 1 \end{cases}$$

9) $$f(x) = \begin{cases} x^2 & x < 0 \\ x + 1 & x \geq 0 \end{cases}$$

10) $$f(x) = \begin{cases} x^2 + 2x & x \leq -1 \\ x & x > -1 \end{cases}$$

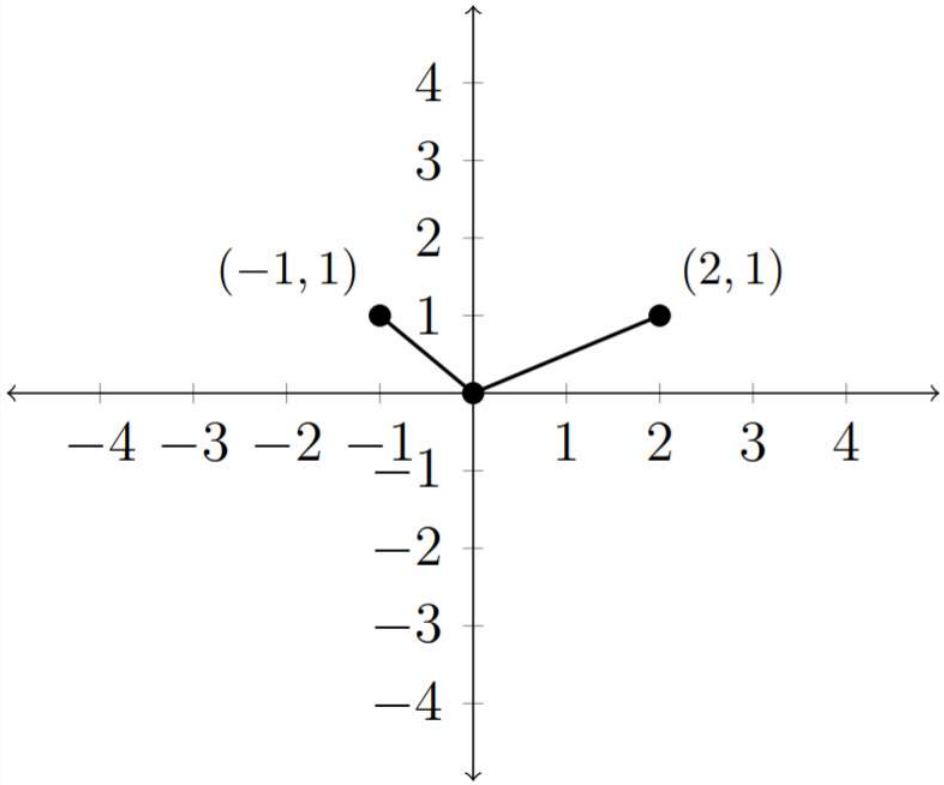
11) $$f(x) = \begin{cases} 2 & x \leq -1 \\ x^2 & x > -1 \end{cases}$$

12) $$f(x) = \begin{cases} 1 - x^2 & x \leq 2 \\ x & x > 2 \end{cases}$$

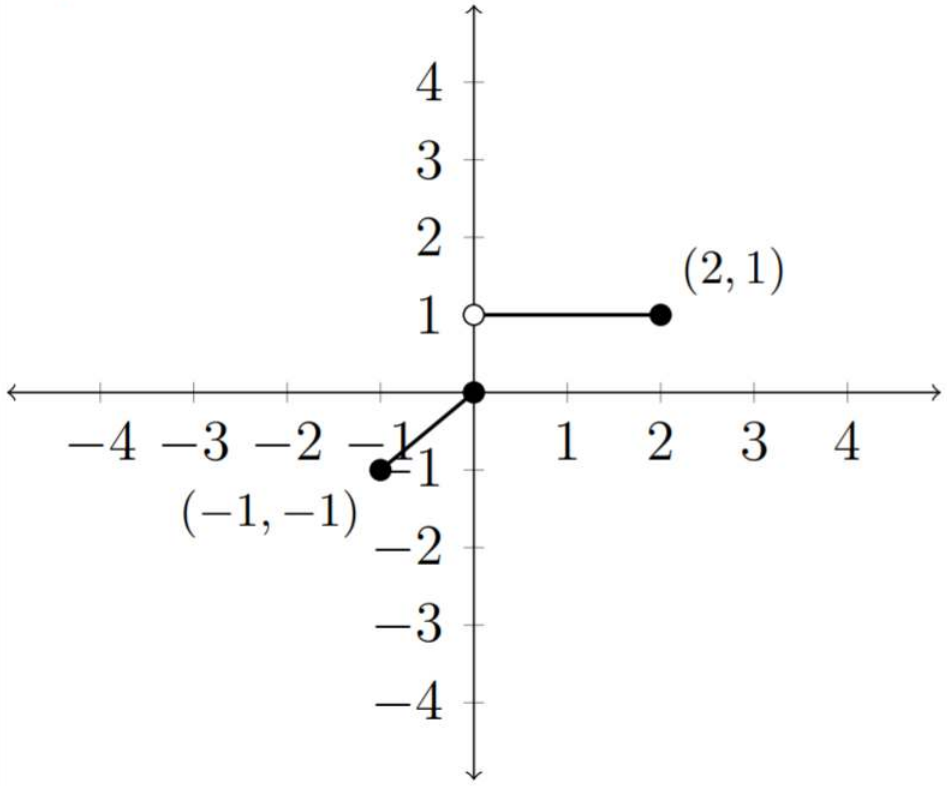
13) $$f(x) = \begin{cases} 4 & x < -2 \\ x^2 & -2 \leq x \leq 2 \\ -x + 6 & x > 2 \end{cases}$$

14) $$f(x) = \begin{cases} 3x & x < 0 \\ x + 1 & 0 \leq x \leq 2 \\ (x-2)^2 & x > 2 \end{cases}$$

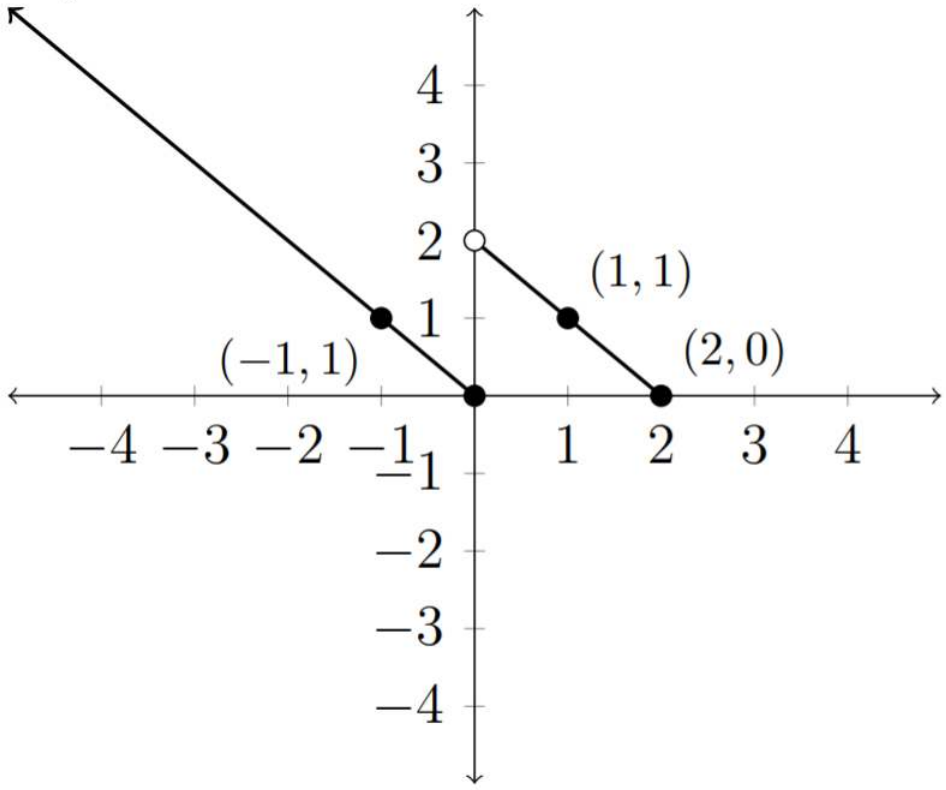
15)



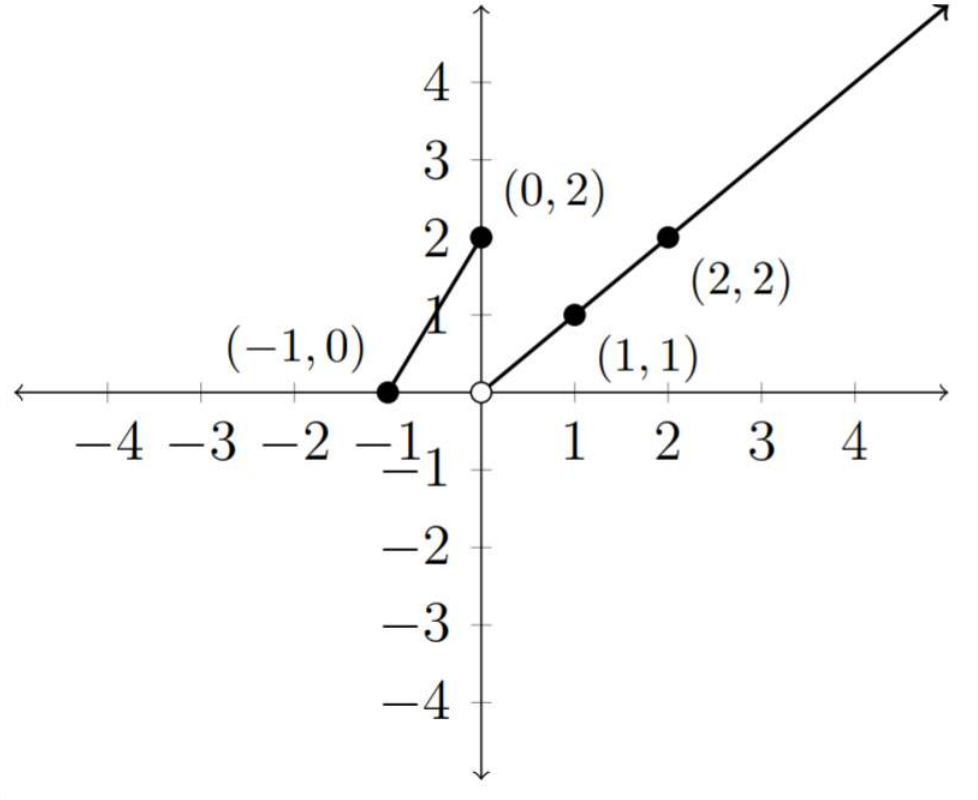
16)



17)



18)



This page titled 2.1.3: Piecewise-defined Functions is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Richard W. Beveridge.

2.1.4: Graphing Equations by Hand

We begin with the definition of an *ordered pair*.

Ordered Pair

The construct (x, y), where x and y are any real numbers, is called an ordered pair of real numbers.

$(4, 3)$, $(-3, 4)$, $(-2, -3)$, and $(3, -1)$ are examples of ordered pairs.

Order Matters

Pay particular attention to the phrase "ordered pairs." Order matters. Consequently, the ordered pair (x, y) is not the same as the ordered pair (y, x), because the numbers are presented in a different order.

The Cartesian Coordinate System

Pictured in Figure 2.1.4.1is a Cartesian Coordinate System. On a grid, we've created two real lines, one horizontal labeled x (we'll refer to this one as the x-axis), and the other vertical labeled y (we'll refer to this one as the y-axis).

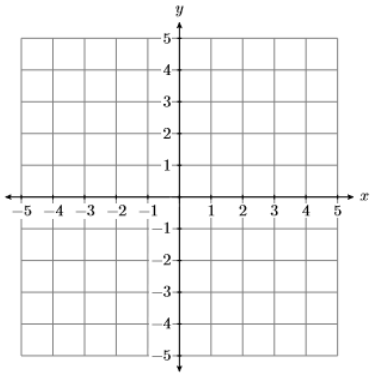


Figure 2.1.4.1: The Cartesian Coordinate System.

Two Important Points:

Here are two important points to be made about the horizontal and vertical axes in Figure 2.1.4.1.

1. As you move from left to right along the horizontal axis (the x-axis in Figure 2.1.4.1), the numbers grow larger. The positive direction is to the right, the negative direction is to the left.
2. As you move from bottom to top along the vertical axis (the y-axis in Figure 2.1.4.1), the numbers grow larger. The positive direction is upward, the negative direction is downward.

Additional Comments:

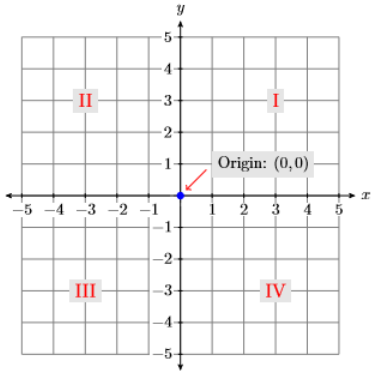


Figure 2.1.4.2: Numbering the quadrants and indicating the coordinates of the origin.

Two additional comments are in order:

3. The point where the horizontal and vertical axes intersect in Figure 2.1.4.2is called the origin of the coordinate system. The origin has coordinates $(0, 0)$.
4. The horizontal and vertical axes divide the plane into four quadrants, numbered $\mathrm{I}, \mathrm{II}, \mathrm{MI}$, and IV (roman numerals for one, two, three, and four), as shown in Figure 2.1.4.2 Note that the quadrants are numbered in a counter-clockwise order.

Note

Rene Descartes (1596-1650) was a French philosopher and mathematician who is well known for the famous phrase"cogito ergo sum" (I think, therefore I am), which appears in his Discours de la methode pour bien conduire sa raison, et chercher la verite dans les sciences (Discourse on the Method of Rightly Conducting the Reason, and Seeking Truth in the Sciences). In that same treatise, Descartes introduces his coordinate system, a method for representing points in the plane via pairs of real numbers. Indeed, the Cartesian plane of modern day is so named in honor of Rene Descartes, who some call the "Father of Modern Mathematics"

Plotting Ordered Pairs

Before we can plot any points or draw any graphs, we first need to set up a Cartesian Coordinate System on a sheet of graph paper? How do we do this? What is required?

how to setup cartesian coordinate system

Draw and label each axis.

If we are going to plot points (x,y), then, on a sheet of graph paper, perform each of the following initial tasks.

1. Use a ruler to draw the horizontal and vertical axes.
2. Label the horizontal axis as the x-axis and the vertical axis as the y-axis.

We don't always label the horizontal axis as the x-axis and the vertical axis as the y-axis. For example, if we want to plot the velocity of an object as a function of time, then we would be plotting points (t, v). In that case, we would label the horizontal axis as the t-axis and the vertical axis as the v-axis.

Indicate the scale on each axis.

3. Label at least one vertical gridline with its numerical value.
4. Label at least one horizontal gridline with its numerical value.

The scales on the horizontal and vertical axes may differ. However, on each axis, the scale must remain consistent. That is, as you count to the right from the origin on the x-axis, if each gridline represents one unit, then as you count to the left from the origin on the x-axis, each gridline must also represent one unit. Similar comments are in order for the y-axis, where the scale must also be consistent, whether you are counting up or down.

The result of this first step is shown in Figure 2.1.4.3

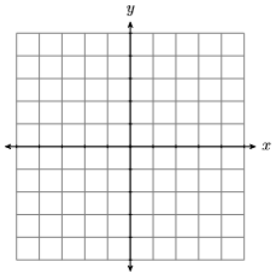


Figure 2.1.4.3: Draw and label each axis.

An example is shown in Figure 2.1.4.4 Note that the scale indicated on the x-axis indicates that each gridline counts as 1-unit as we count from left-to-right. The scale on the y-axis indicates that each gridlines counts as 2-units as we count from bottom-to-top.

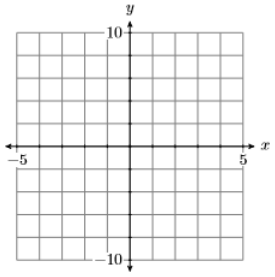


Figure 2.1.4.4: Indicate the scale on each axis.

Now that we know how to set up a Cartesian Coordinate System on a sheet of graph paper, here are two examples of how we plot points on our coordinate system.

To plot the ordered pair $(4, 3)$, start at the origin and move 4 units to the right along the horizontal axis, then 3 units upward in the direction of the vertical axis.

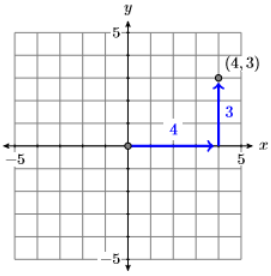


Figure 2.1.4.5: Plot of $(4, 3)$.

To plot the ordered pair $(-2, -3)$, start at the origin and move 2 units to the left along the horizontal axis, then 3 units downward in the direction of the vertical axis.

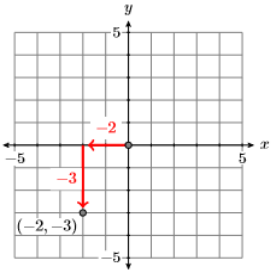


Figure 2.1.4.6: Plot of $(-2, -3)$

Continuing in this manner, each ordered pair (x, y) of real numbers is associated with a unique point in the Cartesian plane. Vice-versa, each point in the Cartesian point is associated with a unique ordered pair of real numbers. Because of this association, we begin to use the words "point" and "ordered pair" as equivalent expressions, sometimes referring to the "point" (x, y) and other times to the "ordered pair" (x, y).

Example 2.1.4.1

Identify the coordinates of the point P in Figure 2.1.4.7.

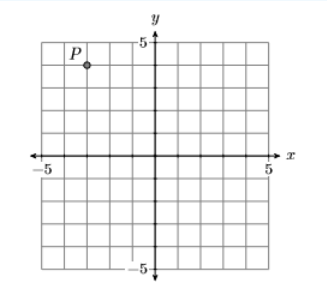


Figure 2.1.4.7: Identify the coordinates of the point P.

Solution

In Figure 2.1.4.8, start at the origin, move 3 units to the left and 4 units up to reach the point P. This indicates that the coordinates of the point P are $(-3, 4)$.

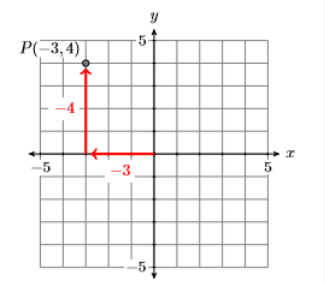
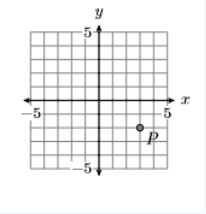


Figure 2.1.4.8: Start at the origin, move 3 units left and 4 units up.

Exercise 2.1.4.1

Identify the coordinates of the point P in the graph below.



Answer

$(3, -2)$

Equations in Two Variables

Note

The variables do not have to always be x and y. For example, the equation $v = 2 + 3.2t$ is an equation in two variables, v and t.

The equation $y = x + 1$ is an equation in two variables, in this case x and y. Consider the point $(x, y) = (2, 3)$. If we substitute 2 for x and 3 for y in the equation $y = x + 1$, we get the following result:

$$\begin{aligned} y &= x + 1 \quad \text{Original equation.} \\ 3 &= 2 + 1 \quad \text{Substitute: 2 for } x, 3 \text{ for } y \\ 3 &= 3 \quad \text{Simplify both sides.} \end{aligned}$$

Because the last line is a true statement, we say that $(2, 3)$ is a solution of the equation $y = x + 1$. Alternately, we say that $(2, 3)$ satisfies the equation $y = x + 1$. On the other hand, consider the point $(x, y) = (-3, 1)$. If we substitute -3 for x and 1 for y in the equation $y = x + 1$, we get the following result.

$$\begin{aligned} y &= x + 1 \quad \text{Original equation.} \\ 1 &= -3 + 1 \quad \text{Substitute: } -3 \text{ for } x, 1 \text{ for } y \\ 1 &= -2 \quad \text{Simplify both sides.} \end{aligned}$$

Because the last line is a false statement, the point $(-3, 1)$ is **not** a solution of the equation $y = x + 1$; that is, the point $(-3, 1)$ does **not** satisfy the equation $y = x + 1$.

Solutions of an equation in two variables

Given an equation in the variables x and y and a point $(x, y) = (a, b)$, if upon subsituting a for x and b for y a true statement results, then the point $(x, y) = (a, b)$ is said to be a solution of the given equation. Alternately, we say that the point $(x, y) = (a, b)$ satisfies the given equation.

Example 2.1.4.2

Which of the ordered pairs $(0, -3)$ and $(1, 1)$ satisfy the equation $y = 3x - 2$?

Solution

Substituting the ordered pairs $(0, -3)$ and $(1, 1)$ into the equation $y = 3x - 2$ lead to the following results:

Consider $(x, y) = (0, -3)$. Substitute 0 for x and -3 for y:

$$\begin{aligned} y &= 3x - 2 \\ -3 &= 3(0) - 2 \\ -3 &= -2 \end{aligned}$$

The resulting statement is false.

Consider $(x, y) = (1, 1)$. Substitute 1 for x and 1 for y:

$$\begin{aligned} y &= 3x - 2 \\ 1 &= 3(1) - 2 \\ 1 &= 1 \end{aligned}$$

The resulting statement is true.

Thus, the ordered pair $(0, -3)$ does not satisfy the equation $y = 3x - 2$, but the ordered pair $(1, 1)$ does satisfy the equation $y = 3x - 2$.

Exercise 2.1.4.2

Which of the ordered pairs $(-1, 3)$ and $(2, 1)$ satisfy the equation $y = 2x + 5$?

Answer

$(-1, 3)$

Graphing Equations in Two Variables

Let's first define what is meant by the graph of an equation in two variables.

The graph of an equation

The graph of an equation is the set of all points that satisfy the given equation.

Example 2.1.4.3

Sketch the graph of the equation $y = x + 1$.

Solution

The definition requires that we plot all points in the Cartesian Coordinate System that satisfy the equation $y = x + 1$. Let's first create a table of points that satisfy the equation. Start by creating three columns with headers x, y, and (x, y), then select some values for x and put them in the first column.

Take the first value of x, namely $x = -3$, and substitute it into the equation $y = x + 1$.

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

x	$y = x + 1$	(x, y)
-3	-2	$(-3, -2)$
-2		
-1		
0		
1		
2		
3		

Thus, when $x = -3$, we have $y = -2$. Enter this value into the table.

Continue substituting each tabular value of x into the equation $y = x + 1$ and use each result to complete the corresponding entries in the table.

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

x	$y = x + 1$	(x, y)
-3	-2	$(-3, -2)$
-2	-1	$(-2, -1)$
-1	0	$(-1, 0)$
0	1	$(0, 1)$
1	2	$(1, 2)$
2	3	$(2, 3)$
3	4	$(3, 4)$

The last column of the table now contains seven points that satisfy the equation $y = x + 1$. Plot these points on a Cartesian Coordinate System (see Figure \(\PageIndex{9}\)).

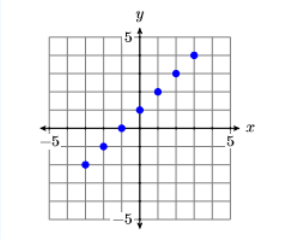


Figure 2.1.4.9: Seven points that satisfy the equation $y = x + 1$.

In Figure 2.1.4.9 we have plotted seven points that satisfy the given equation $y = x + 1$. However, the definition requires that we plot all points that satisfy the equation. It appears that a pattern is developing in Figure 2.1.4.9 but let's calculate and plot a few more points in order to be sure. Add the x-values -2.5, -1.5, -0.5, 0.5, 1.5, and 2.5 to the x-column of the table, then use the equation $y = x + 1$ to evaluate y at each one of these x-values.

$$y = -2.5 + 1 = -1.5$$
$$y = -1.5 + 1 = -0.5$$
$$y = -0.5 + 1 = 0.5$$
$$y = 0.5 + 1 = 1.5$$
$$y = 1.5 + 1 = 2.5$$
$$y = 2.5 + 1 = 3.5$$

x	$y = x + 1$	(x, y)
-2.5	-1.5	$(-2.5, -1.5)$
-1.5	-0.5	$(-1.5, -0.5)$
-0.5	0.5	$(-0.5, 0.5)$
0.5	1.5	$(0.5, 1.5)$
1.5	2.5	$(1.5, 2.5)$
2.5	3.5	$(2.5, 3.5)$

Add these additional points to the graph in Figure 2.1.4.9to produce the image shown in Figure 2.1.4.10

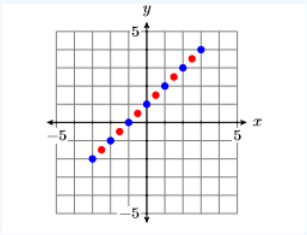


Figure 2.1.4.10: Adding additional points to the graph of $y = x + 1$.

There are an infinite number of points that satisfy the equation $y = x + 1$. In Figure 2.1.4.10 we've plotted only 13 points that satisfy the equation. However, the collection of points plotted in Figure 2.1.4.10suggest that if we were to plot the remainder of the points that satisfy the equation $y = x + 1$, we would get the graph of the line shown in Figure 2.1.4.11

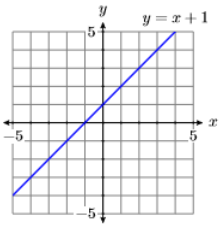
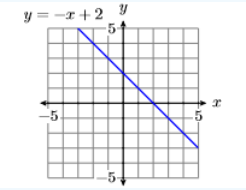


Figure 2.1.4.11: The graph of $y = x + 1$ is a line.

Exercise 2.1.4.3

Sketch the graph of the equation $y = -x + 2$.

Answer



Guidelines and Requirements

Example 2.1.4.3 suggests that we should use the following guidelines when sketching the graph of an equation.

Guidelines for drawing the graph of an equation

When asked to draw the graph of an equation, perform each of the following steps:

1. Set up and calculate a table of points that satisfy the given equation.
2. Set up a Cartesian Coordinate System on graph paper and plot the points in your table on the system. Label each axis (usually x and y) and indicate the scale on each axis.
3. If the number of points plotted are enough to envision what the shape of the final curve will be, then draw the remaining points that satisfy the equation as imagined. Use a ruler if you believe the graph is a line. If the graph appears to be a curve, freehand the graph without the use of a ruler.
4. If the number of plotted points do not provide enough evidence to envision the final shape of the graph, add more points to your table, plot them, and try again to envision the final shape of the graph. If you still cannot predict the eventual shape of the graph, keep adding points to your table and plotting them until you are convinced of the final shape of the graph.

Here are some additional requirements that must be followed when sketching the graph of an equation.

Graph paper, lines, curves, and rulers.

The following are requirements for this class:

5. All graphs are to be drawn on graph paper.
6. All lines are to be drawn with a ruler. This includes the horizontal and vertical axes.
7. If the graph of an equation is a curve instead of a line, then the graph should be drawn freehand, without the aid of a ruler.

Using the TABLE Feature of the Graphing Calculator

As the equations become more complicated, it can become quite tedious to create tables of points that satisfy the equation. Fortunately, the graphing calculator has a TABLE feature that enables you to easily construct tables of points that satisfy the given equation.

Example 2.1.4.4

Use the graphing calculator to help create a table of points that satisfy the equation $y = x^2 - 7$. Plot the points in your table. If you don't feel that there is enough evidence to envision what the final shape of the graph will be, use the calculator to add more points to your table and plot them. Continue this process until your are convinced of the final shape of the graph.

Solution



Figure 2.1.4.12: The graphing calculator.

The first step is to load the equation $y = x^2 - 7$ into the **Y=** menu of the graphing calculator. The topmost row of buttons on your calculator (see Figure 2.1.4.12) have the following appearance:

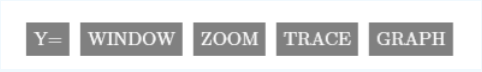
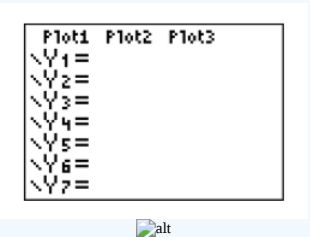


Figure 2.1.4.13. Use the up-and-down arrow keys (see Figure 2.1.4.15) to move the cursor after **Y1**= in the **Y**= menu, then use the following keystrokes to enter the equation $y = x^2 - 7$. The result is shown in Figure 2.1.4.14.





alt

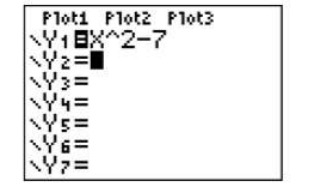


Figure 2.1.4.13: Opening the **Y**= menu. ***Figure* 2.1.4.14:** Entering the equation $y = x^2 - 7$.



Figure 2.1.4.15: Arrow keys move the cursor.

The next step is to “set up” the table. First, note that the calculator has symbolism printed on its case above each of its buttons. Above the WINDOW button you’ll note the phrase TBLSET. Note that it is in the same color as the 2ND button. Thus, to open the setup window for the table, enter the following keystrokes.

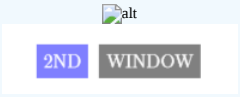


Figure 2.1.4.16) equal to the first x-value you want to see in your table. In this case, enter -4 after TblStart. SetTbl to the increment you want for your x-values. In this case, set Tbl equal to 1. Finally, set both the independent and dependent variables to “automatic.” In each case, use the arrow keys to highlight the word AUTO and press ENTER. The result is shown in Figure 2.1.4.16.

Next, note the word TABLE above the GRAPH button is in the same color as the 2ND key. To open the TABLE, enter the following keystrokes.



Figure 2.1.4.17. Note that you can use the up-and-down arrow keys to scroll through the table.



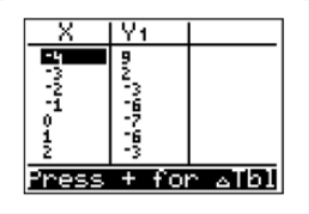


Figure 2.1.4.16: Opening the TBLSET menu. **Figure 2.1.4.17:** Table of points satisfying $y = x^2 - 7$.

Next, enter the results from your calculator’s table into a table on a sheet of graph paper, then plot the points in the table. The results are shown in Figure 2.1.4.18

x	$y = x^2 - 7$	(x, y)
-4	9	$(-4, 9)$
-3	2	$(-3, 2)$
-2	-3	$(-2, -3)$
-1	-6	$(-1, -6)$
0	-7	$(0, -7)$
1	-6	$(1, -6)$
2	-3	$(2, -3)$
3	2	$(3, 2)$
4	9	$(4, 9)$

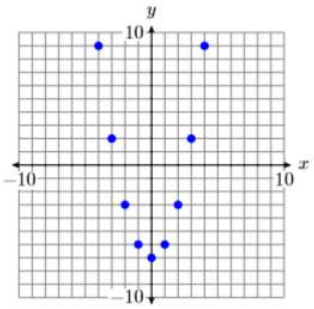


Figure 2.1.4.18: Plotting nine points that satisfy the equation $y = x^2 - 7$.

In Figure 2.1.4.18, the eventual shape of the graph of $y = x^2 - 7$ may be evident already, but let’s add a few more points to our table and plot them. Open the table “setup” window again by pressing 2ND WINDOW. Set TblStart to -4 again, then set the increment Tbl to 0.5. The result is shown in Figure 2.1.4.19



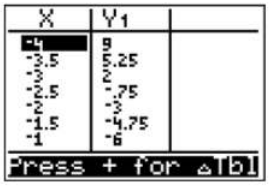


Figure 2.1.4.19: SetTbl equal to 0.5. **Figure 2.1.4.20:** More points satisfying $y = x^2 - 7$.

Add these new points to the table on your graph paper and plot them (see Figure 2.1.4.21).

x	$y = x^2 - 7$	(x, y)
−3.5	5.25	(−3.5, 5.25)
−2.5	−0.75	(−2.5, −0.75)
−1.5	−4.75	(−1.5, −4.75)
−0.5	−6.75	(−0.5, −6.75)
0.5	−6.75	(0.5, −6.75)
1.5	−4.75	(1.5, −4.75)
2.5	−0.75	(2.5, −0.75)
3.5	5.25	(3.5, 5.25)

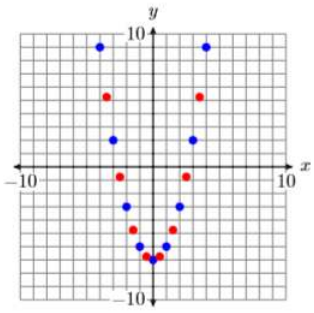


Figure 2.1.4.21: Adding additional points to the graph of $y = x^2 - 7$.

There are an infinite number of points that satisfy the equation $y = x^2 - 7$. In Figure 2.1.4.21, we've plotted only 17 points that satisfy the equation $y = x^2 - 7$. However, the collection of points in Figure 2.1.4.21 suggest that if we were to plot the remainder of the points that satisfy the equation $y = x^2 - 7$, the result would be the curve (called a parabola) shown in Figure 2.1.4.22

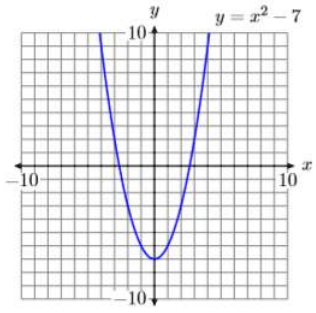


Figure 2.1.4.22: The graph of $y = x^2 - 7$ is a curve called a parabola.

This page titled 2.1.4: Graphing Equations by Hand is shared under a CC BY-NC-ND 3.0 license and was authored, remixed, and/or curated by David Arnold via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: Graphing Equations by Hand** by David Arnold is licensed CC BY-NC-ND 3.0. Original source: https://math.libretexts.org/api/deki/files/80188/ElementaryAlgebra.pdf.

2.2: Linear Functions

2.2: Linear Functions is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.2.1: Slope of a Line

Learning Objectives

By the end of this section, you will be able to:

- Find the slope of a line
- Graph a line given a point and the slope
- Graph a line using its slope and intercept
- Choose the most convenient method to graph a line
- Graph and interpret applications of slope–intercept
- Use slopes to identify parallel and perpendicular lines

Before you get started, take this readiness quiz.

1. Simplify: $\frac{(1-4)}{(8-2)}$.
   If you missed this problem, review [link].
2. Divide: $\frac{0}{4}$, $\frac{4}{0}$.
   If you missed this problem, review [link].
3. Simplify: $\frac{15}{-3}$, $\frac{-15}{3}$, $\frac{-15}{-3}$.
   If you missed this problem, review [link].

Find the Slope of a Line

When you graph linear equations, you may notice that some lines tilt up as they go from left to right and some lines tilt down. Some lines are very steep and some lines are flatter.

In mathematics, the measure of the steepness of a line is called the *slope* of the line.

The concept of slope has many applications in the real world. In construction the pitch of a roof, the slant of the plumbing pipes, and the steepness of the stairs are all applications of slope. and as you ski or jog down a hill, you definitely experience slope.

We can assign a numerical value to the slope of a line by finding the ratio of the rise and run. The *rise* is the amount the vertical distance changes while the *run* measures the horizontal change, as shown in this illustration. Slope is a rate of change. See Figure.

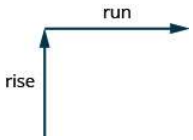


Figure 2.2.1.1.

SLOPE OF A LINE

The slope of a line is $m = \frac{\text{rise}}{\text{run}}$.

The rise measures the vertical change and the run measures the horizontal change.

To find the slope of a line, we locate two points on the line whose coordinates are integers. Then we sketch a right triangle where the two points are vertices and one side is horizontal and one side is vertical.

To find the slope of the line, we measure the distance along the vertical and horizontal sides of the triangle. The vertical distance is called the *rise* and the horizontal distance is called the *run*,

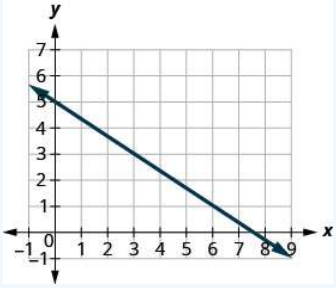
FIND THE SLOPE OF A LINE FROM ITS GRAPH USING $m = \frac{\text{rise}}{\text{run}}$

1. Locate two points on the line whose coordinates are integers.
2. Starting with one point, sketch a right triangle, going from the first point to the second point.
3. Count the rise and the run on the legs of the triangle.

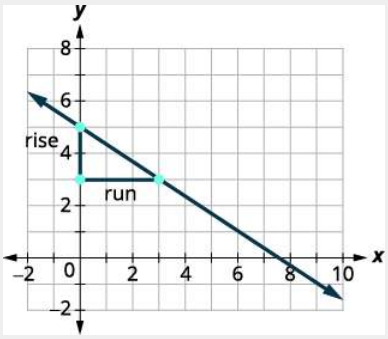
4. Take the ratio of rise to run to find the slope: $m = \frac{\text{rise}}{\text{run}}$.

Example 2.2.1.1

Find the slope of the line shown.

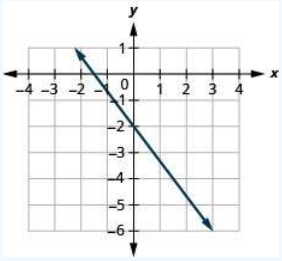


Solution

Locate two points on the graph whose coordinates are integers.	$(0, 5)$ and $(3, 3)$
Starting at $(0, 5)$, sketch a right triangle to $(3, 3)$ as shown in this graph.	
Count the rise— since it goes down, it is negative.	The rise is -2.
Count the run.	The run is 3.
Use the slope formula.	$m = \frac{\text{rise}}{\text{run}}$
Substitute the values of the rise and run.	$m = \frac{-2}{3}$
Simplify.	$m = -\frac{2}{3}$
	The slope of the line is $-\frac{2}{3}$.
	So y decreases by 2 units as x increases by 3 units.

Try It 2.2.1.2

Find the slope of the line shown.

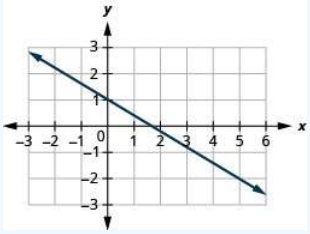


Answer

$-\frac{4}{3}$

? Try It 2.2.1.3

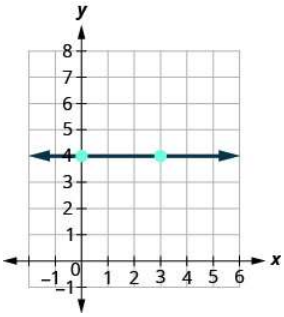
Find the slope of the line shown.



Answer

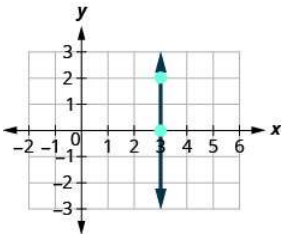
$-\frac{3}{5}$

How do we find the slope of horizontal and vertical lines? To find the slope of the horizontal line, $y = 4$, we could graph the line, find two points on it, and count the rise and the run. Let's see what happens when we do this, as shown in the graph below.



What is the rise?	The rise is 0.
What is the run?	The run is 3.
What is the slope?	$m = \frac{\text{rise}}{\text{run}}$
	$m = \frac{0}{3}$
	$m = 0$
	The slope of the horizontal line $y = 4$ is 0.

Let's also consider a vertical line, the line $x = 3$, as shown in the graph.



What is the rise?	The rise is 2.
What is the run?	The run is 0.
What is the slope?	$m = \frac{\text{rise}}{\text{run}}$
	$m = \frac{2}{0}$

The slope is undefined since division by zero is undefined. So we say that the slope of the vertical line $x = 3$ is undefined.

All horizontal lines have slope 0. When the *y*-coordinates are the same, the rise is 0.

The slope of any vertical line is undefined. When the *x*-coordinates of a line are all the same, the run is 0.

SLOPE OF A HORIZONTAL AND VERTICAL LINE

The slope of a horizontal line, $y = b$, is 0.

The slope of a vertical line, $x = a$, is undefined.

Example 2.2.1.4

Find the slope of each line: a. $x = 8$ b. $y = -5$.

Solution

a. $x = 8$ This is a vertical line. Its slope is undefined.
b. $y = -5$ This is a horizontal line. It has slope 0.

Try It 2.2.1.5

Find the slope of the line: $x = -4$.

Answer

undefined

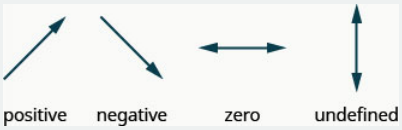
Try It 2.2.1.6

Find the slope of the line: $y = 7$.

Answer

0

QUICK GUIDE TO THE SLOPES OF LINES



Sometimes we'll need to find the slope of a line between two points when we don't have a graph to count out the rise and the run. We could plot the points on grid paper, then count out the rise and the run, but as we'll see, there is a way to find the slope without graphing. Before we get to it, we need to introduce some algebraic notation.

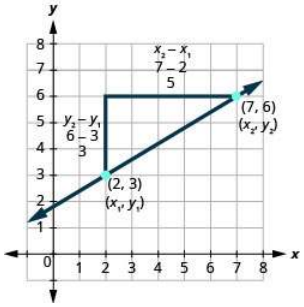
We have seen that an ordered pair (x, y) gives the coordinates of a point. But when we work with slopes, we use two points. How can the same symbol (x, y) be used to represent two different points? Mathematicians use subscripts to distinguish the points.

(x_1, y_1) read "x sub 1, y sub 1"
(x_2, y_2) read "x sub 2, y sub 2"

We will use (x_1, y_1) to identify the first point and (x_2, y_2) to identify the second point.

If we had more than two points, we could use (x_3, y_3), (x_4, y_4), and so on.

Let's see how the rise and run relate to the coordinates of the two points by taking another look at the slope of the line between the points $(2, 3)$ and $(7, 6)$, as shown in this graph.



Since we have two points, we will use subscript notation. $\begin{pmatrix} x_1, & y_1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x_2, & y_2 \\ 7 & 6 \end{pmatrix}$

$m = \frac{\text{rise}}{\text{run}}$

On the graph, we counted the rise of 3 and the run of 5. $m = \frac{3}{5}$

Notice that the rise of 3 can be found by subtracting the y-coordinates, 6 and 3, and the run of 5 can be found by subtracting the x-coordinates 7 and 2.

We rewrite the rise and run by putting in the coordinates. $m = \frac{6-3}{7-2}$

But 6 is y_2, the y-coordinate of the second point and 3 is y_1, the y-coordinate of the first point. So we can rewrite the slope using subscript notation. $m = \frac{y_2-y_1}{7-2}$

Also 7 is the x-coordinate of the second point and 2 is the x-coordinate of the first point. So again we rewrite the slope using subscript notation. $m = \frac{y_2-y_1}{x_2-x_1}$

We've shown that $m = \frac{y_2-y_1}{x_2-x_1}$ is really another version of $m = \frac{\text{rise}}{\text{run}}$. We can use this formula to find the slope of a line when we have two points on the line.

SLOPE OF A LINE BETWEEN TWO POINTS

The slope of the line between two points (x_1, y_1) and (x_2, y_2) is: $m = \frac{y_2-y_1}{x_2-x_1}$.

The slope is:

y of the second point minus y of the first point

over

x of the second point minus x of the first point

Example 2.2.1.7

Use the slope formula to find the slope of the line through the points $(-2, -3)$ and $(-7, 4)$.

Solution

We'll call $(-2,-3)$ point #1and $(-7,4)$ point #2. $\begin{pmatrix} x_1, & y_1 \\ -2 & -3 \end{pmatrix} \begin{pmatrix} x_2, & y_2 \\ -7 & 4 \end{pmatrix}$

Use the slope formula. $m = \frac{y_2 - y_1}{x_2 - x_1}$

Substitute the values.

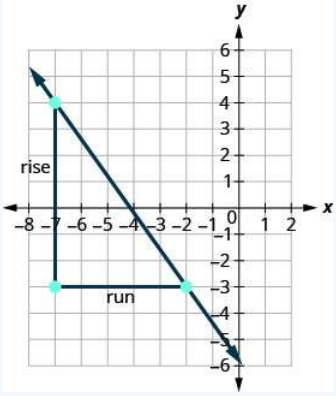
y of the second point minus y of the first point

x of the second point minus x of the first point $m = \frac{4-(-3)}{-7-(-2)}$

Simplify $m = \frac{7}{-5}$

$m = \frac{-7}{5}$

Let's verify this slope on the graph shown.



$$m = \frac{\text{rise}}{\text{run}}$$

$$m = \frac{7}{-5}$$

$$m = \frac{-7}{5}$$

Try It 2.2.1.8

Use the slope formula to find the slope of the line through the pair of points: $(-3, 4)$ and $(2, -1)$.

Answer

-1

Try It 2.2.1.9

Use the slope formula to find the slope of the line through the pair of points: $(-2, 6)$ and $(-3, -4)$.

Answer

10

Graph a Line Given a Point and the Slope

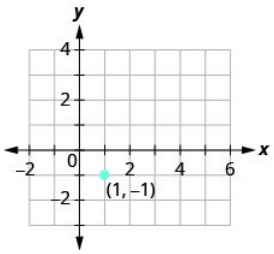
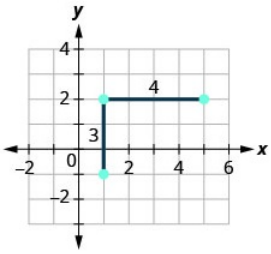
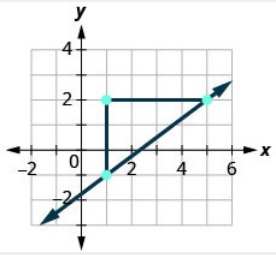
Up to now, in this chapter, we have graphed lines by plotting points, by using intercepts, and by recognizing horizontal and vertical lines.

We can also graph a line when we know one point and the slope of the line. We will start by plotting the point and then use the definition of slope to draw the graph of the line.

Example 2.2.1.10: How to graph a Line Given a Point and the Slope

Graph the line passing through the point $(1, -1)$ whose slope is $m = \frac{3}{4}$.

Solution

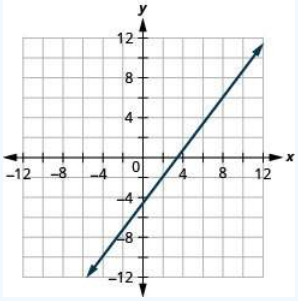
Step 1. Plot the given point.	Plot $(1, -1)$	
Step 2. Use the slope formula $m = \frac{\text{rise}}{\text{run}}$ to identify the rise and the run.	Identify the rise and the run.	$m = \frac{3}{4}$ $\frac{\text{rise}}{\text{run}} = \frac{3}{4}$ rise = 3 run = 4
Step 3. Starting at the given point, count out the rise and run to mark the second point.	Start at $(1, -1)$ and count the rise and the run. Up 3 units, right 4 units.	
Step 4. Connect the points with a line.	Connect the two points with a line.	

You can check your work by finding a third point. Since the slope is $m = \frac{3}{4}$, it can also be written as $m = \frac{-3}{-4}$ (negative divided by negative is positive!). Go back to $(1, -1)$ and count out the rise, -3, and the run, -4.

Try It 2.2.1.11

Graph the line passing through the point $(2, -2)$ with the slope $m = \frac{4}{3}$.

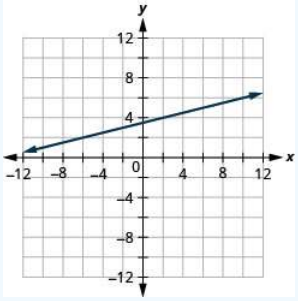
Answer



Try It 2.2.1.12

Graph the line passing through the point $(-2, 3)$ with the slope $m = \frac{1}{4}$.

Answer



GRAPH A LINE GIVEN A POINT AND THE SLOPE.

1. Plot the given point.
2. Use the slope formula $m = \frac{\text{rise}}{\text{run}}$ to identify the rise and the run.
3. Starting at the given point, count out the rise and run to mark the second point.
4. Connect the points with a line.

Graph a Line Using its Slope and Intercept

We have graphed linear equations by plotting points, using intercepts, recognizing horizontal and vertical lines, and using one point and the slope of the line. Once we see how an equation in slope–intercept form and its graph are related, we'll have one more method we can use to graph lines.

See Figure. Let's look at the graph of the equation $y = 12x + 3$ and find its slope and *y*-intercept.

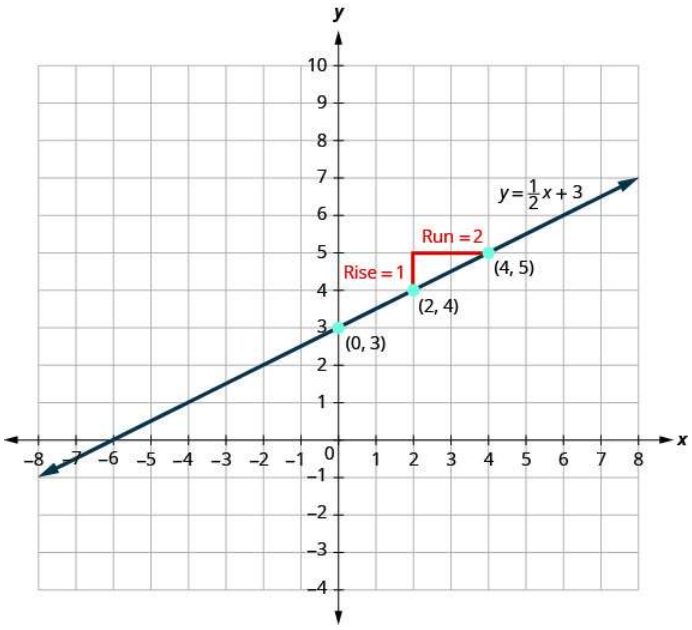


Figure 2.2.1.2

The red lines in the graph show us the rise is 1 and the run is 2. Substituting into the slope formula:

$$m = \frac{\text{rise}}{\text{run}}$$

$$m = \frac{1}{2}$$

The *y*-intercept is $(0, 3)$.

Look at the equation of this line. $y = \frac{1}{2}x + 3$

Look at the slope and *y*-intercept. slope $m = \frac{1}{2}$ and y-intercept $(0, 3)$

When a linear equation is solved for *y*, the coefficient of the *x* term is the slope and the constant term is the *y*-coordinate of the *y*-intercept. We say that the equation $y = \frac{1}{2}x + 3$ is in slope–intercept form. Sometimes the slope–intercept form is called the "*y*-form."

$m = \frac{1}{2}$; *y*-intercept is (0, 3)

$y = \frac{1}{2}x + 3$

$y = mx + b$

SLOPE INTERCEPT FORM OF AN EQUATION OF A LINE

The slope–intercept form of an equation of a line with slope *m* and *y*-intercept, $(0, b)$ is $y = mx + b$.

Let's practice finding the values of the slope and *y*-intercept from the equation of a line.

Example 2.2.1.13

Identify the slope and *y*-intercept of the line from the equation: a. $y = -\frac{4}{7}x - 2$ b. $x + 3y = 9$

Solution

a. We compare our equation to the slope–intercept form of the equation.

Write the slope–intercept form of the equation of the line.	$y = mx + b$
Write the equation of the line.	$y = -\frac{4}{7}x - 2$
Identify the slope.	$m = -\frac{4}{7}$
Identify the y-intercept.	y=intercept is $(0, -2)$

b. When an equation of a line is not given in slope–intercept form, our first step will be to solve the equation for y.

Solve for y.	$x + 3y = 9$
Subtract x from each side.	$3y = -x + 9$
Divide both sides by 3.	$\frac{3y}{3} = \frac{-x+9}{3}$
Simplify.	$y = -\frac{1}{3}x + 3$
Write the slope–intercept form of the equation of the line.	$y = mx + b$
Write the equation of the line.	$y = -\frac{1}{3}x + 3$
Identify the slope.	$m = -\frac{1}{3}$
Identify the y-intercept.	y=intercept is $(0, 3)$

Try It 2.2.1.14

Identify the slope and y-intercept from the equation of the line. a. $y = \frac{2}{5}x - 1$ b. $x + 4y = 8$

Answer

a. $m = \frac{2}{5}$; $(0, -1)$
b. $m = -\frac{1}{4}$; $(0, 2)$

Try It 2.2.1.15

Identify the slope and y-intercept from the equation of the line. a. $y = -\frac{4}{3}x + 1$ b. $3x + 2y = 12$

Answer

a. $m = -\frac{4}{3}$; $(0, 1)$
b. $m = -\frac{3}{2}$; $(0, 6)$

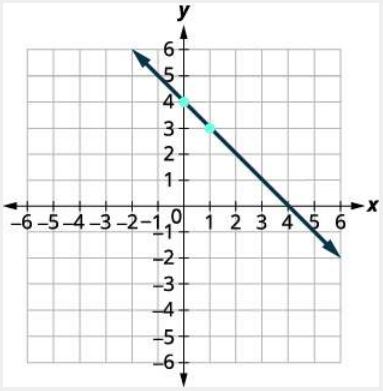
We have graphed a line using the slope and a point. Now that we know how to find the slope and y-intercept of a line from its equation, we can use the y-intercept as the point, and then count out the slope from there.

Example 2.2.1.16

Graph the line of the equation $y = -x + 4$ using its slope and y-intercept.

Solution

	$y = mx + b$
The equation is in slope–intercept form.	$y = -x + 4$

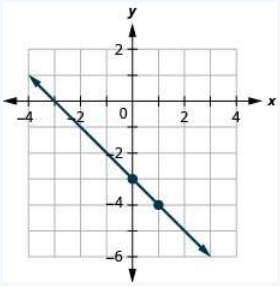
Identify the slope and y-intercept.	$m = -1$ y-intercept is $(0, 4)$
Plot the y-intercept.	See the graph.
Identify the rise over the run.	$m = \frac{-1}{1}$
Count out the rise and run to mark the second point.	rise -1, run 1 

Draw the line as shown in the graph.

Try It 2.2.1.17

Graph the line of the equation $y = -x - 3$ using its slope and y-intercept.

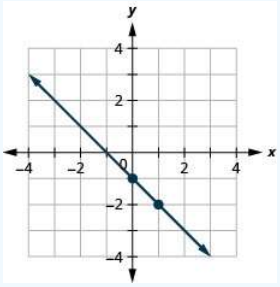
Answer



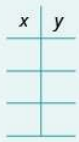
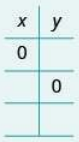
Try It 2.2.1.18

Graph the line of the equation $y = -x - 1$ using its slope and y-intercept.

Answer



Now that we have graphed lines by using the slope and y-intercept, let's summarize all the methods we have used to graph lines.

Methods to Graph Lines			
Point Plotting	**Slope–Intercept**	**Intercepts**	**Recognize Vertical and Horizontal Lines**
	$y = mx + b$		
Find three points. Plot the points, make sure they line up, then draw the line.	Find the slope and *y*-intercept. Start at the *y*-intercept, then count the slope to get a second point.	Find the intercepts and a third point. Plot the points, make sure they line up, then draw the line.	The equation has only one variable. $x = a$ vertical $y = b$ horizontal

Choose the Most Convenient Method to Graph a Line

Now that we have seen several methods we can use to graph lines, how do we know which method to use for a given equation?

While we could plot points, use the slope–intercept form, or find the intercepts for *any* equation, if we recognize the most convenient way to graph a certain type of equation, our work will be easier.

Generally, plotting points is not the most efficient way to graph a line. Let's look for some patterns to help determine the most convenient method to graph a line.

Here are five equations we graphed in this chapter, and the method we used to graph each of them.

	Equation	**Method**
#1	$x = 2$	Vertical line
#2	$y = -1$	Horizontal line
#3	$-x + 2y = 6$	Intercepts
#4	$4x - 3y = 12$	Intercepts
#5	$y = -x + 4$	Slope–intercept

Equations #1 and #2 each have just one variable. Remember, in equations of this form the value of that one variable is constant; it does not depend on the value of the other variable. Equations of this form have graphs that are vertical or horizontal lines.

In equations #3 and #4, both *x* and *y* are on the same side of the equation. These two equations are of the form $Ax + By = C$. We substituted $y = 0$ to find the *x*- intercept and $x = 0$ to find the *y*-intercept, and then found a third point by choosing another value for *x* or *y*.

Equation #5 is written in slope–intercept form. After identifying the slope and *y*-intercept from the equation we used them to graph the line.

This leads to the following strategy.

STRATEGY FOR CHOOSING THE MOST CONVENIENT METHOD TO GRAPH A LINE

Consider the form of the equation.

- If it only has one variable, it is a vertical or horizontal line.
  - $x = a$ is a vertical line passing through the *x*-axis at *a*.
  - $y = b$ is a horizontal line passing through the *y*-axis at *b*.
- If *y* is isolated on one side of the equation, in the form $y = mx + b$, graph by using the slope and *y*-intercept.
  - Identify the slope and *y*-intercept and then graph.
- If the equation is of the form $Ax + By = C$, find the intercepts.
  - Find the *x*- and *y*-intercepts, a third point, and then graph.

? Example 2.2.1.19

Determine the most convenient method to graph each line:

ⓐ $y=5$ ⓑ $4x-5y=20$ ⓒ $x=-3$ ⓓ $y=-\frac{5}{9}x+8$

Answer

ⓐ $y=5$
This equation has only one variable, *y*. Its graph is a horizontal line crossing the *y*-axis at 5.
ⓑ $4x-5y=20$
This equation is of the form $Ax+By=C$. The easiest way to graph it will be to find the intercepts and one more point.
ⓒ $x=-3$
There is only one variable, *x*. The graph is a vertical line crossing the *x*-axis at -3.
ⓓ $y=-\frac{5}{9}x+8$
Since this equation is in $y=mx+b$ form, it will be easiest to graph this line by using the slope and *y*-intercepts.

? Example 2.2.1.20

Determine the most convenient method to graph each line:

ⓐ $3x+2y=12$ ⓑ $y=4$ ⓒ $y=\frac{1}{5}x-4$ ⓓ $x=-7$.

Answer

ⓐ intercepts ⓑ horizontal line ⓒ slope-intercept ⓓ vertical line

? Example 2.2.1.21

Determine the most convenient method to graph each line:

ⓐ $x=6$ ⓑ $y=-\frac{3}{4}x+1$ ⓒ $y=-8$ ⓓ $4x-3y=-1$.

Answer

ⓐ vertical line ⓑ slope-intercept ⓒ horizontal line
ⓓ intercepts

Graph and Interpret Applications of Slope-Intercept

Many real-world applications are modeled by linear equations. We will take a look at a few applications here so you can see how equations written in slope-intercept form relate to real-world situations.

Usually, when a linear equation models uses real-world data, different letters are used for the variables, instead of using only *x* and *y*. The variable names remind us of what quantities are being measured.

Also, we often will need to extend the axes in our rectangular coordinate system to bigger positive and negative numbers to accommodate the data in the application.

? Example 2.2.1.22

The equation $F=\frac{9}{5}C+32$ is used to convert temperatures, *C*, on the Celsius scale to temperatures, *F*, on the Fahrenheit scale.

ⓐ Find the Fahrenheit temperature for a Celsius temperature of 0.

ⓑ Find the Fahrenheit temperature for a Celsius temperature of 20.

ⓒ Interpret the slope and *F*-intercept of the equation.

ⓓ Graph the equation.

Answer

ⓐ

Find the Fahrenheit temperature for a Celsius temperature of 0.	$F = \frac{9}{5}C + 32$
Find F when C=0.	$F = \frac{9}{5}(0) + 32$
Simplify.	$F = 32$

ⓑ

Find the Fahrenheit temperature for a Celsius temperature of 20.	$F = \frac{9}{5}C + 32$
Find F when C=20.	$F = \frac{9}{5}(20) + 32$
Simplify.	$F = 36 + 32$
Simplify.	$F = 68$

ⓒ

Interpret the slope and *F*-intercept of the equation.

Even though this equation uses *F* and *C*, it is still in slope-intercept form.

$y = mx + b$

$F = mC + b$

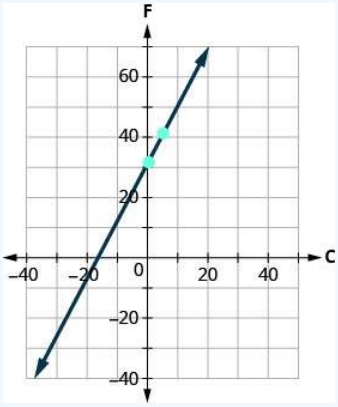
$F = \frac{9}{5}C + 32$

The slope, $\frac{9}{5}$, means that the temperature Fahrenheit (*F*) increases 9 degrees when the temperature Celsius (*C*) increases 5 degrees.

The *F*-intercept means that when the temperature is 0° on the Celsius scale, it is 32° on the Fahrenheit scale.

ⓓ Graph the equation.

We'll need to use a larger scale than our usual. Start at the *F*-intercept $(0, 32)$, and then count out the rise of 9 and the run of 5 to get a second point as shown in the graph.



? Example 2.2.1.23

The equation \(h=2s+50\) is used to estimate a woman's height in inches, *h*, based on her shoe size, *s*.

ⓐ Estimate the height of a child who wears women's shoe size 0.

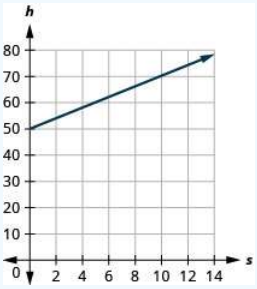
ⓑ Estimate the height of a woman with shoe size 8.

ⓒ Interpret the slope and *h*-intercept of the equation.

ⓓ Graph the equation.

Answer

ⓐ 50 inches
ⓑ 66 inches
ⓒ The slope, 2, means that the height, *h*, increases by 2 inches when the shoe size, *s*, increases by 1. The *h*-intercept means that when the shoe size is 0, the height is 50 inches.
ⓓ



Example 2.2.1.24

The equation \(T=\frac{1}{4}n+40\) is used to estimate the temperature in degrees Fahrenheit, *T*, based on the number of cricket chirps, *n*, in one minute.

ⓐ Estimate the temperature when there are no chirps.

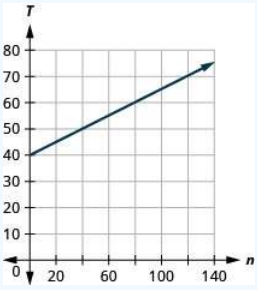
ⓑ Estimate the temperature when the number of chirps in one minute is 100.

ⓒ Interpret the slope and *T*-intercept of the equation.

ⓓ Graph the equation.

Answer

ⓐ 40 degrees
ⓑ 65 degrees
ⓒ The slope, $\frac{1}{4}$, means that the temperature Fahrenheit (*F*) increases 1 degree when the number of chirps, *n*, increases by 4. The *T*-intercept means that when the number of chirps is 0, the temperature is 40°.
ⓓ



The cost of running some types business have two components—a *fixed cost* and a *variable cost*. The fixed cost is always the same regardless of how many units are produced. This is the cost of rent, insurance, equipment, advertising, and other items that must be paid regularly. The variable cost depends on the number of units produced. It is for the material and labor needed to produce each item.

Example 2.2.1.25

Sam drives a delivery van. The equation $C = 0.5m + 60$ models the relation between his weekly cost, *C*, in dollars and the number of miles, *m*, that he drives.

ⓐ Find Sam's cost for a week when he drives 0 miles.

ⓑ Find the cost for a week when he drives 250 miles.

ⓒ Interpret the slope and *C*-intercept of the equation.

ⓓ Graph the equation.

Answer

ⓐ

Find Sam's cost for a week when he drives 0 miles.	$C = 0.5m + 60$
Find C when m=0.	$C = 0.5(0) + 60$
Simplify.	$C = 60$
	Sam's costs are \$60 when he drives 0 miles.

ⓑ

Find Sam's cost for a week when he drives 250 miles.	$C = 0.5m + 60$
Find C when m=250.	$C = 0.5(250) + 60$
Simplify.	$C = 185$
	Sam's costs are \$185 when he drives 250 miles.

ⓒ Interpret the slope and *C*-intercept of the equation.

$y = mx + b$

$C = 0.5n + 60$

The slope, 0.5, means that the weekly cost, *C*, increases by \$0.50 when the number of miles driven, *n*, increases by 1.
The *C*-intercept means that when the number of miles driven is 0, the weekly cost is \$60.

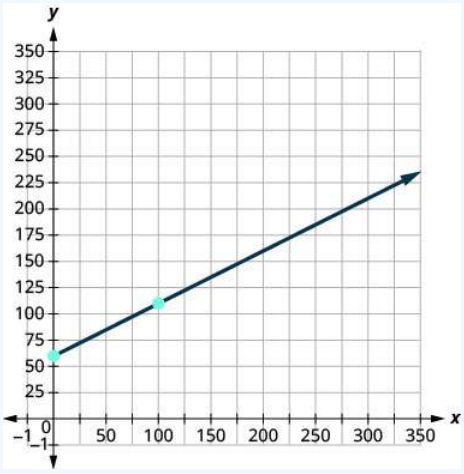
ⓓ Graph the equation.

We'll need to use a larger scale than our usual. Start at the *C*-intercept $(0, 60)$.

To count out the slope $m = 0.5$, we rewrite it as an equivalent fraction that will make our graphing easier.

	$m = 0.5$
Rewrite as a fraction.	$m = \frac{0.5}{1}$
Multiply numerator and denominator by 100	$m = \frac{0.5(100)}{1(100)}$
Simplify.	$m = \frac{50}{100}$

So to graph the next point go up 50 from the intercept of 60 and then to the right 100. The second point will be $(100, 110)$



Example 2.2.1.26

Stella has a home business selling gourmet pizzas. The equation $C = 4p + 25$ models the relation between her weekly cost, C, in dollars and the number of pizzas, p, that she sells.

ⓐ Find Stella's cost for a week when she sells no pizzas.

ⓑ Find the cost for a week when she sells 15 pizzas.

ⓒ Interpret the slope and C-intercept of the equation.

ⓓ Graph the equation.

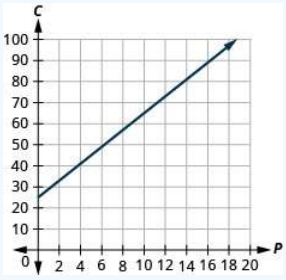
Answer

ⓐ $25

ⓑ $85

ⓒ The slope, 4, means that the weekly cost, C, increases by $4 when the number of pizzas sold, p, increases by 1. The C-intercept means that when the number of pizzas sold is 0, the weekly cost is $25.

ⓓ



Example 2.2.1.27

Loreen has a calligraphy business. The equation $C = 1.8n + 35$ models the relation between her weekly cost, C, in dollars and the number of wedding invitations, n, that she writes.

ⓐ Find Loreen's cost for a week when she writes no invitations.

ⓑ Find the cost for a week when she writes 75 invitations.

ⓒ Interpret the slope and C-intercept of the equation.

ⓓ Graph the equation.

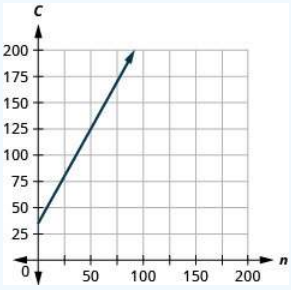
Answer

ⓐ $35

ⓑ $170

ⓒ The slope, 1.8, means that the weekly cost, C, increases by $\$1.80$when the number of invitations, n, increases by 1. The C-intercept means that when the number of invitations is 0, the weekly cost is $35.

ⓓ



Use Slopes to Identify Parallel and Perpendicular Lines

Two lines that have the same slope are called **parallel lines**. Parallel lines have the same steepness and never intersect.

We say this more formally in terms of the rectangular coordinate system. Two lines that have the same slope and different *y*-intercepts are called parallel lines. See Figure.

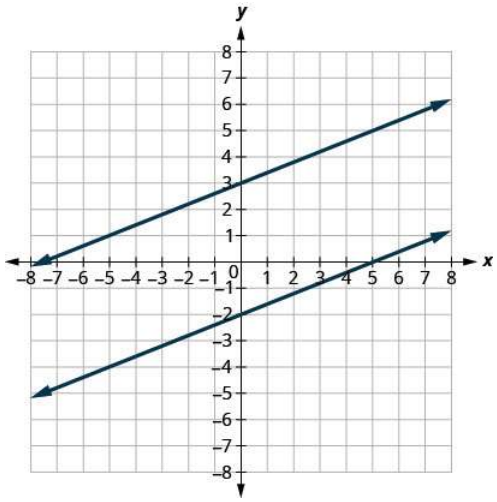


Figure 2.2.1.3

Verify that both lines have the same slope, $m = \frac{2}{5}$, and different *y*-intercepts.

What about vertical lines? The slope of a vertical line is undefined, so vertical lines don't fit in the definition above. We say that vertical lines that have different *x*-intercepts are parallel, like the lines shown in this graph.

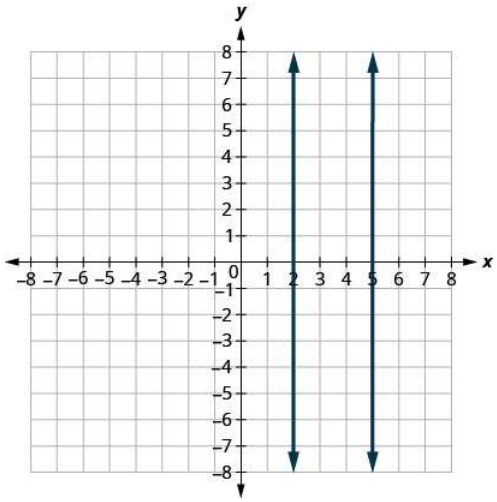


Figure 2.2.1.4

PARALLEL LINES

Parallel lines are lines in the same plane that do not intersect.

- Parallel lines have the same slope and different *y*-intercepts.
- If $m1$ and $m2$ are the slopes of two parallel lines then $m1 = m2$.
- Parallel vertical lines have different *x*-intercepts

Since parallel lines have the same slope and different *y*-intercepts, we can now just look at the slope–intercept form of the equations of lines and decide if the lines are parallel.

Example 2.2.1.28

Use slopes and y-intercepts to determine if the lines are parallel:

ⓐ $3x-2y=6$ and $y=\frac{3}{2}x+1$ ⓑ $y=2x-3$ and $-6x+3y=-9$.

Answer

ⓐ

	$3x-2y=6$	and	$y=\frac{3}{2}x+1$
	$-2y=-3x+6$		
Solve the first equation for y.	$\frac{-2y}{-2}=\frac{-3x+6}{-2}$		
The equation is now in slope–intercept form.	$y=\frac{3}{2}x-3$		
The equation of the second line is already in slope–intercept form.			$y=\frac{3}{2}x+1$
	$y=\frac{3}{2}x-3$		$y=\frac{3}{2}x+1$
Identify the slope and y-intercept of both lines.	$y=mx+b$		$y=mx+b$
	$m=\frac{3}{2}$		$y=\frac{3}{2}$
	y-intercept is $(0,-3)$		y-intercept is $(0,1)$

The lines have the same slope and different y-intercepts and so they are parallel.
You may want to graph the lines to confirm whether they are parallel.

ⓑ

	$y=2x-3$	and	$-6x+3y=-9$
The first equation is already in slope–intercept form.	$y=2x-3$		
			$-6x+3y=-9$
			$3y=6x-9$
Solve the second equation for y.			$\frac{3y}{3}=\frac{6x-9}{3}$
			$y=2x-3$
The second equation is now in slope–intercept form.			$y=2x-3$
	$y=2x-3$		$y=2x-3$
Identify the slope and y-intercept of both lines.	$y=mx+b$		$y=mx+b$
	$m=2$		$m=2$
	y-intercept is $(0,-3)$		y-intercept is $(0,-3)$

The lines have the same slope, but they also have the same y-intercepts. Their equations represent the same line and we say the lines are coincident. They are not parallel; they are the same line.

Example 2.2.1.29

Use slopes and y-intercepts to determine if the lines are parallel:

ⓐ $2x+5y=5$ and $y=-\frac{2}{5}x-4$ ⓑ $y=-\frac{1}{2}x-1$ and $x+2y=-2$.

Answer

ⓐ parallel ⓑ not parallel; same line

Example 2.2.1.30

Use slopes and *y*-intercepts to determine if the lines are parallel:

ⓐ $4x-3y=6$ and $y=\frac{4}{3}x-1$ ⓑ $y=\frac{3}{4}x-3$ and $3x-4y=12$.

Answer

ⓐ parallel ⓑ not parallel; same line

Example 2.2.1.31

Use slopes and *y*-intercepts to determine if the lines are parallel:

ⓐ $y=-4$ and $y=3$ ⓑ $x=-2$ and $x=-5$.

Answer

ⓐ $y=-4$ and $y=3$

We recognize right away from the equations that these are horizontal lines, and so we know their slopes are both 0.
Since the horizontal lines cross the *y*-axis at y=−4y=−4 and at y=3,y=3, we know the *y*-intercepts are (0,−4)(0,−4) and (0,3).
(0,3).
The lines have the same slope and different *y*-intercepts and so they are parallel.

ⓑ $x=-2$ and $x=-5$

We recognize right away from the equations that these are vertical lines, and so we know their slopes are undefined.
Since the vertical lines cross the *x*-axis at $x=-2$ and $x=-5$, we know the *y*-intercepts are $(-2,0)$ and $(-5,0)$.
The lines are vertical and have different *x*-intercepts and so they are parallel.

Example 2.2.1.32

Use slopes and *y*-intercepts to determine if the lines are parallel:

ⓐ $y=8$ and $y=-6$ ⓑ $x=1$ and $x=-5$.

Answer

ⓐ parallel ⓑ parallel

Example 2.2.1.33

Use slopes and *y*-intercepts to determine if the lines are parallel:

ⓐ $y=1$ and $y=-5$ ⓑ $x=8$ and $x=-6$.

Answer

ⓐ parallel ⓑ parallel

Let's look at the lines whose equations are $y=\frac{1}{4}x-1$ and $y=-4x+2$, shown in Figure.

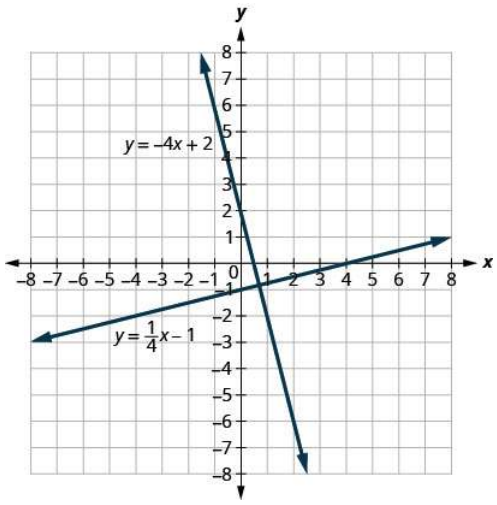


Figure 2.2.1.5

These lines lie in the same plane and intersect in right angles. We call these lines perpendicular.

If we look at the slope of the first line, $m_1 = \frac{1}{4}$, and the slope of the second line, $m_2 = -4$, we can see that they are *negative reciprocals* of each other. If we multiply them, their product is -1.

$$m_1 \cdot m_2$$
$$14(-4)$$
$$-1$$

This is always true for **perpendicular lines** and leads us to this definition.

PERPENDICULAR LINES

Perpendicular lines are lines in the same plane that form a right angle.

- If m_1 and m_2 are the slopes of two perpendicular lines, then:
  - their slopes are negative reciprocals of each other, $m_1 = -\frac{1}{m_2}$.
  - the product of their slopes is -1, $m_1 \cdot m_2 = -1$.
- A vertical line and a horizontal line are always perpendicular to each other

We were able to look at the slope–intercept form of linear equations and determine whether or not the lines were parallel. We can do the same thing for perpendicular lines.

We find the slope–intercept form of the equation, and then see if the slopes are opposite reciprocals. If the product of the slopes is -1, the lines are perpendicular.

Example 2.2.1.34

Use slopes to determine if the lines are perpendicular:

ⓐ $y = -5x - 4$ and $x - 5y = 5$ ⓑ $7x + 2y = 3$ and $2x + 7y = 5$

Answer

ⓐ

The first equation is in slope–intercept form.Solve the second equation fory.Identify the slope of each line.y=−5x−4yym1=−5x−4=mx+b=−5x−5y−5y−5y=5=−x+5=−x+5−5=15x−1yym2=15x−1=mx+b=15The first equation is in slope–intercept form.y=−5x−4Solve the second equation fory.x−5y=5−5y=−x+5−5y−5=−x+5−5y=15x−1Identify the slope of each line.y=−5x−4y=mx+bm1=−5y=15x−1y=mx+bm2=15

The slopes are negative reciprocals of each other, so the lines are perpendicular. We check by multiplying the slopes, Since

−5(15)=−1,−5(15)=−1, it checks.

ⓑ
Solve the equations fory.Identify the slope of each line.7x+2y2y2y2y=3=−7x+3=−7x+32=−72x+32ym1=mx+b=−722x+7y7y7y7y=5=−2x+5=−2x+57=−27x+57ym1=mx+b=−27Solve the equations fory.7x+2y=32y=−7x+32y2=−7x+32y=−72x+322x+7y=57y=−2x+57y7=−2x+57y=−27x+57Identify the slope of each line.y=mx+bm1=−72y=mx+bm1=−27
The slopes are reciprocals of each other, but they have the same sign. Since they are not negative reciprocals, the lines are not perpendicular.

? Example 2.2.1.3

Use slopes to determine if the lines are perpendicular:

ⓐ $y=-3x+2$ and $x-3y=4$ ⓑ $5x+4y=1$ and $4x+5y=3$.

Answer

ⓐ perpendicular ⓑ not perpendicular

? Example 2.2.1.3

Use slopes to determine if the lines are perpendicular:

ⓐ $y=2x-5$ and $x+2y=-6$ ⓑ $2x-9y=3$ and $9x-2y=1$.

Answer

ⓐ perpendicular ⓑ not perpendicular

Key Concepts

- **Slope of a Line**
  - The slope of a line is $m=\frac{\text{rise}}{\text{run}}$.
  - The rise measures the vertical change and the run measures the horizontal change.

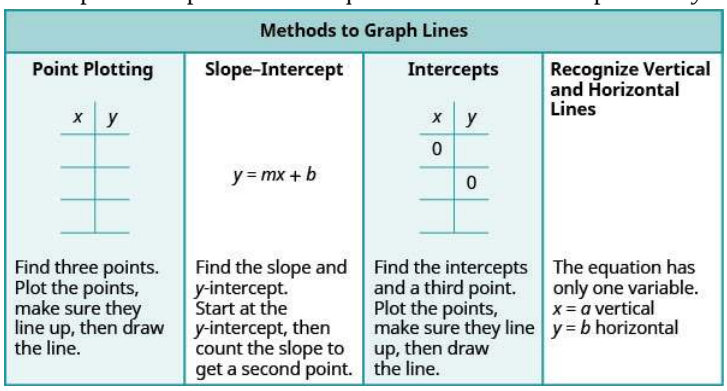
- **How to find the slope of a line from its graph using $m=\frac{\text{rise}}{\text{run}}$.**
  1. Locate two points on the line whose coordinates are integers.
  2. Starting with one point, sketch a right triangle, going from the first point to the second point.
  3. Count the rise and the run on the legs of the triangle.
  4. Take the ratio of rise to run to find the slope: $m=\frac{\text{rise}}{\text{run}}$.
- **Slope of a line between two points.**
  - The slope of the line between two points (x_1,y_1) and (x_2,y_2) is:

$$m=\frac{y_2-y_1}{x_2-x_1}$$

.

- **How to graph a line given a point and the slope.**
  1. Plot the given point.
  2. Use the slope formula $m=\frac{\text{rise}}{\text{run}}$ to identify the rise and the run.
  3. Starting at the given point, count out the rise and run to mark the second point.

4. Connect the points with a line.

- **Slope Intercept Form of an Equation of a Line**
  - The slope–intercept form of an equation of a line with slope m and y-intercept, $(0, b)$ is $y = mx + b$
  - 

- **Parallel Lines**
  - Parallel lines are lines in the same plane that do not intersect.
    Parallel lines have the same slope and different y-intercepts.
    If m_1 and m_2 are the slopes of two parallel lines then $m_1 = m_2$.
    Parallel vertical lines have different x-intercepts.
- **Perpendicular Lines**
  - Perpendicular lines are lines in the same plane that form a right angle.
  - If m_1 and m_2 are the slopes of two perpendicular lines, then:
    their slopes are negative reciprocals of each other, $m_1 = -\frac{1}{m_2}$.
    the product of their slopes is -1, $m_1 \cdot m_2 = -1$.
  - A vertical line and a horizontal line are always perpendicular to each other.

Glossary

parallel lines

Parallel lines are lines in the same plane that do not intersect.

perpendicular lines

Perpendicular lines are lines in the same plane that form a right angle.

This page titled 2.2.1: Slope of a Line is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **3.3: Slope of a Line** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/intermediate-algebra-2e.

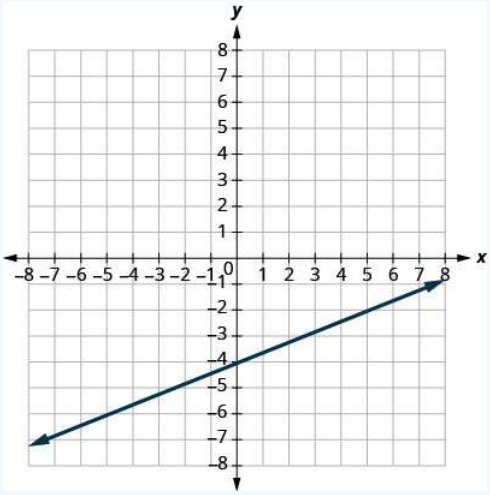
2.2.1E: Exercises

Practice Makes Perfect

Find the Slope of a Line

In the following exercises, find the slope of each line shown.

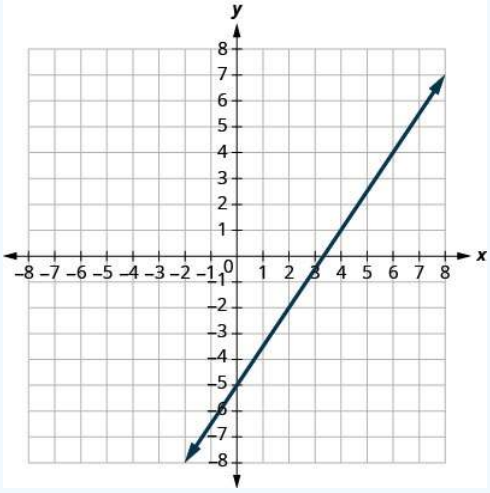
1.



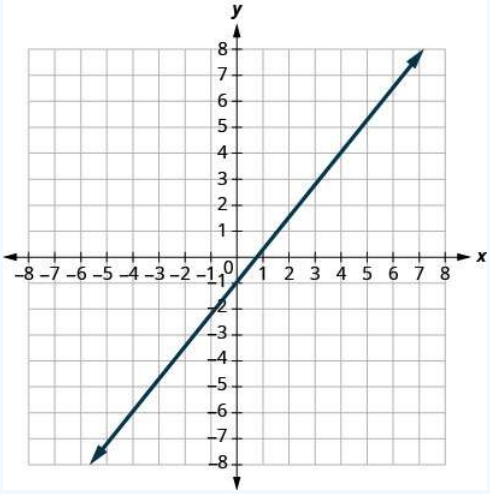
Answer

$m = \frac{2}{5}$

2.



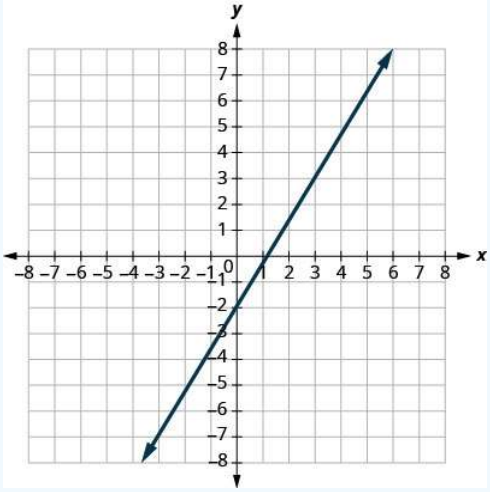
3.



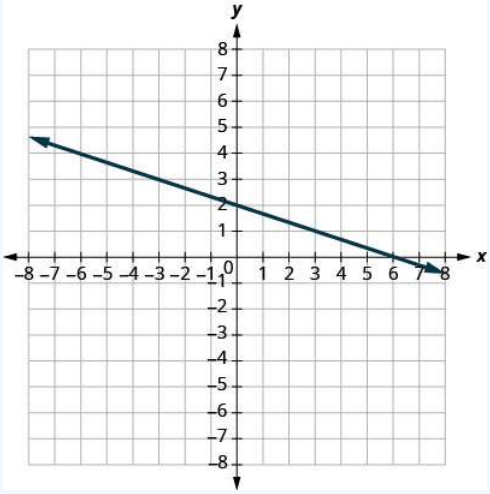
Answer

$m = \frac{5}{4}$

4.



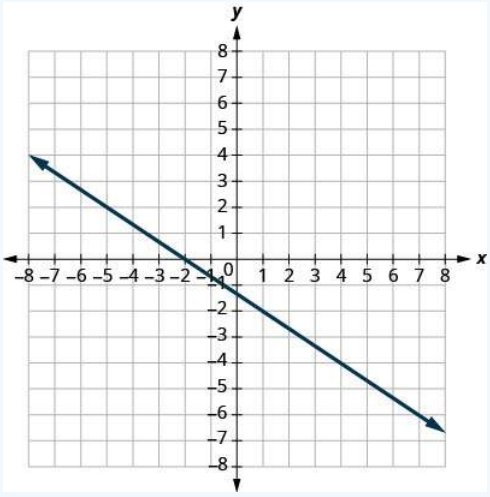
5.



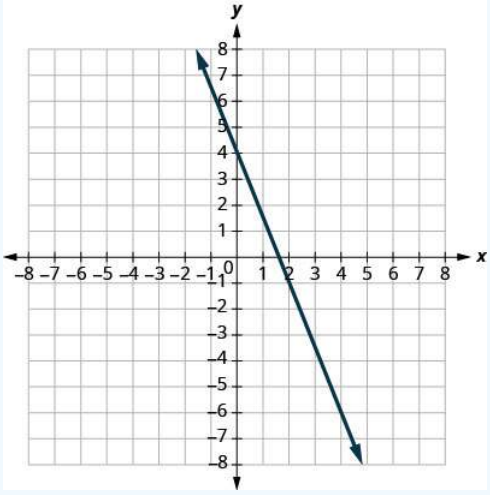
Answer

$m = -\frac{1}{3}$

6.



7.

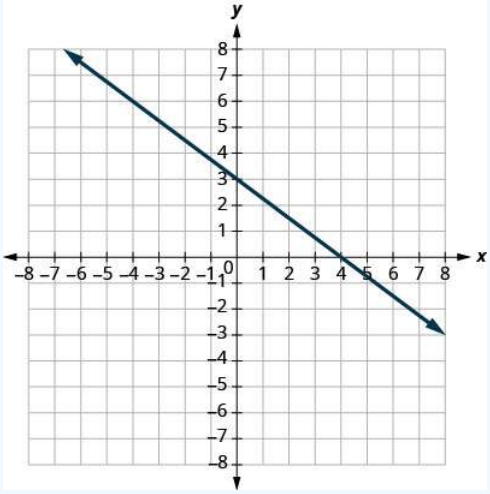


Answer

$m = -\frac{5}{2}$

https://math.libretexts.org/@go/page/139511

8.



In the following exercises, find the slope of each line.

9. $y = 3$

Answer

$m = 0$

10. $y = -2$

11. $x = -5$

Answer

undefined

12. $x = 4$

In the following exercises, use the slope formula to find the slope of the line between each pair of points.

13. $(2, 5),\ (4, 0)$

Answer

$m = -\frac{5}{2}$

14. $(3, 6),\ (8, 0)$

15. $(-3, 3),\ (4, -5)$

Answer

$m = -\frac{8}{7}$

16. $(-2, 4),\ (3, -1)$

17. $(-1, -2),\ (2, 5)$

Answer

$m = \frac{7}{3}$

18. $(-2, -1),\ (6, 5)$

19. $(4, -5),\ (1, -2)$

Answer

$m = -1$

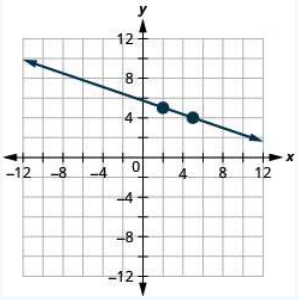
20. $(3, -6),\ (2, -2)$

Graph a Line Given a Point and the Slope

In the following exercises, graph each line with the given point and slope.

21. $(2, 5)$; $m = -\frac{1}{3}$

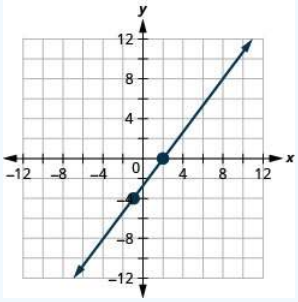
Answer



22. $(1, 4)$; $m = -\frac{1}{2}$

23. $(-1, -4)$; $m = \frac{4}{3}$

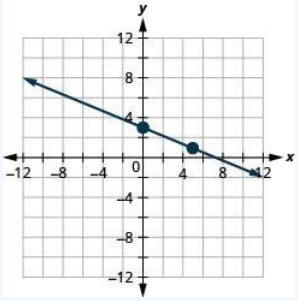
Answer



24. $(-3, -5)$; $m = \frac{3}{2}$

25. y-intercept: $(0, 3)$; $m = -\frac{2}{5}$

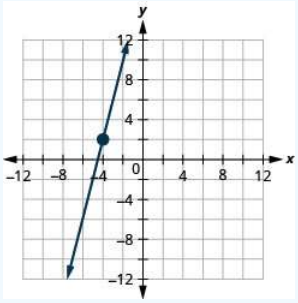
Answer



26. x-intercept: $(-2, 0)$; $m = \frac{3}{4}$

27. $(-4, 2)$; $m = 4$

Answer



28. $(1, 5)$; $m = -3$

Graph a Line Using Its Slope and Intercept

In the following exercises, identify the slope and *y*-intercept of each line.

29. $y = -7x + 3$

Answer

$m = -7$; $(0, 3)$

30. $y = 4x - 10$

31. $3x + y = 5$

Answer

$m = -3$; $(0, 5)$

32. $4x + y = 8$

33. $6x + 4y = 12$

Answer

$m = -\frac{3}{2}$; $(0, 3)$

34. $8x + 3y = 12$

35. $5x - 2y = 6$

Answer

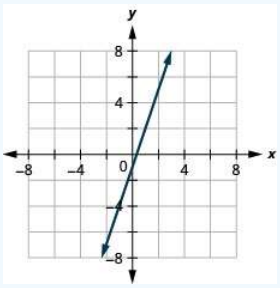
$m = \frac{5}{2}$; $(0, -3)$

36. $7x - 3y = 9$

In the following exercises, graph the line of each equation using its slope and *y*-intercept.

37. $y = 3x - 1$

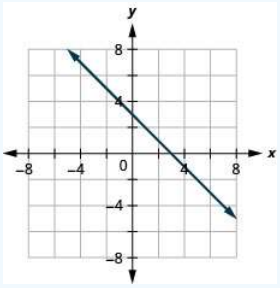
Answer



38. $y = 2x - 3$

39. $y = -x + 3$

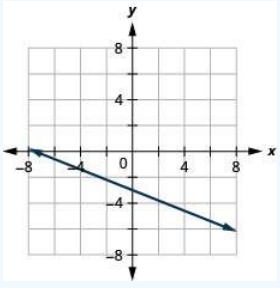
Answer



40. $y = -x - 4$

41. $y = -\frac{2}{5}x - 3$

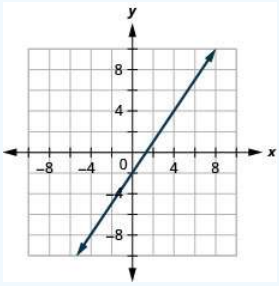
Answer



42. $y = -\frac{3}{5}x + 2$

43. $3x - 2y = 4$

Answer



44. $3x - 4y = 8$

Choose the Most Convenient Method to Graph a Line

In the following exercises, determine the most convenient method to graph each line.

45. $x = 2$

Answer

vertical line

46. $y = 5$

47. $y = -3x + 4$

Answer

slope-intercept

48. $x - y = 5$

49. $x - y = 1$

Answer

intercepts

50. $y = \frac{2}{3}x - 1$

51. $3x - 2y = -12$

Answer

intercepts

52. $2x - 5y = -10$

Graph and Interpret Applications of Slope–Intercept

53. The equation $P = 31 + 1.75w$ models the relation between the amount of Tuyet's monthly water bill payment, P, in dollars, and the number of units of water, w, used.

a. Find Tuyet's payment for a month when 0 units of water are used.

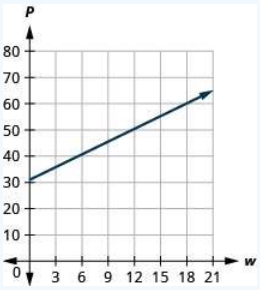
b. Find Tuyet's payment for a month when 12 units of water are used.

c. Interpret the slope and P-intercept of the equation.

d. Graph the equation.

Answer

a. $\$31$
b. $\$52$
c. The slope, 1.75, means that the payment, P, increases by $\$1.75$ when the number of units of water used, w, increases by 1. The P-intercept means that when the number units of water Tuyet used is 0, the payment is $\$31$.
d.



54. The equation $P = 28 + 2.54w$ models the relation between the amount of Randy's monthly water bill payment, P, in dollars, and the number of units of water, w, used.

a. Find the payment for a month when Randy used 0 units of water.

b. Find the payment for a month when Randy used 15 units of water.

c. Interpret the slope and P-intercept of the equation.

d. Graph the equation.

55. Bruce drives his car for his job. The equation $R = 0.575m + 42$ models the relation between the amount in dollars, R, that he is reimbursed and the number of miles, m, he drives in one day.

a. Find the amount Bruce is reimbursed on a day when he drives 0 miles.

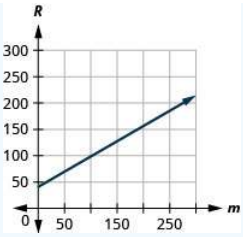
b. Find the amount Bruce is reimbursed on a day when he drives 220 miles.

c. Interpret the slope and R-intercept of the equation.

d. Graph the equation.

Answer

a. $\$42$
b. $\$168.50$
c. The slope, 0.575, means that the amount he is reimbursed, R, increases by $\$0.575$ when the number of miles driven, m, increases by 1. The R-intercept means that when the number miles driven is 0, the amount reimbursed is $\$42$.
d.



56. Janelle is planning to rent a car while on vacation. The equation $C = 0.32m + 15$ models the relation between the cost in dollars, C, per day and the number of miles, m, she drives in one day.

a. Find the cost if Janelle drives the car 0 miles one day.

b. Find the cost on a day when Janelle drives the car 400 miles.

c. Interpret the slope and C-intercept of the equation.

d. Graph the equation.

57. Cherie works in retail and her weekly salary includes commission for the amount she sells. The equation $S = 400 + 0.15c$ models the relation between her weekly salary, S, in dollars and the amount of her sales, c, in dollars.

a. Find Cherie's salary for a week when her sales were $\$0$.

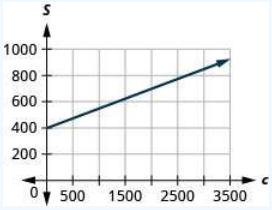
b. Find Cherie's salary for a week when her sales were $\$3,600$

c. Interpret the slope and S-intercept of the equation.

d. Graph the equation.

Answer

a. $\$400$
b. $\$940$
c. The slope, 0.15, means that Cherie's salary, S, increases by $\$0.15$ for every $\$1$ increase in her sales. The S-intercept means that when her sales are $\$0$, her salary is $\$400$
d.



58. Patel's weekly salary includes a base pay plus commission on his sales. The equation $S = 750 + 0.09c$ models the relation between his weekly salary, S, in dollars and the amount of his sales, c, in dollars.

a. Find Patel's salary for a week when his sales were 0.

b. Find Patel's salary for a week when his sales were $18,540$

c. Interpret the slope and S-intercept of the equation.

d. Graph the equation.

59. Costa is planning a lunch banquet. The equation $C = 450 + 28g$ models the relation between the cost in dollars, C, of the banquet and the number of guests, g.

a. Find the cost if the number of guests is 40.

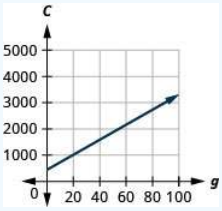
b. Find the cost if the number of guests is 80.

c. Interpret the slope and C-intercept of the equation.

d. Graph the equation.

Answer

a. $\$1570$
b. $\$5690$
c. The slope gives the cost per guest. The slope, 28, means that the cost, C, increases by $\$28$ when the number of guests increases by 1. The C-intercept means that if the number of guests was 0, the cost would be $\$450$
d.



60. Margie is planning a dinner banquet. The equation $C = 750 + 42g$ models the relation between the cost in dollars, C, of the banquet and the number of guests, g.

a. Find the cost if the number of guests is 50.

b. Find the cost if the number of guests is 100.

c. Interpret the slope and C-intercept of the equation.

d. Graph the equation.

Use Slopes to Identify Parallel and Perpendicular Lines

In the following exercises, use slopes and y-intercepts to determine if the lines are parallel, perpendicular, or neither.

61. $y = \frac{3}{4}x - 3\,;\, 3x - 4y = -2$

Answer

parallel

62. $3x - 4y = -2\,;\, y = \frac{3}{4}x - 3$

63. $2x - 4y = 6\,;\, x - 2y = 3$

Answer

neither

64. $8x + 6y = 6\,;\, 12x + 9y = 12$

65. $x = 5\,;\, x = -6$

Answer

parallel

66. $x = -3\,;\, x = -2$

67. $4x-2y=5; 3x+6y=8$

Answer

perpendicular

68. $8x-2y=7; 3x+12y=9$

69. $3x-6y=12; 6x-3y=3$

Answer

neither

70. $9x-5y=4; 5x+9y=-1$

71. $7x-4y=8; 4x+7y=14$

Answer

perpendicular

72. $5x-2y=11; 5x-y=7$

73. $3x-2y=8; 2x+3y=6$

Answer

perpendicular

74. $2x+3y=5; 3x-2y=7$

75. $3x-2y=1; 2x-3y=2$

Answer

neither

76. $2x+4y=3; 6x+3y=2$

77. $y=2; y=6$

Answer

parallel

78. $y=-1; y=2$

Writing Exercises

79. How does the graph of a line with slope $m=12$ differ from the graph of a line with slope $m=2$?

Answer

Answers will vary.

80. Why is the slope of a vertical line “undefined”?

81. Explain how you can graph a line given a point and its slope.

Answer

Answers will vary.

82. Explain in your own words how to decide which method to use to graph a line.

Self Check

a. After completing the exercises, use this checklist to evaluate your mastery of the objectives of this section.

I can...	Confidently	With some help	No-I don't get it!
find the slope of a line.			
graph a line given a point and the slope.			
graph a line using its slope and intercept.			
choose the most convenient method to graph a line.			
graph and interpret applications of slope-intercept.			
use slopes to identify parallel and perpendicular lines.			

b. After reviewing this checklist, what will you do to become confident for all objectives?

This page titled 2.2.1E: Exercises is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **3.3E: Exercises** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/intermediate-algebra-2e.

2.2.2: Find the Equation of a Line

Learning Objectives

By the end of this section, you will be able to:

- Find an equation of the line given the slope and y-intercept
- Find an equation of the line given the slope and a point
- Find an equation of the line given two points
- Find an equation of a line parallel to a given line
- Find an equation of a line perpendicular to a given line

Before you get started, take this readiness quiz.

1. Solve: $\frac{2}{5}(x+15)$.
   If you missed this problem, review [link].
2. Simplify: $-3(x-(-2))$.
   If you missed this problem, review [link].
3. Solve for y: $y-3=-2(x+1)$.
   If you missed this problem, review [link].

How do online companies know that “you may also like” a particular item based on something you just ordered? How can economists know how a rise in the minimum wage will affect the unemployment rate? How do medical researchers create drugs to target cancer cells? How can traffic engineers predict the effect on your commuting time of an increase or decrease in gas prices? It’s all mathematics.

The physical sciences, social sciences, and the business world are full of situations that can be modeled with linear equations relating two variables. To create a mathematical model of a linear relation between two variables, we must be able to find the equation of the line. In this section, we will look at several ways to write the equation of a line. The specific method we use will be determined by what information we are given.

Find an Equation of the Line Given the Slope and *y*-Intercept

We can easily determine the slope and intercept of a line if the equation is written in slope-intercept form, $y=mx+b$. Now we will do the reverse—we will start with the slope and *y*-intercept and use them to find the equation of the line.

Example 2.2.2.1

Find the equation of a line with slope -9 and *y*-intercept $(0,-4)$.

Solution

Since we are given the slope and *y*-intercept of the line, we can substitute the needed values into the slope-intercept form, $y=mx+b$.

Name the slope.	$m=-9$
Name the *y*-intercept.	y=intercept is $(0,-4)$
Substitute the values into $y=mx+b$.	$y=mx+b$
	$y=-9x+(-4)$
	$y=-9x-4$

Try It 2.2.2.2

Find the equation of a line with slope $\frac{2}{5}$ and y-intercept $(0, 4)$.

Answer

$y = \frac{2}{5}x + 4$

Try It 2.2.2.3

Find the equation of a line with slope -1 and y-intercept $(0, -3)$.

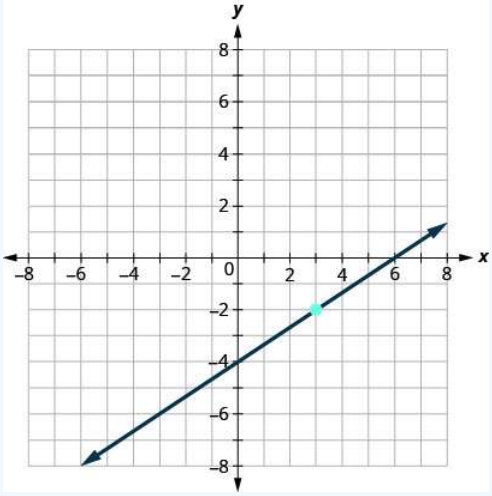
Answer

$y = -x - 3$

Sometimes, the slope and intercept need to be determined from the graph.

Example 2.2.2.4

Find the equation of the line shown in the graph.



Solution

We need to find the slope and y-intercept of the line from the graph so we can substitute the needed values into the slope-intercept form, $y = mx + b$.

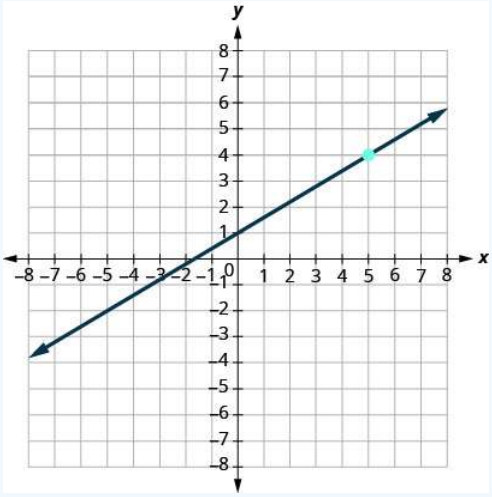
To find the slope, we choose two points on the graph.

The y-intercept is $(0, -4)$ and the graph passes through $(3, -2)$.

Find the slope, by counting the rise and run.	$m = \frac{\text{rise}}{\text{run}}$
	$m = \frac{2}{3}$
Find the y-intercept.	y=intercept is $(0, -4)$
Substitute the values into $y = mx + b$.	$y = mx + b$
	$y = \frac{2}{3}x - 4$

Try It 2.2.2.5

Find the equation of the line shown in the graph.

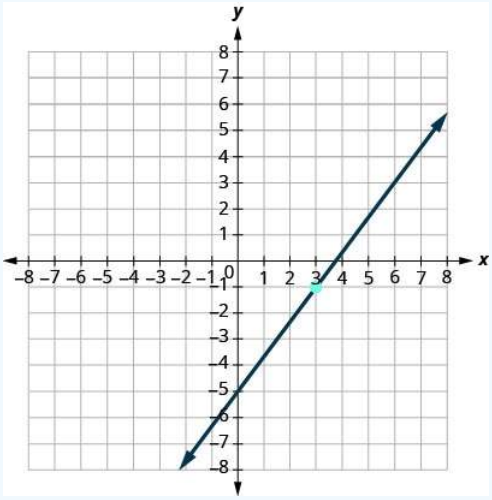


Answer

$y = \frac{3}{5}x + 1$

Try It 2.2.2.6

Find the equation of the line shown in the graph.



Answer

$y = \frac{4}{3}x - 5$

Find an Equation of the Line Given the Slope and a Point

Finding an equation of a line using the slope-intercept form of the equation works well when you are given the slope and *y*-intercept or when you read them off a graph. But what happens when you have another point instead of the *y*-intercept?

We are going to use the slope formula to derive another form of an equation of the line.

Suppose we have a line that has slope m and that contains some specific point (x_1, y_1) and some other point, which we will just call (x, y). We can write the slope of this line and then change it to a different form.

	m	$=$	$\frac{y-y_1}{x-x_1}$
Multiply both sides of the equation by $x-x_1$.	$m(x-x_1)$	$=$	$\left(\frac{y-y_1}{x-x_1}\right)(x-x_1)$
Simplify.	$m(x-x_1)$	$=$	$y-y_1$
Rewrite the equation with theyterms on the left.	$y-y_1$	$=$	$m(x-x_1)$

This format is called the **point-slope form** of an equation of a line.

POINT-SLOPE FORM OF AN EQUATION OF A LINE

The **point-slope form** of an equation of a line with slope m and containing the point (x_1, y_1) is:

$$y-y_1=m(x-x_1)$$

We can use the point-slope form of an equation to find an equation of a line when we know the slope and at least one point. Then, we will rewrite the equation in slope-intercept form. Most applications of linear equations use the the slope-intercept form.

How to Find an Equation of a Line Given a Point and the Slope

Example 2.2.2.7

Find an equation of a line with slope $m=-\frac{1}{3}$ that contains the point $(6,-4)$. Write the equation in slope-intercept form.

Solution

Step 1. Identify the slope.	The slope is given.	$m=-\frac{1}{3}$
Step 2. Identify the point.	The point is given.	$\begin{pmatrix} x_1, & y_1 \\ 6 & -4 \end{pmatrix}$
Step 3. Substitute the values into the point-slope form, $y-y_1=m(x-x_1)$	Simplify.	$y-y_1=m(x-x_1)$ $y-(-4)=-\frac{1}{3}(x-6)$ $y+4=-\frac{1}{3}x+2$
Step 4. Write the equation in slope intercept form.		$y=-\frac{1}{3}x-2$

Try It 2.2.2.8

Find the equation of a line with slope $m=-\frac{2}{5}$ and containing the point $(10,-5)$.

Answer

$y=-\frac{2}{5}x-1$

Try It 2.2.2.9

Find the equation of a line with slope $m=-\frac{3}{4}$, and containing the point $(4,-7)$.

Answer

$y=-\frac{3}{4}x-4$

We list the steps for easy reference.

TO FIND AN EQUATION OF A LINE GIVEN THE SLOPE AND A POINT.

1. Identify the slope.
2. Identify the point.
3. Substitute the values into the point-slope form, $y-y_1=m(x-x_1)$.
4. Write the equation in slope-intercept form.

Example 2.2.2.10

Find an equation of a horizontal line that contains the point $(-2,-6)$. Write the equation in slope-intercept form.

Solution

Every horizontal line has slope 0. We can substitute the slope and points into the point-slope form, $y-y_1=m(x-x_1)$.

Identify the slope.	$m=0$
Identify the point.	$\begin{pmatrix} x_1, & y_1 \\ -2 & -6 \end{pmatrix}$
Substitute the values into $y-y_1=m(x-x_1)$. Simplify.	$y-y_1=m(x-x_1)$ $y-(-6)=0(x--2)$ $y+6=0$
Write in slope-intercept form.	It is in y-form, but could be written $y=-6$.

Did we end up with the form of a horizontal line, $y=a$?

Try It 2.2.2.11

Find the equation of a horizontal line containing the point $(-3,8)$.

Answer

$y=8$

Try It 2.2.2.12

Find the equation of a horizontal line containing the point $(-1,4)$.

Answer

$y=4$

Find an Equation of the Line Given Two Points

When real-world data is collected, a linear model can be created from two data points. In the next example we'll see how to find an equation of a line when just two points are given.

So far, we have two options for finding an equation of a line: slope-intercept or point-slope. When we start with two points, it makes more sense to use the point-slope form.

But then we need the slope. Can we find the slope with just two points? Yes. Then, once we have the slope, we can use it and one of the given points to find the equation.

How to Find the Equation of a Line Given Two Points

Example 2.2.2.13

Find an equation of a line that contains the points $(-3, -1)$ and $(2, -2)$ Write the equation in slope-intercept form.

Solution

Step 1. Find the slope using the given points.	Find the slope of the line through $(-3. -1)$ and $(2, -2)$.	$m = \frac{y_2 - y_1}{x_2 - x_1}$ $m = \frac{-2-(-1)}{2-(-3)}$ $m = \frac{-1}{5}$ $m = -\frac{1}{5}$
Step 2. Choose one point.	Choose either point.	$\begin{pmatrix} x_1, & y_1 \\ 2 & -2 \end{pmatrix}$
Step 3. Substitute the values into the point-slope form, $y - y_1 = m(x - x_1)$.	Simplify.	$y - y_1 = m(x - x_1)$ $y - (-2) = -\frac{1}{5}(x - 2)$ $y + 2 = -\frac{1}{5}x + \frac{2}{5}$
Step 4. Write the equation in slope-intercept form.		$y = -\frac{1}{5}x - \frac{8}{5}$

Try It 2.2.2.14

Find the equation of a line containing the points $(-2, -4)$ and $(1, -3)$.

Answer

$y = \frac{1}{3}x - \frac{10}{3}$

Try It 2.2.2.15

Find the equation of a line containing the points $(-4, -3)$ and $(1, -5)$.

Answer

$y = -\frac{2}{5}x - \frac{23}{5}$

The steps are summarized here.

TO FIND AN EQUATION OF A LINE GIVEN TWO POINTS.

1. Find the slope using the given points. $m = \frac{y_2 - y_1}{x_2 - x_1}$
2. Choose one point.
3. Substitute the values into the point-slope form: $y - y_1 = m(x - x_1)$.
4. Write the equation in slope-intercept form.

Example 2.2.2.16

Find an equation of a line that contains the points $(-3, 5)$ and $(-3, 4)$. Write the equation in slope-intercept form.

Solution

Again, the first step will be to find the slope.

Find the slope of the line through$(-3, 5)$and$(-3, 4)$.

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

$$m = \frac{4 - 5}{-3 - (-3)}$$

$$m = \frac{-1}{0}$$

The slope is undefined.

This tells us it is a vertical line. Both of our points have an *x*-coordinate of -3. So our equation of the line is $x = -3$. Since there is no *y*, we cannot write it in slope-intercept form.

You may want to sketch a graph using the two given points. Does your graph agree with our conclusion that this is a vertical line?

Try It 2.2.2.17

Find the equation of a line containing the points $(5, 1)$ and $(5, -4)$.

Answer

$x = 5$

Try It 2.2.2.18

Find the equation of a line containing the points $(-4, 4)$ and $(-4, 3)$.

Answer

$x = -4$

We have seen that we can use either the slope-intercept form or the point-slope form to find an equation of a line. Which form we use will depend on the information we are given.

To Write an Equation of a Line

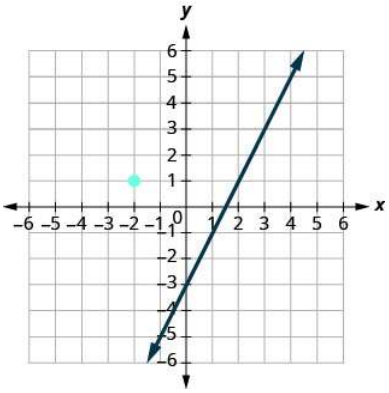
If given:	Use:	Form:
Slope and *y*-intercept	slope-intercept	$y = mx + b$
Slope and a point	point-slope	$y - y_1 = m(x - x_1)$
Two points	point-slope	$y - y_1 = m(x - x_1)$

Find an Equation of a Line Parallel to a Given Line

Suppose we need to find an equation of a line that passes through a specific point and is parallel to a given line. We can use the fact that parallel lines have the same slope. So we will have a point and the slope—just what we need to use the point-slope equation.

First, let's look at this graphically.

This graph shows $y = 2x - 3$. We want to graph a line parallel to this line and passing through the point $(-2, 1)$.

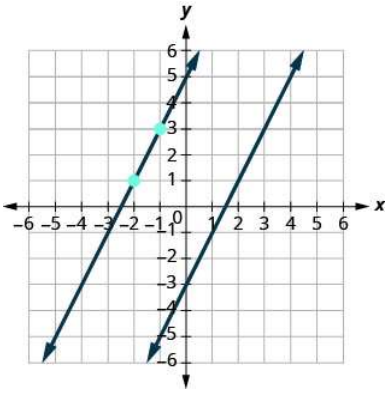


We know that parallel lines have the same slope. So the second line will have the same slope as $y = 2x - 3$. That slope is $m_{\parallel} = 2$. We'll use the notation $m_{\parallel}$ to represent the slope of a line parallel to a line with slope m. (Notice that the subscript $\parallel$ looks like two parallel lines.)

The second line will pass through $(-2, 1)$ and have $m = 2$.

To graph the line, we start at $(-2, 1)$ and count out the rise and run.

With $m = 2$ (or $m = \frac{2}{1}$), we count out the rise 2 and the run 1. We draw the line, as shown in the graph.



Do the lines appear parallel? Does the second line pass through$(-2, 1)$?

We were asked to graph the line, now let's see how to do this algebraically.

We can use either the slope-intercept form or the point-slope form to find an equation of a line. Here we know one point and can find the slope. So we will use the point-slope form.

How to Find the Equation of a Line Parallel to a Given Line and a Point

Example 2.2.2.19

Find an equation of a line parallel to $y = 2x - 3$ that contains the point $(-2, 1)$. Write the equation in slope-intercept form.

Solution

Step 1. Find the slope of the given line.	The given line is in slope-intercept form, $y = 2x - 3$	$m = 2$
Step 2. Find the slope of the parallel line.	Parallel lines have the same slope.	$m_{\parallel} = 2$
Step 3. Identify the point.	The given point is $(-2, 1)$.	$\begin{pmatrix} x_1, & y_1 \\ -2 & 1 \end{pmatrix}$

Step 4. Substitute the values into the point-slope form, $y-y_1=m(x-x_1)$.	Simplify.	$y-y_1=m(x-x_1)$ $y-(1)=2(x-(-2))$ $y-1=2(x+2)$ $y-1=2x+4$
Step 5. Write the equation in slope-intercept form.		$y=2x+5$

Look at graph with the parallel lines shown previously. Does this equation make sense? What is the *y*-intercept of the line? What is the slope?

? Try It 2.2.2.20

Find an equation of a line parallel to the line $y=3x+1$ that contains the point $(4,2)$. Write the equation in slope-intercept form.

Answer

$y=3x-10$

? Try It 2.2.2.21

Find an equation of a line parallel to the line $y=12x-3$ that contains the point $(6,4)$.

Write the equation in slope-intercept form.

Answer

$y=\frac{1}{2}x+1$

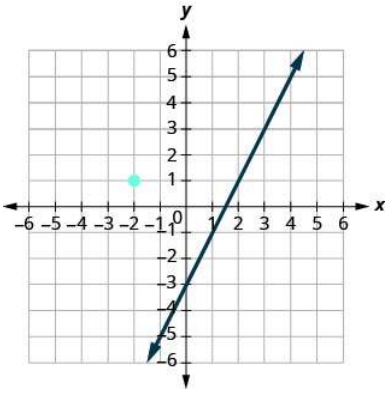
FIND AN EQUATION OF A LINE PARALLEL TO A GIVEN LINE.

1. Find the slope of the given line.
2. Find the slope of the parallel line.
3. Identify the point.
4. Substitute the values into the point-slope form: $y-y_1=m(x-x_1)$.
5. Write the equation in slope-intercept form.

Find an Equation of a Line Perpendicular to a Given Line

Now, let's consider perpendicular lines. Suppose we need to find a line passing through a specific point and which is perpendicular to a given line. We can use the fact that perpendicular lines have slopes that are negative reciprocals. We will again use the point-slope equation, like we did with parallel lines.

This graph shows $y=2x-3$. Now, we want to graph a line perpendicular to this line and passing through $(-2,1)$.



We know that perpendicular lines have slopes that are negative reciprocals.

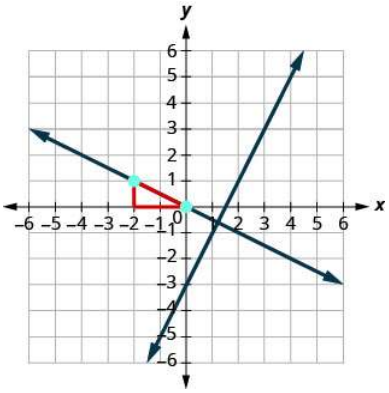
We'll use the notation $m_{\perp}$ to represent the slope of a line perpendicular to a line with slope m. (Notice that the subscript $\perp$ looks like the right angles made by two perpendicular lines.)

$$y = 2x - 3 \quad \text{perpendicular line}$$

$$m = 2 \quad m_{\perp} = -\frac{1}{2}$$

We now know the perpendicular line will pass through $(-2, 1)$ with $m_{\perp} = -\frac{1}{2}$.

To graph the line, we will start at $(-2, 1)$ and count out the rise -1 and the run 2. Then we draw the line.



Do the lines appear perpendicular? Does the second line pass through$(-2, 1)$?

We were asked to graph the line, now, let's see how to do this algebraically.

We can use either the slope-intercept form or the point-slope form to find an equation of a line. In this example we know one point, and can find the slope, so we will use the point-slope form.

How to Find the Equation of a Line Perpendicular to a Given Line and a Point

Example 2.2.2.22

Find an equation of a line perpendicular to $y = 2x - 3$ that contains the point $(-2, 1)$. Write the equation in slope-intercept form.

Solution

Step 1. Find the slope of the given line.	The given line is in slope-intercept form, $y = 2x - 3$	$m = 2$
Step 2. Find the slope of the perpendicular line.	The slopes of perpendicular lines are negative reciprocals.	$m_{\perp} = -\frac{1}{2}$

Step 3. Identify the point.	The given point is $(-2, 1)$.	$\begin{pmatrix} x_1, & y_1 \\ -2 & 1 \end{pmatrix}$
Step 4. Substitute the values into the point-slope form, $y-y_1=m(x-x_1)$.	Simplify.	$y-y_1=m(x-x_1)$ $y-(1)=-\frac{1}{2}(x-(-2))$ $y-1=-\frac{1}{2}(x+2)$ $y-1=-\frac{1}{2}x-1$
Step 5. Write the equation in slope-intercept form.		$y=-\frac{1}{2}x$

Try It 2.2.2.23

Find an equation of a line perpendicular to the line $y=3x+1$ that contains the point $(4, 2)$. Write the equation in slope-intercept form.

Answer

$y=-\frac{1}{3}x+\frac{10}{3}$

Try It 2.2.2.24

Find an equation of a line perpendicular to the line $y=12x-3$ that contains the point $(6, 4)$. Write the equation in slope-intercept form.

Answer

$y=-2x+16$

FIND AN EQUATION OF A LINE PERPENDICULAR TO A GIVEN LINE.

1. Find the slope of the given line.
2. Find the slope of the perpendicular line.
3. Identify the point.
4. Substitute the values into the point-slope form, $y-y_1=m(x-x_1)$.
5. Write the equation in slope-intercept form.

Example 2.2.2.24

Find an equation of a line perpendicular to $x=5$ that contains the point $(3, -2)$. Write the equation in slope-intercept form.

Solution

Again, since we know one point, the point-slope option seems more promising than the slope-intercept option. We need the slope to use this form, and we know the new line will be perpendicular to $x=5$. This line is vertical, so its perpendicular will be horizontal. This tells us the $m_{\perp}=0$.

Step 1. Identify the point.	The given point is $(3, -2)$.	$\begin{pmatrix} x_1, & y_1 \\ 3 & -2 \end{pmatrix}$
Step 2. Identify the slope of the perpendicular line.		$m_{\perp}=0$

Step 3. Substitute the values into the point-slope form, $y-y_1=m(x-x_1)$.	Simplify.	$y-y_1=m(x-x_1)$ $y-(-2)=0(x-(3))$
Step 4. Simplify		$y+2=0$ $y=-2$

Sketch the graph of both lines. On your graph, do the lines appear to be perpendicular?

Try It 2.2.2.25

Find an equation of a line that is perpendicular to the line $x=4$ that contains the point $(4,-5)$. Write the equation in slope-intercept form.

Answer

$y=-5$

Try It 2.2.2.26

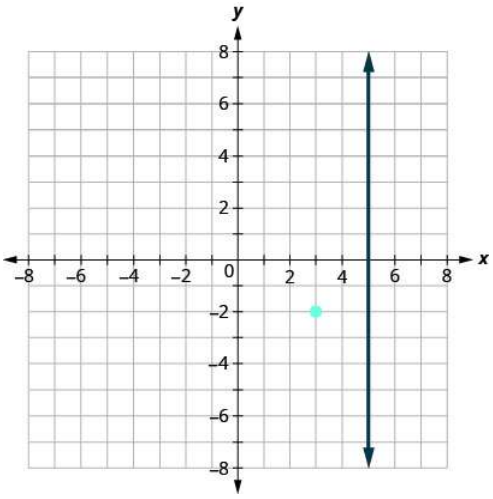
Find an equation of a line that is perpendicular to the line $x=2$ that contains the point $(2,-1)$. Write the equation in slope-intercept form.

Answer

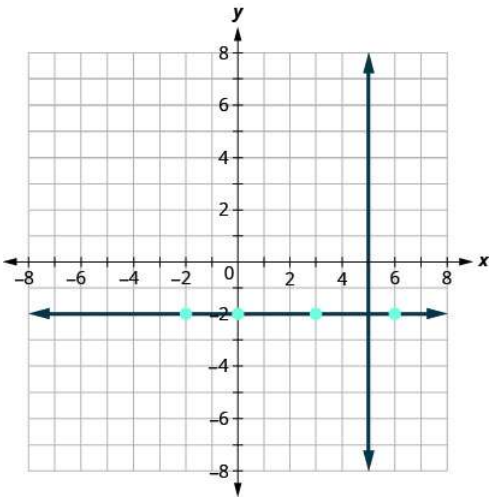
$y=-1$

In the Example above, we used the point-slope form to find the equation. We could have looked at this in a different way.

We want to find a line that is perpendicular to $x=5$ that contains the point $(3,-2)$. This graph shows us the line $x=5$ and the point $(3,-2)$.



We know every line perpendicular to a vertical line is horizontal, so we will sketch the horizontal line through $(3,-2)$.



Do the lines appear perpendicular?

If we look at a few points on this horizontal line, we notice they all have *y*-coordinates of -2. So, the equation of the line perpendicular to the vertical line $x = 5$ is $y = -2$.

Example 2.2.2.27

Find an equation of a line that is perpendicular to $y = -3$ that contains the point $(-3, 5)$. Write the equation in slope-intercept form.

Solution

The line $y = -3$ is a horizontal line. Any line perpendicular to it must be vertical, in the form $x = a$. Since the perpendicular line is vertical and passes through $(-3, 5)$, every point on it has an *x*-coordinate of -3. The equation of the perpendicular line is $x = -3$.

You may want to sketch the lines. Do they appear perpendicular?

Try It 2.2.2.28

Find an equation of a line that is perpendicular to the line $y = 1$ that contains the point $(-5, 1)$. Write the equation in slope-intercept form.

Answer

$x = -5$

Try It 2.2.2.29

Find an equation of a line that is perpendicular to the line $y = -5$ that contains the point $(-4, -5)$. Write the equation in slope-intercept form.

Answer

$x = -4$

Access these online resources for additional instruction and practice with finding the equation of a line.

- Write an Equation of Line Given its slope and Y-Intercept
- Using Point Slope Form to Write the Equation of a Line, Find the equation given slope and point
- Find the equation given two points
- Find the equation of perpendicular and parallel lines

Key Concepts

- **How to find an equation of a line given the slope and a point.**
  1. Identify the slope.
  2. Identify the point.
  3. Substitute the values into the point-slope form,$y-y_1=m(x-x_1)$.
  4. Write the equation in slope-intercept form.

- **How to find an equation of a line given two points.**
  1. Find the slope using the given points. $m=\frac{y_2-y_1}{x_2-x_1}$
  2. Choose one point.
  3. Substitute the values into the point-slope form: $y-y_1=m(x-x_1)$.
  4. Write the equation in slope-intercept form.

To Write an Equation of a Line		
If given:	**Use:**	**Form:**
Slope and *y*-intercept	**slope-intercept**	$y=mx+b$
Slope and a point	**point-slope**	$y-y_1=m(x-x_1)$
Two points	**point-slope**	$y-y_1=m(x-x_1)$

- **How to find an equation of a line parallel to a given line.**
  1. Find the slope of the given line.
  2. Find the slope of the parallel line.
  3. Identify the point.
  4. Substitute the values into the point-slope form: $y-y_1=m(x-x_1)$.
  5. Write the equation in slope-intercept form
- **How to find an equation of a line perpendicular to a given line.**
  1. Find the slope of the given line.
  2. Find the slope of the perpendicular line.
  3. Identify the point.
  4. Substitute the values into the point-slope form, $y-y_1=m(x-x_1)$.
  5. Write the equation in slope-intercept form.

Glossary

point-slope form

The point-slope form of an equation of a line with slope *m* and containing the point (x_1,y_1) is $y-y_1=m(x-x_1)$.

This page titled 2.2.2: Find the Equation of a Line is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **3.4: Find the Equation of a Line** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/intermediate-algebra-2e.

2.3: Linear Models

2.3: Linear Models is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.3.1: Linear Models

Sebastian waves good-bye to his brother, who is talking to a group of his friends approximately 20 feet away. Sebastian then begins to walk away from his brother at a constant rate of 4 feet per second. Let's model the distance separating the two brothers as a function of time.

Our first approach will be graphical. We will let the variable d represent the distance (in feet) between the brothers and the variable t represent the amount of time (in seconds) that has passed since Sebastian waved good-bye to his brother. Because the distance separating the brothers depends on the amount of time that has passed, we will say that the distance d is the dependent variable and the time t is the independent variable.

It is somewhat traditional in the modeling process to place the independent variable on the horizontal axis and the dependent variable on the vertical axis. This is not a hard and fast rule, more a matter of personal taste, but we will follow this rule in our example nonetheless. Thus, we will place distance on the vertical axis and time on the horizontal axis, as shown in Figure 2.3.1.1. Notice that we've labeled each axis with its variable representation and included the units, an important practice.

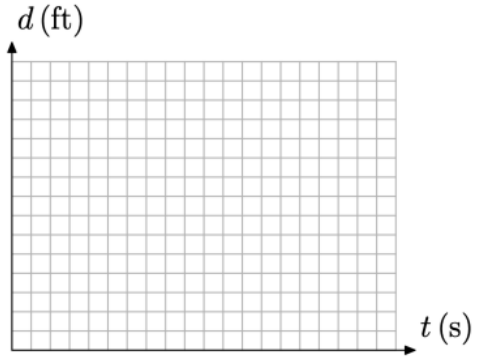


Figure 2.3.1.1 Distance depends upon time.

Warning

The label on the horizontal axis, t(s), might look like function notation to some readers. This is not the case. Rather, the variable t represents time, and the (s) in parentheses that follows represents seconds, a standard abbreviation in physics. Similar comments are in order for the label d (ft). The variable d represents distance, and the (ft) in parentheses that follows represent feet, another standard abbreviation in physics.

There are a number of different ways that you can label the axes of your graph with units appropriate for the problem at hand. For example, consider the technique presented in Figure 2.3.1.2 where the labels are placed to the left of the vertical axis and underneath the horizontal axis. Another difference is the fact that the unit abbreviations in Figure 1 are spelled out in their entirety in Figure 2.3.1.2

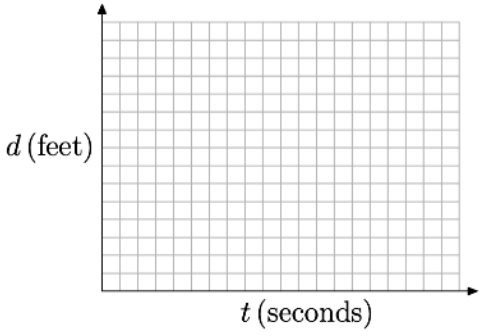


Figure 2.3.1.2 Distance depends upon time.

Some instructors prefer that you rotate the distance label on the vertical axis ninety degrees, so that it appears sideways. Others prefer that you label the ends of each axis with the variable, as we have done in Figure 2.3.1.1, but spell out the units in their

entirety alongside each axis as we've done in Figure 2.3.1.2 The list of preferences goes on and on.

Note

It is important to have a conversation with your instructor in order to determine what your instructor's expectations are when it comes to labeling the axes and indicating the units on your graphs.

We prefer to label the axes as shown in Figure 2.3.1.1, and we will try to be consistent to this standard throughout the remainder of the text, though we might stray to alternate forms of labeling from time to time

We must now scale each axis appropriately, a task that is harder than it first seems. A poor choice of scale can make the task ahead more difficult than it needs to be. We will choose a scale for each axis with the following thoughts in mind.

Guidelines for Scaling Axes

Here is some good advice to follow when scaling the dependent and independent axes.

1. We want to avoid postage-stamp-sized graphs. A large graph is easier to interpret than one that is cramped in a small corner of our graph paper.
2. It is not necessary to have the same scale on each axis, but once a scale is chosen, you must remain consistent.
3. We want to choose a scale that correlates easily with the given rate.

Sebastian is walking away from his brother at a constant rate of 4 feet per second. Let's let each box on the vertical axis represent 4 feet and every two boxes on the horizontal axis represent 1 second, as shown in Figure 2.3.1.3

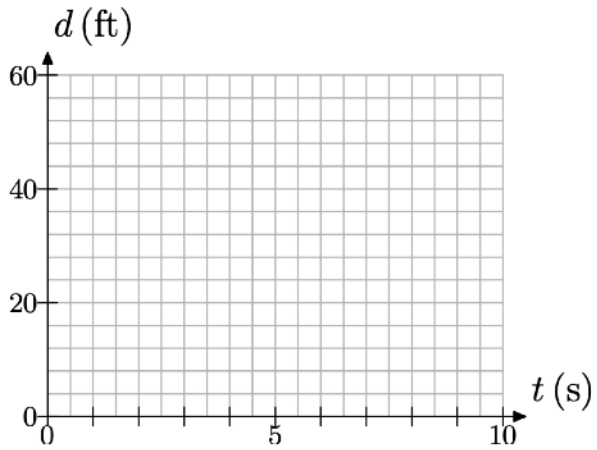


Figure 2.3.1.3 Scaling each axis to accommodate the rate.

At time t = 0, Sebastian is separated from his brother by a distance of d = 20 feet. This corresponds to the point (t, d) = (0, 20) shown in Figure 2.3.1.4(a).

Next, Sebastian walks away from his brother at a constant rate of 4 feet per second. This means that for every second of time that elapses, the distance between the brothers increases by 4 feet. Starting at the point (0, 20), move 1 second (two boxes) to the right and 4 feet (1 box) upward to the point (1, 24), as shown in Figure 2.3.1.4(b).

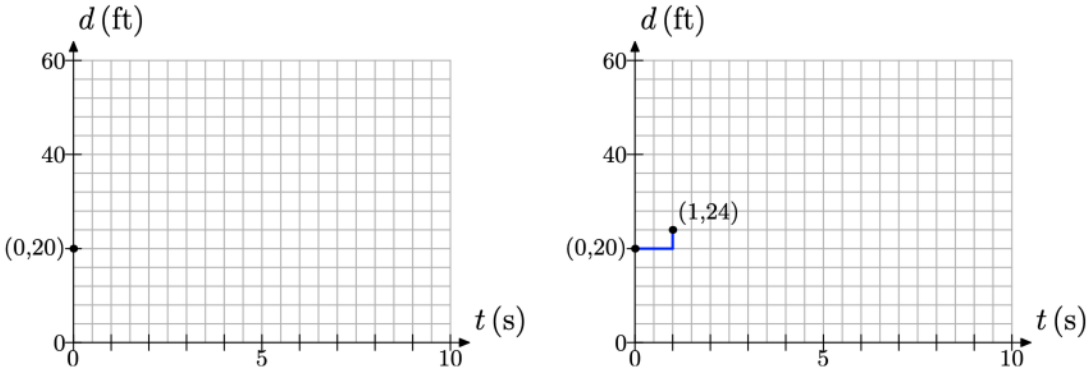


Figure 2.3.1.4

The rate of separation is a constant 4 feet per second. So, continue indefinitely in the manner of Figure Figure 2.3.1.4(b), moving 1 second (2 boxes) to the right, then 4 feet upward (1 box). This will produce the linear relationship between distance and time suggested in Figure 2.3.1.5(a).

If we assume that the distance is a continuous function of time, a legitimate assumption due to the fact that the distance is increasing continuously at a constant rate of 4 feet per second, then we can replace the discrete set of data points in Figure 2.3.1.5(a) with the line shown in Figure 2.3.1.5(b).

The line in Figure 2.3.1.5(b) is a continuous model. It can be drawn with a simple stroke of the pencil, without the tip of the pencil ever leaving contact with our graph paper. On the other hand, the set of points in Figure 2.3.1.4(a) is a discrete model. After plotting a point, our pencil must break contact with our graph paper before plotting the next point. This is the essential difference between a discrete model and a continuous model.

In this case, the continuous model is a more accurate representation of the distance between the brothers. We say this because the distance between them is increasing at a constant rate of 4 feet per second, or 2 feet every half second, or 1 foot every quarter second, etc. Shortly, we will show an example where this sort of continuous model is unreasonable.

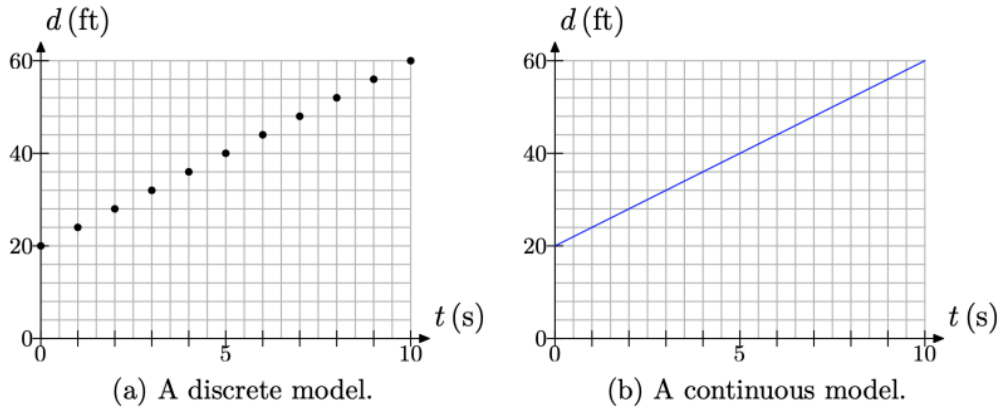


Figure 2.3.1.5 Constant rate yields a linear relationship.

Now that we've modeled the distance between the brothers with a graph, we can use the graph to make predictions. For example, to determine the distance between the brothers after 8 seconds, locate 8 seconds on the time axis, draw a vertical arrow to the line, then a horizontal arrow to the distance axis, as shown in Figure 2.3.1.6

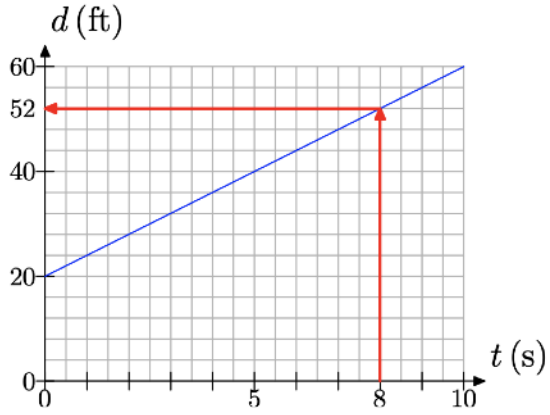


Figure 2.3.1.6 Predicting the distance between the brothers after 8 seconds.

However, suppose that we want to determine the distance between the brothers after 2 minutes. As the graph in Figure 2.3.1.5(b) only models the distance over the first 10 seconds, we would have to redraw the graph over the first 2 minutes (120 seconds) to determine the answer. We did not plan ahead for this contingency, so perhaps we can model the distance between the brothers in another way, one that will more easily predict the distance between the brothers after an arbitrary amount of time t.

To this end, we search for a pattern that describes the distance d between the brothers as a function of time t. Because the distance between the brothers is increasing at a rate of 4 feet per second, we note that:

- At t = 0 seconds, the distance between the brothers is d = 20 feet.
- At t = 1 second, the distance between the brothers is d = 24 feet.
- At t = 2 seconds, the distance between the brothers is d = 28 feet.
- At t = 3 seconds, the distance between the brothers is d = 32 feet.

We summarize these results in Table 2.3.1.1(a).

However, you don't want to simplify the distances as we have in Table 2.3.1.1(a), because you hide the pattern or the relationship between the distance d and the time t. It is more efficient to seek a relationship between distance and time in the following manner. After t = 1 second, the distance increases by 1 increment of 4 feet, so d = 20 + 4(1). After t = 2 seconds, the distance increases by 2 increments of 4 feet, so d = 20 + 4(2). Continuing in this manner, we have:

• At t = 3 seconds, the distance between the brothers is d = 20 + 4(3) feet.

• At t = 4 seconds, the distance between the brothers is d = 20 + 4(4) feet.

These results are summarized in Table 2.3.1.1(b).

t	d
0	20
1	24
2	28
3	32

Table 2.3.1.1(a) Determining a model equation.

t	d
0	20
1	20 + 4(1)
2	20 + 4(2)
3	20 + 4(3)

Table 2.3.1.1(b) Determining a model equation.

Unlike Table 2.3.1.1(a), Table 2.3.1.1(b) reveals a relationship between distance d and time t that can be described by the equation

$$d = 20 + 4t$$

The careful reader will check that equation (3) reveals the correct distances for t = 0, 1, 2, and 3 seconds, as recorded in Table 2.3.1.1(a). Two important observations can be made about equation (3).

1. The 20 in d = 20 + 4t is the initial distance between the brothers and corresponds to the point (0, 20) in Figure 2.3.1.4(a).
2. The 4 in N = 20 + 4t is the rate at which the distance between the brothers is increasing (4 feet per second).

Moreover, equation (3) can be used to predict the distance between the brothers at 2 minutes. First, convert t = 2 minutes to t = 120 seconds, then substitute this number in our model equation (3).

$$d = 20 + 4(120) = 500$$

Thus, the distance between the brothers after 2 minutes is d = 500 feet.

We can also write the equation d = 20 + 4t using function notation.

$$d(t) = 20 + 4t$$

Then, to find the distance between the brothers at the end of 2 minutes, we would perform the following calculation.

$$d(120) = 20 + 4(120)$$
$$d(120) = 500$$

Unlike function notation, when the result is written d = 500 feet, note how one piece of information is hidden, namely the time. With function notation, we interpret d(120) = 500 to mean “the distance between the two brothers after 120 seconds is 500 feet.” Note how both the distance and the time are available in the notation d(120) = 500.

Modeling the Discrete with the Continuous

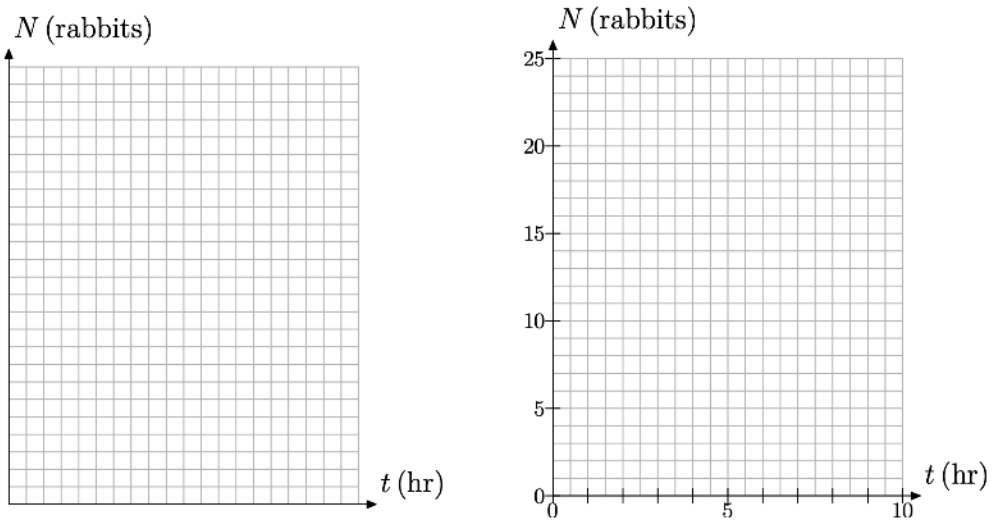
Jenny builds a rabbit hutch behind her barn. She places 25 rabbits in the hutch, then locks the door and leaves. Unfortunately, there is a flaw in the design of the hutch and the rabbits begin to escape at a constant rate of 5 rabbits every 2 hours. Again, we’ll model the number N of rabbits remaining in the hutch as a function of time t. First, we propose a graphical model.

Note that the number of rabbits remaining in the hutch depends on the amount of time that has passed. This makes the number N of rabbits remaining in the hutch the dependent variable, which we will place on the vertical axis in Figure 2.3.1.7(a). Time t is the independent variable and is placed on the horizontal axis.

We’ll again choose a scale for our axes that accommodates the fact that the rabbit population is decreasing at a constant rate of 5 rabbits every 2 hours. In Figure 2.3.1.7(b), we let each box on the vertical axis represent 1 rabbit, while two boxes on the horizontal axis represents 1 hour. We could just as easily let each box on the horizontal axis represent one hour, but our choice makes a graph that is a bit larger. Larger graphs are a bit easier to read and interpret.

At time t = 0 hours, the rabbit population is N = 25 rabbits. This fact is represented by the point (t, N) = (0, 25) in Figure 2.3.1.8(a). Because the rabbit population decreases at a constant rate of 5 rabbits every 2 hours, we start at the point (0, 25), then move 2 hours (4 boxes) to the right, and 5 rabbits (5 boxes) down to the point (2, 20), also shown in Figure 2.3.1.8(a).

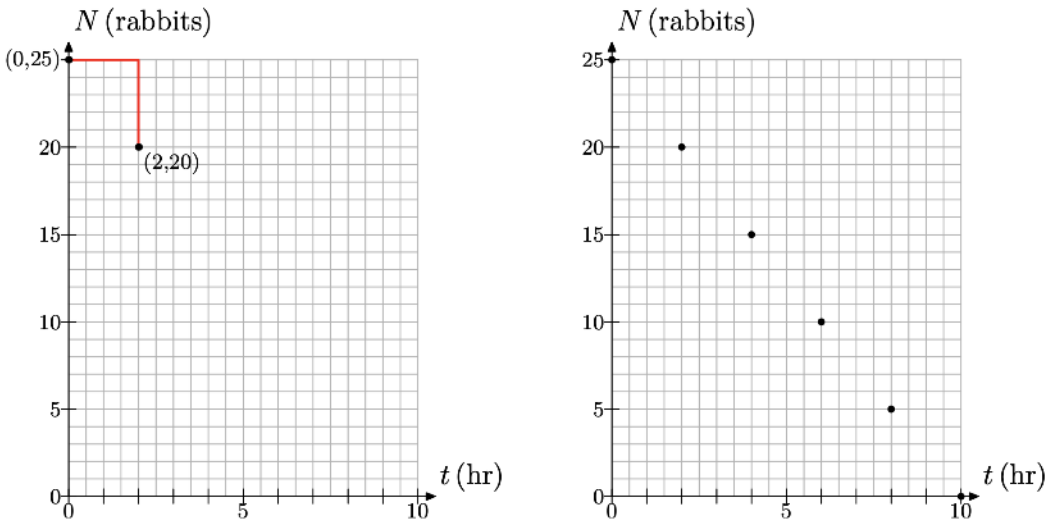
The rate at which the rabbits are decreasing is constant at 5 rabbits every 2 hours, so continue indefinitely in the manner of Figure 2.3.1.8(a), moving 2 hours (4 boxes) to the right, then 5 rabbits (5 boxes) downward. This will produce the linear relationship between the number of rabbits N and the time t shown in Figure 2.3.1.8(b).



(a) Dependent and independent variables.

(b) Scaling the axes to accommodate the rate.

Figure 2.3.1.7

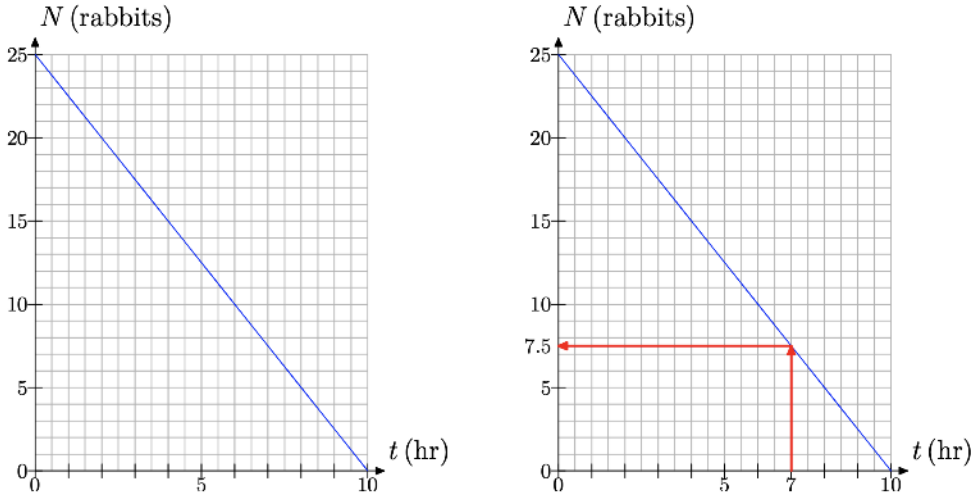


(a) Initial rabbit population is 25. Five rabbits escape every 2 hours.

(b) A discrete model.

Figure 2.3.1.8

We can draw a line through the data points in Figure 2.3.1.8(b) to produce the continuous model in Figure 2.3.1.9(a). However, we need to be aware of the shortcoming imposed by this continuous approximation. For example, consider the prediction in Figure 2.3.1.9(b). Is it reasonable to say that 7.5 rabbits remain in the hutch after 7 hours?



(a) A continuous model.

(b) Predicting the rabbit population after 7 hours.

Figure 2.3.1.9

In our first model, the distance between the brothers can be any real number, so a continuous model was appropriate. However, in the case of Jenny's rabbit hutch, the remaining population must be a whole number of rabbits (unless a fox gets in), so modeling the population with the continuous line in Figure 9(b) is at best an approximation of reality. However, mathematicians will frequently model a discrete situation with a continuous model. As long as we are aware of its limitations, we can still use the model to make reasonable predictions. For example, we might say that there are approximately 7 rabbits left in the hutch after 7 hours.

We saw the advantage of using function notation at the end of our previous model, so let's employ function notation a bit earlier in this model. We will let

$$N(t) = \text{ the number of rabbits remaining after } t \text{ hours.}$$

Initially, at time t = 0, there are 25 rabbits in the hutch. Thus, we write

$$N(0) = 25$$

It might be easier to think of losing 5 rabbits every 2 hours as being equivalent to losing "on average" 2.5 rabbits every hour. Thus, at the end of 1 hour, the number of rabbits decreases by one increment of 2.5 rabbits, and we write

$$N(1) = 25 - 2.5(1)$$

At the end of 2 hours, the rabbit population decreases by 2 increments of 2.5 rabbits and we can write

$$N(2) = 25 - 2.5(2)$$

At the end of 3 hours, the rabbit population decreases by 3 increments of 2.5 rabbits and we can write

$$N(3) = 25 - 2.5(3)$$

A clear pattern develops, particularly when we summarize these results in Table 2.3.1.2

t	N(t)
0	25
1	25 − 2.5(1)
2	25 − 2.5(2)
3	25 − 2.5(3)

Table 2.3.1.2Determining a model equation.

Table 2.3.1.2reveals a relationship between the number of rabbits N and time t that can be described by the equation

$$N(t) = 25 - 2.5t$$

$$N(t) = 25 - 2.5t$$

The careful reader will again check that equation (4) returns the correct number of rabbits at times t = 0, 1, 2, and 3, as recorded in Table 2.3.1.2

There are two important observations we can make about equation (4).

1. The 25 in $N(t) = 25 - 2.5t$ is the initial rabbit population and corresponds to the point (0, 25) in Figure 2.3.1.8(a).
2. The −2.5 in $N(t) = 25 - 2.5t$ is the rate at which the rabbit population is decreasing "on average" (2.5 rabbits per hour).

The equation (4) can be used to predict the number of rabbits remaining in the hutch after t = 7 hours. Simply substitute t = 7 in equation (4).

$$N(7) = 25 - 2.5(7) = 7.5$$

It is important to note that the prediction made by the model equation is identical to that made by the model graph in Figure 2.3.1.9(b).

However, again note that this equation is a continuous model, and its prediction that 7.5 rabbits remain in the hutch is not realistic (unless that fox got loose again). However, if we are aware of the model's shortcomings, the equation can still be used as a good predictive tool. For example, we might again say that approximately 7 rabbits remain in the hutch after 7 hours. This can be written $N(7) \approx 7$, which means that "after 7 hours, there are approximately 7 rabbits remaining in the hutch."

Determining the Equation Model from the Graph

Mrs. Burke sets up a motion detector at the front of her classroom, then positions one of her students a fixed distance from the detector and asks the student to approach the detector at a constant speed. The detector measures the distance d (in meters) of the student from the detector as a function of time t (in seconds). The graph of distance d versus time t is given in Figure 2.3.1.10

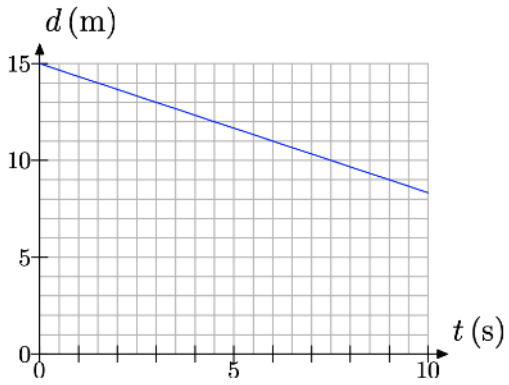
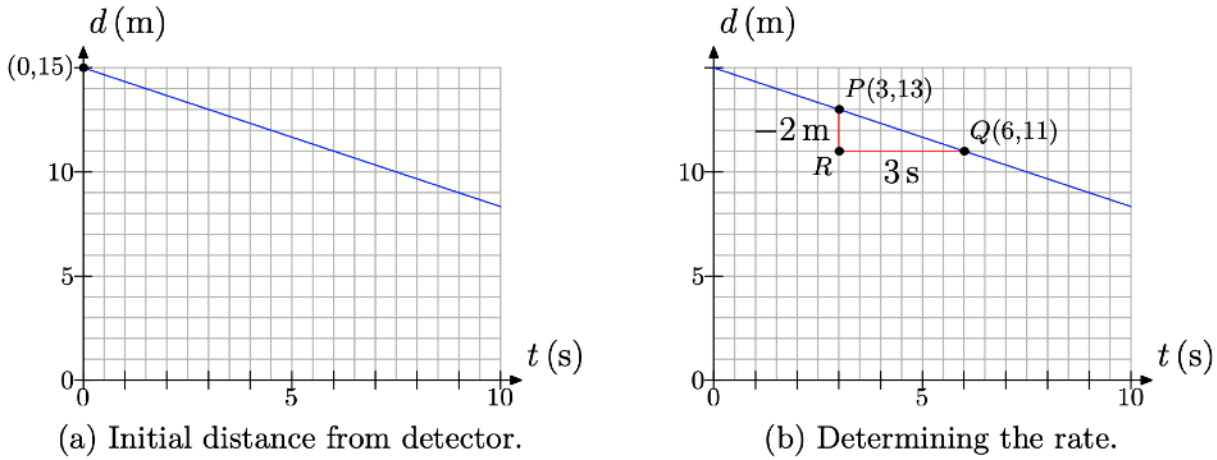


Figure 2.3.1.10 The detector measures the distance between student and detector versus time.

It is a simple matter to determine the student's initial distance from the detector. We need only determine the value of d at time t = 0 seconds. The result is located at the point (0, 15), as shown in Figure 2.3.1.11(a). Thus, the student sets up at an initial distance of 15 meters from the detector.

To determine the rate at which the student approaches the detector, we need to do a bit more work. Examine the graph and pick two points on the line. It makes things a bit easier if you pick points on the line that are situated at the intersection of two grid lines, but as we will show, this is not necessary. With this thought in mind, we've picked the points P(3, 13) and Q(6, 11) on the line, as shown in Figure 2.3.1.11(b).



(a) Initial distance from detector. (b) Determining the rate.

Figure 2.3.1.11. Determining initial distance and rate.

Draw a right triangle $\triangle PQR$ with sides parallel to the axes, as shown in Figure 2.3.1.11(b). Determine the length of each side of the right triangle.

- Side PR is 2 boxes in length, but each box represents 1 meter, so side P R represents a decrease of 2 meters in distance from the detector. That is why we've used the minus sign in labeling the side PR with −2m in Figure 2.3.1.11(b).
- Side RQ is 6 boxes in length, but 2 boxes represents 1 second, so side RQ represents an increase of 3 seconds in time. That is why we've labeled side RQ with 3 s in Figure 2.3.1.11(b).

Thus, the distance between the student and detector is decreasing at a rate of 2 meters every 3 seconds.

What would happen if we picked two different points on the line? Consider the case in Figure 2.3.1.12, where we've picked the points on the line at P(3, 13) and Q(9, 9). We've also decided to draw right triangle $\triangle PQR$ on the opposite side of the line. However, note again that the sides of the right triangle $\triangle PQR$ are parallel to the horizontal and vertical axes.

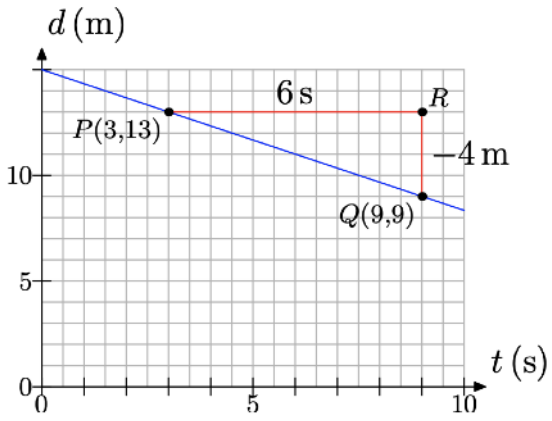


Figure 2.3.1.12 Determining the rate.

Determine the length of each side of triangle $\triangle PQR$.

- Side PR is 12 boxes in length, but 2 boxes represent 1 second, so side PR represents an increase of 6 seconds in time. That is why we've labeled side P R with 6 s in Figure 2.3.1.12
- Side RQ is 4 boxes in length, but each box represents 1 meter, so side RQ represents a decrease of 4 meters in distance from the detector. That is why we've used a minus sign in labeling the side RQ with −4 m in Figure 2.3.1.12

Thus, the distance between the student and detector is decreasing at a rate of 4 meters every 6 seconds. In symbols, we would write that the rate is

$$\text{Rate} = \frac{-4\text{m}}{6\text{s}} = -\frac{4}{6}\text{m/s}$$

Note, however, that this reduces to

$$\text{Rate } = -\frac{2}{3}\text{m/s}$$

which is identical to the rate found earlier when using the points P and Q in Figure 2.3.1.11(b).

The fact that these rates are equivalent is due to the fact that the triangles (\triangle P Q R\) in Figure 2.3.1.11(b) and Figure 2.3.1.12 are similar triangles, so their sides are proportional. Thus, it doesn't matter which two points you pick on the line, nor does it matter which side of the line you place your right triangle. Thus, the only requirement is that you draw a right triangle with sides parallel to the coordinate axes.

Finally, let's see if we can develop a model equation. We will define

$$d(t) = \text{ the distance from the detector at time } t$$

Initially, the student is 15 meters from the detector. That is, at time t = 0, the distance from the detector is 15 meters. In symbols, we write

$$d(0) = 15$$

The distance decreases at a rate of 2 meters every 3 seconds. This is equivalent to saying that the distance decreases 2/3 meters every second. At the end of 1 second, the distance has decreased by 1 increment of 2/3 meters, so the distance from the detector is given by

$$d(1) = 15 - \frac{2}{3}(1)$$

At the end of 2 seconds, the distance has decreased by 2 increments of 2/3 meters, so the distance from the detector is given by

$$d(2) = 15 - \frac{2}{3}(2)$$

At the end of 3 seconds, the distance has decreased by 3 increments of 2/3 meters, so the distance from the detector is given by

$$d(3) = 15 - \frac{2}{3}(3)$$

A clear pattern emerges, particularly if you summarize the results as we have in Table 2.3.1.3

t	d(t)
0	15
1	15 − (2/3)(1)
2	15 − (2/3)(2)
3	15 − (2/3)(3)

Table 2.3.1.3 Determining a model equation.

Table 2.3.1.3 reveals that the linear relationship (see Figure 2.3.1.10) between the distance d from the detector at time t can be modeled by the equation

$$d(t) = 15 - \frac{2}{3}t$$

Again, the careful reader will check that equation (5) returns the correct distance d at the times t = 0, 1, 2, and 3 recorded in Table 2.3.1.3

There are two important observations to be made about equation (5).

1. The 15 in d(t) = 15−(2/3)t is the initial distance from the detector and corresponds to the point (0, 15) in Figure 2.3.1.11(a).

2. The −2/3 in d(t) = 15−(2/3)t is the rate at which the distance between the student and detector is changing as determined in Figure 2.3.1.11(b). It is negative because the distance is decreasing with time.

Equation (5) can be used to make predictions. For example, to determine the distance between the student and detector at the end of 9 seconds, insert t = 9 into equation (5).

$$d(9) = 15 - \frac{2}{3}(9) = 15 - 6 = 9$$

Of course, the notation d(9) = 9 is interpreted to mean “the distance between the student and the detector after 9 seconds is 9 meters.

This page titled 2.3.1: Linear Models is shared under a CC BY-NC-SA 2.5 license and was authored, remixed, and/or curated by David Arnold via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: Linear Models** by David Arnold is licensed CC BY-NC-SA 2.5. Original source: https://web.archive.org/web/20200814023923/http://msenux2.redwoods.edu/IntAlgText/.

2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know)

The end of the month is approaching, and bills are coming due. As you sit at your kitchen table trying to figure out your budget for next month, you wonder whether you will be able to afford that concert you had been planning on attending. Some of your costs remain unchanged from month to month, such as your rent, Internet service, gym membership, and insurance. Other costs tend to fluctuate with your usage, such as your utilities, cellphone bill, vehicle fuel, and the amount of money spent on recreational activities. Together, these regular and irregular costs total to next month's costs.

Examining a few recent paycheque stubs, you calculate the average monthly net income you bring home from your hourly cashier position at Sobeys. The exact amount of each paycheque, of course, depends on how many hours you work. Besides your short-term costs, you need to start saving for next year's tuition. Therefore, your budget needs to include regular deposits into your savings account to meet that goal. Once you have put your bills, paycheques, and goals together, you hope that your budget will balance. If there is a shortfall, you will have to miss out on those concert tickets.

Budgeting at work is no different in principle from your home budget. Businesses also need to recognize the different types of costs they incur each month, some of which remain the same and some of which fluctuate. Businesses need to pay for these costs by generating revenues, which correspond to your paycheque. As with your education goals, businesses also require profits to grow. A business needs to understand all of these numbers so it can plan its activities realistically.

This section explores the various types of costs and establishes a model relating total costs to total revenues to determine total profitability levels. You will then apply this model to see how the sale of an individual product contributes to covering costs and how each product individually contributes to overall profitability.

Types of Costs

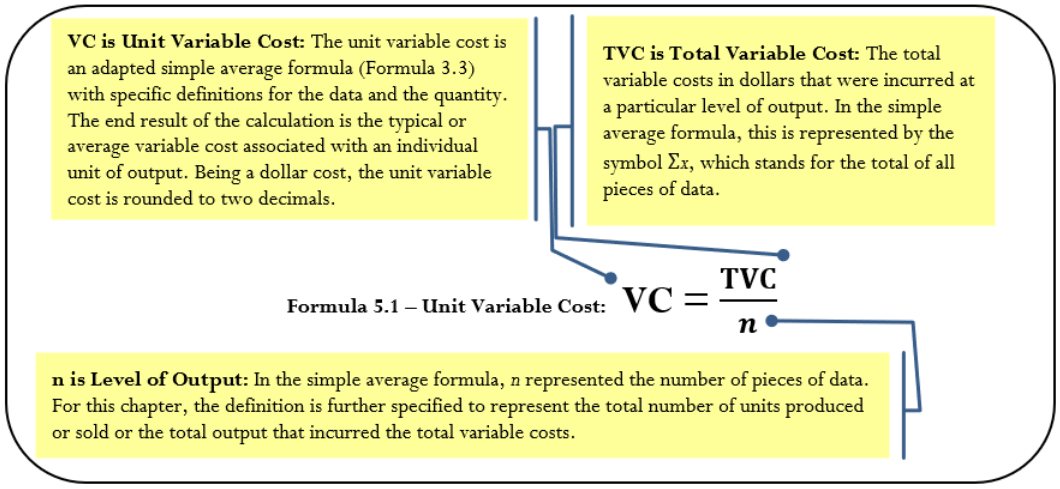
A **cost** is an outlay of money required to produce, acquire, or maintain a product, which includes both physical goods and services. Costs can come in three forms:

1. A **fixed cost** is a cost that does not change with the level of production or sales (call this "output" for short). In other words, whether the business outputs nothing or outputs 10,000 units, these costs remain the same. Some examples include rent, insurance, property taxes, salaries unrelated to production (such as management), production equipment, office furniture, and much more. **Total fixed costs** are the sum of all fixed costs that a business incurs.
2. A **variable cost** is a cost that changes with the level of output. In other words, if the business outputs nothing there is no variable cost. However, if the business outputs just one unit (or more) then a cost appears. Some examples include material costs of products, production labour (hourly or piecework wages), sales commissions, repairs, maintenance, and more. **Total variable costs** are the sum of all variable costs that a business incurs at a particular level of output.
3. A **blended cost** is a cost that comprises both fixed cost and variable cost components. In other words, a portion of the total cost remains unchanged while another portion depends on the output. For calculation purposes, you must separate a blended cost into its fixed and variable cost components. A few examples will illustrate the concept of blended costs:

- Residential natural gas bills from Manitoba Hydro include a fixed charge per month of $14 plus charges for cubic metres of actual consumption based on transportation, distribution, and primary and supplemental gas rates. In this situation, the $14 is a fixed cost while the actual consumption of natural gas is a variable cost.
- A cell phone bill includes a fixed charge for the phone service plus any additional charges for usage, such as long distance, texting, or data.
- If employees are paid a salary plus commission, then their salaries represent fixed costs while their commissions are a variable cost.

The Formula

In calculating business costs, fixed costs are commonly calculated on a total basis only since the business incurs these costs regardless of any production. However, variable costs are commonly calculated both on a total and per-unit basis to reveal the overall cost along with the cost associated with any particular unit of output. When these variable costs are assigned on an individual basis it is called a **unit variable cost**. The calculation of unit variable cost has a further benefit because it allows managers to explore how the total business costs vary at different levels of output.



Formula 5.1

How It Works

Follow these steps to calculate the unit variable cost:

- **Step 1**: Identify all fixed, variable, and blended costs, along with the level of output. For variable costs, understand any important elements of how the cost is structured. For blended costs, separate the costs into variable and fixed components.
- **Step 2**: Calculate the total variable cost (TVC) by totaling all variable costs at the indicated level of output. This involves taking any known unit variable costs and multiplying each by the level of output.
- **Step 3**: Divide the total variable cost by the total level of output by applying Formula 5.1.

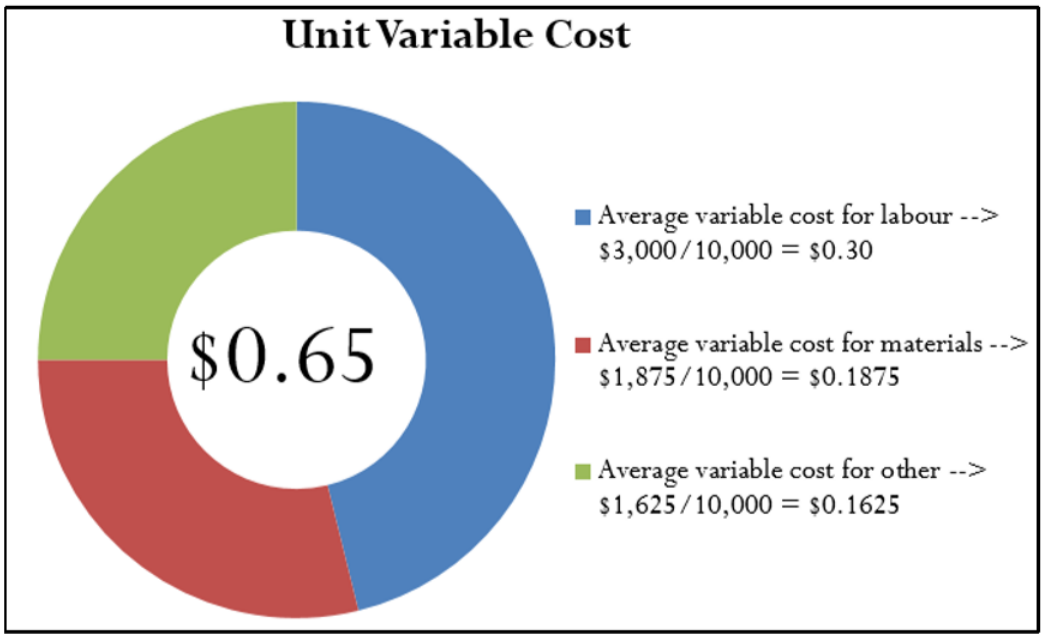
Assume a company produces 10,000 units and wants to know its unit variable cost. It incurs production labor costs of $3,000, material costs of $1,875, and other variable costs totaling $1,625.

Step 1: In this case, all costs are variable costs (production labour and material costs are always variable). The level of output is $n = 10,000$ units.

Step 2: Total all variable costs together to get TVC

$$TVC = \$3,000 + \$1,875 + \$1,625 = \$6,500$$

Step 3: Apply Formula 5.1 to arrive at $VC = \$6,500/10,000 = \$0.65$This means that, on average, the variable cost associated with one unit of production is $0.65.



Important Notes

The definitions of variable, fixed, and blended costs along with their typical associated examples represent a simplified view of how the real world operates. The complexities involved in real-world business costs complicate the fundamentals of managing a business. Therefore, here is how this textbook addresses the complexities of atypical operations, changing fixed costs, and decreasing unit variable costs:

1. **Atypical Operations**. Although there are "normal" ways that businesses operate, there are also businesses that have atypical operations. What is a fixed cost to one business may be a variable cost to another. For example, rent is usually a fixed cost. However, some rental agreements include a fixed base cost plus a commission on the operational output. These rental agreements form blended costs. This textbook does not venture into any of these atypical costs and instead focuses on common cost categorizations.
2. **Changing Fixed Costs**. In real-world applications, fixed costs do not remain flat at all levels of output. As output increases, fixed costs tend to move upwards in steps. For example, at a low level of output only one manager (on salary) may be needed. As output increases, eventually another manager needs to be hired, perhaps one for every 20,000 units produced. In other words, up to 20,000 units the fixed costs would be constant, but at 20,001 units the fixed costs take a step upwards as another manager is added. The model presented in this textbook does not address these upward steps and treats fixed costs as a constant at all levels of output.
3. **Decreasing Unit Variable Costs**. Production tends to realize efficiencies as the level of output rises, resulting in the unit variable cost dropping. This is commonly known as achieving economies of scale. As a consumer, you often see a similar concept in your retail shopping. If you purchase one can of soup, it may cost $1. However, if you purchase a bulk tray of 12 cans of soup it may cost only $9, which works out to 75¢ per can. This price is lower partly because the retailer incurs lower costs, such as fewer cashiers to sell 12 cans to one person than to sell one can each to 12 different people. Now apply this analogy to production. Producing one can of soup costs 75¢. However, a larger production run of 12 soup cans may incur a cost of only $6 instead of $9 because workers and machines can multitask. This means the unit variable cost would decrease by 25¢ per can. However, the model in this textbook assumes that unit variable costs always remain constant at any given level of output.

You are considering starting your own home-based Internet business. After a lot of research, you have gathered the following financial information:

Dell computer	$214.48 monthly lease payments

Office furniture (desk and chair)	$186.67 monthly rental
Shaw high-speed Internet connection	$166.88 per month
Your wages	$30 per hour
Utilities	$13 per month plus $0.20 per hour usage
Software (and ongoing upgrades)	$20.00 per month
Business licenses and permits	$27.00 per month
Google click-through rate	$10.00 per month + $0.01 per click payable as total clicks per sale

Generating and fulfilling sales of 430 units involves 80 hours of work per month. Based on industry response rates, your research also shows that to achieve your sales you require a traffic volume of 34,890 Google clicks. On a monthly basis, calculate the total fixed cost, total variable cost, and unit variable cost.

Solution

Calculate the monthly total fixed costs (TFC), total variable costs (TVC), and the unit variable cost (VC).

What You Already Know

There are three different types of costs. Other known information includes:

$n = 430$ units Hours = 80

Google clicks = 34,890

How You Will Get There

Step 1:

Sort the costs into fixed and variable. Separate the components for blended costs. You can total the fixed costs to arrive at the total fixed costs, or TFC.

Step 2:

Calculate variable costs, then total them to arrive at the total variable cost, or TVC.

Step 3:

Apply Formula 5.1 to calculate the unit variable cost.

Perform

Step 1:

Fixed Costs

Dell computer	$214.48
Office furniture	$186.67
Shaw high-speed Internet connection	$166.88
Utilities (blended cost)	$13.00 only
Software	$20.00
Business licenses / permits	$27.00
Google click (blended cost)	$10.00
TOTAL FIXED COSTS	TFC = $638.03

Variable Costs

Wages	$30.00 per hour
Utilities (blended cost)	$0.20 per hour
Google click (blended cost)	$0.01 per click

Step 2:

$$\begin{aligned} TVC &= (\$30.00 \times \text{ hours }) + (\$0.20 \times \text{ hours }) + (\$0.01 \times \text{ Google clicks }) \\ &= (\$30.00 \times 80) + (\$0.20 \times 80) + (\$0.01 \times 34,890) = \$2,764.90 \end{aligned}$$

Step 3:

$$VC = \frac{\$2,764.90}{430} = \$6.43$$

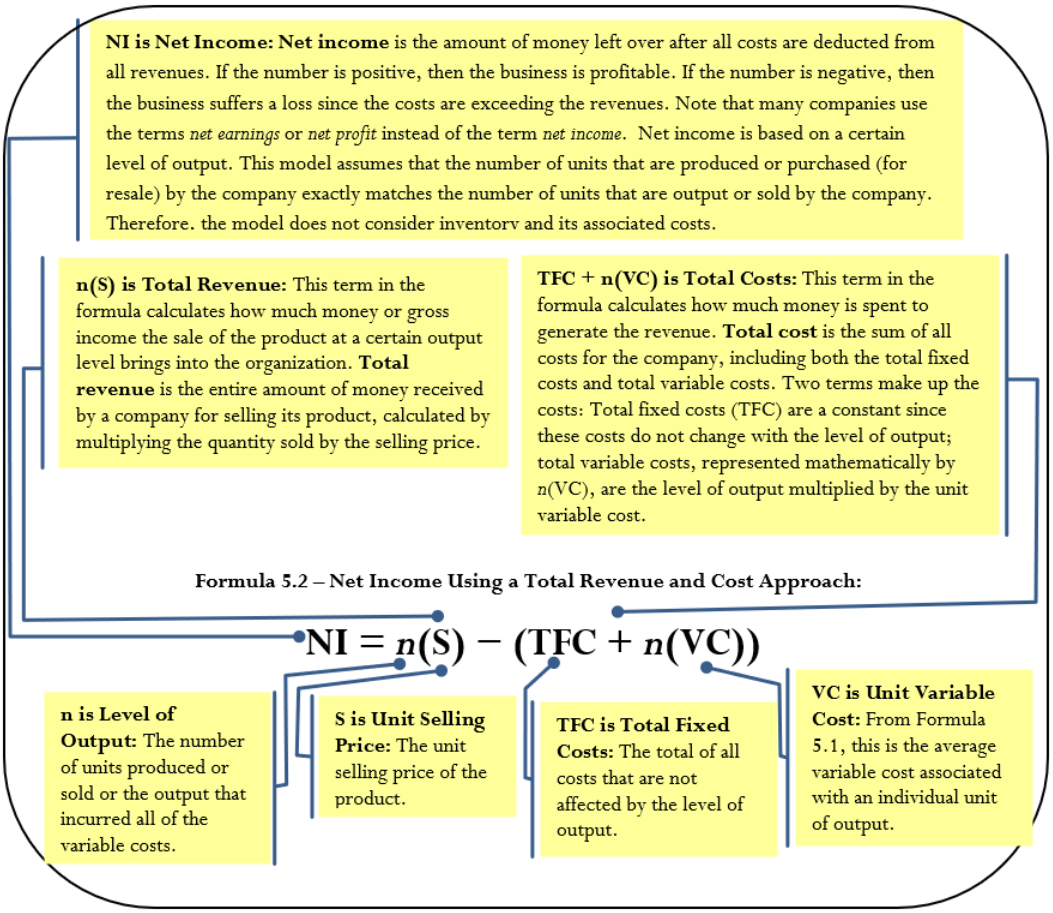
Seven components make up the fixed costs: the computer, furniture, Internet service, fixed utilities, software, licenses/permits, and the fixed Google cost, all totaling $638.03. Three components make up the variable costs: wages, hourly utilities, and Google clicks, totaling $2,764.90. The average unit variable cost is $6.43 per unit sold.

Net Income Using a Total Revenue and Total Cost Approach

Most businesses are "for-profit" businesses, meaning that they operate to make money. In Example 2.3.2.1, you figured out the total fixed costs, total variable costs, and the unit variable cost for your Internet business. However, you left unanswered one of the most important questions in business: If you sell the planned 430 units, are you profitable? Is there any money left after you pay for all of those fixed and variable costs? You must also remember that 430 units is just an estimate. What happens if you sell only 350 units? What happens if you sell 525 units? What happens if you decide to pay yourself a higher wage?

There are no guarantees in business, and the future is always uncertain. Successful business managers plan for the future and perform many “what-if” scenarios to answer questions such as those above. This section develops a model for calculating total net income based on total revenues and total costs. The model allows managers to analyze various scenarios and determine the impact on profitability.

The Formula



Formula 5.2

How It Works

Follow these steps to calculate net income using a total revenue and total income approach:

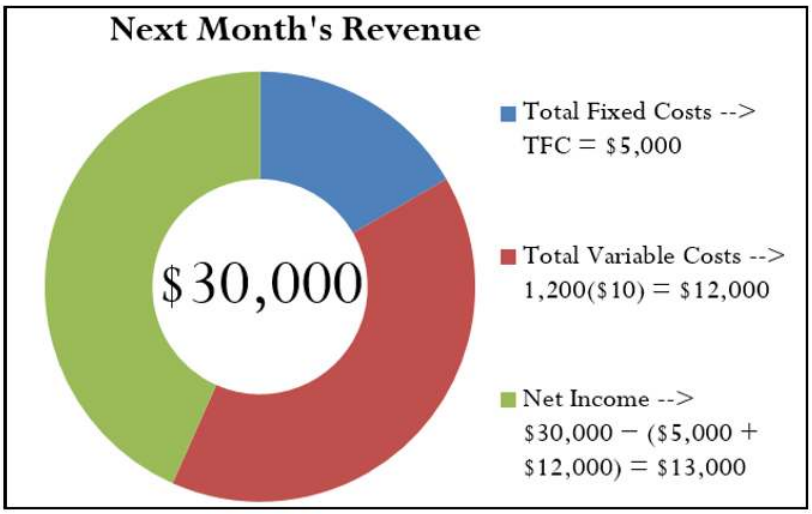
- **Step 1**: Calculate the total revenue. This requires identifying the unit selling price of the product and multiplying it by the level of sales.
- **Step 2**: Calculate your total costs. This requires identifying and separating costs into fixed and variable components. Fixed costs are totaled to arrive at the total fixed cost. Total variable costs are either known or can be calculated through multiplying the unit variable cost by the level of output.
- **Step 3**: Calculate the net income by applying Formula 5.2.

For example, assume that last month a company incurred total variable costs of $10,000 in the course of producing 1,000 units. For next month it forecasts total fixed costs of $5,000 and all variable costs remaining unchanged. Projected production for next month is 1,200 units selling for $25 each. You want to estimate next month's net income.

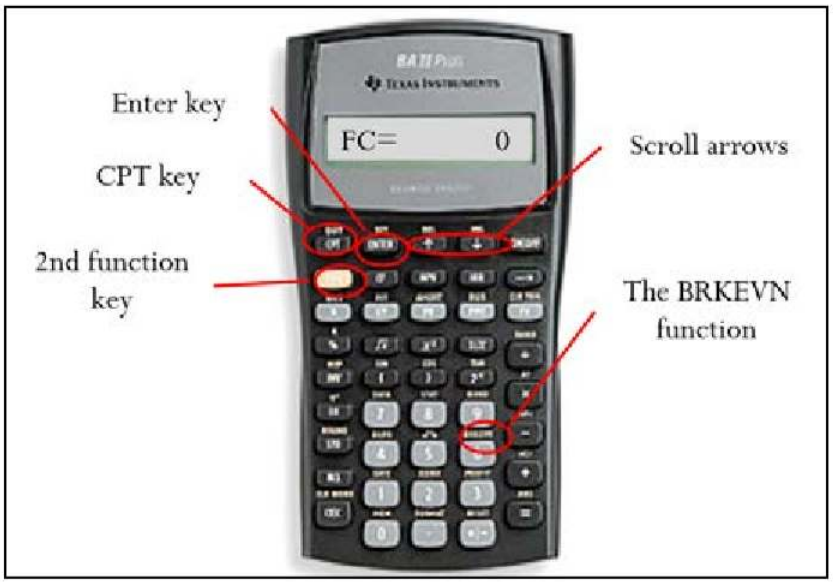
Step 1: Using Formula 5.2, you calculate total revenue from $n(S)$, or the total level of output multiplied by the price of the product. If you project sales of 1,200 units (n) at $25 each (S), then the total forecasted revenue is $1,200(\$25) = \$30,000$

Step 2: Total fixed costs, or TFC, are $5,000. To get the total variable costs, you must resolve $n(VC)$. You calculate the unit variable cost, or VC, with Formula 5.1. Using the current month figures, you see $VC = \$10,000 \div 1,000 = \$10$. If the projected level of output is 1,200 units, then the total variable costs are $1,200(\$10) = \$12,000$

Step 3: Applying Formula 5.2 you have $NI = \text{Total Revenue} - \text{Total costs} = \$30,000 - (\$5,000 + \$12,000) = \$13,000$. Based on the numbers, you forecast net income of $13,000 for next month.



Important Notes



The Texas Instruments BAII Plus calculator is programmed with a version of Formula 5.2. The function is called Brkevn, and you access it by pressing 2nd and then the number six key. The relationship between the formula symbols and the calculator symbols is displayed in the table below.

Variable	Formula 5.2 Notation	Calculator Notation
Total fixed cost	TFC	FC
Unit variable cost	VC	VC
Price per unit	S	P (for price)
Net income	NI	PFT (for profit)
Level of output	n	Q (for quantity)

To solve Formula 5.2 for the PFT or any other variable, enter data into all of the above variables except one. Keying in a variable requires inputting the value and pressing Enter. Use ↑ and ↓ to scroll through the display. When you are ready, scroll to the unknown variable and press CPT.

Paths To Success

An easy way to remember Formula 5.2 is to understand what the formula represents. As explained, the calculation of $n(S)$ multiplies quantity by price to produce the total revenue. The $(TFC + n(VC)$ takes the total fixed costs and adds the total variable costs (which is a function of quantity multiplied by unit variable cost) to arrive at the total cost. Therefore, Formula 5.2 expressed more simply is:

$$\text{Net Income} = \text{Total Revenue} - \text{Total Cost}$$

? Exercise 2.3.2.1: Give It Some Thought

In the following situations, explain what would happen to net income and why.

a. The selling price is raised.
b. The hourly wages of production workers are increased to match the increase in the consumer price index.
c. The level of output decreases.
d. Your insurance company lowers your insurance premiums because your company has had no claims in the past year.

Answer

a. Net income increases because the total revenues increase with no similar increase in costs.
b. Net income decreases because the hourly wages of production workers are variable costs. If the costs go up, then less money is left over.
c. Net income decreases because both total revenue and total variable costs drop; however, the fixed costs remain the same, so there is less revenue to cover proportionally higher costs.
d. Net income increases because the insurance premium is a fixed cost that has become smaller. With lower costs, more money is left over.

Recall from Example 2.3.2.1 that you are considering starting your own home-based Internet business. The following information is known:

FC = \$638.03, VC = \$6.43, forecasted n = 430

Based on these numbers, calculate:

a. The forecasted net income if your price per unit is \$10.
b. The dollar change in net income if you decide to pay yourself a higher wage of \$35.38 per hour instead of \$30.00 per hour while still working 80 hours. Note the total variable costs excluding wages were \$364.90.
c. The dollar change in net income if sales are 30% lower than your initial forecast, ignoring part (b) calculations.

Solution

a. You need to determine net income (NI) using the revenue and cost numbers.
b. Recalculate the total variable cost (TVC) to determine the impact of the higher wage. Then calculate the change in net income (NI).
c. Modify the level of output (n) and calculate the change in net income (NI).

What You Already Know

The costs, level of output, and price are known:

FC = \$638.03, VC = \$6.43, forecasted n = 430, S = \$10

For part (b), the total variable costs excluding wages are known:

TVC (excluding wages) = \$364.90

New wage = \$35.38

Hours = 80

For part (c), the change in the level of output is known:

$\Delta\% = -30\%$

How You Will Get There

a.

Step 1:

Calculate total revenue at the forecasted output.

Step 2:

Calculate total costs at the forecasted output.

Step 3:

Apply Formula 5.2.

b.

Step 1:

Total revenue is unchanged since the output has not changed.

Step 2:

Fixed costs are unchanged. Recalculate the total variable costs by adding the new wages to the total variable costs excluding wages. By rearranging Formula 5.1, you can see $TVC = n(VC)$. Therefore, substitute TVC in place of $n(VC)$ in Formula 5.2.

Step 3:

Apply Formula 5.2 to recalculate the new net income. Compare to net income in part (a) and determine the change.

c.

Step 1:

The revised level of output is a percent change calculation where $\Delta\% = -30\%$ and $OLD = 430$. Solve Formula 3.1 for New. Once New is known, recalculate total revenue.

Step 2:

Fixed costs are unchanged. Recalculate total variables costs using the revised level of output.

Step 3:

Apply Formula 5.2 to recalculate the new net income. Compare to net income in part (a) and determine the change.

Perform

a.

Step 1:

$$\text{Total revenue } = n(S) = 430(\$10) = \$4,300$$

Step 2:

$$TFC = \$638.03; TVC = n(VC) = 430(\$6.43) = \$2,764.90$$

Step 3:

$$NI = s4,300 - (\$638.03 + \$2,764.90) = \$897.07$$

b.

Step 1:

$$\text{Total revenue } = \$4,300$$

Step 2:

$$TFC = \$638.03; TVC = \$364.90 + (\$35.38 \times 80 \text{ hours }) = \$3,195.30$$

Step 3:

$$NI = \$4,300 - (\$638.03 + \$3,195.30) = \$466.67$$
$$\text{Change in } NI = \$466.67 - \$897.07 = -\$430.40$$

c.

Step 1:

$$-30\% = \frac{\text{New} - 430}{430} \times 100$$
$$-129 = \text{New} - 430$$
$$301 = \text{New}$$
$$\text{Total revenue} = 301(\$10) = \$3,010$$

Step 2:

$$TFC = \$638.03; TVC = 301(\$6.43) = \$1,935.43$$

Step 3:

$$NI = \$3,010 - (\$638.03 + \$1,935.43) = \$436.54$$
$$\text{Change in } NI = \$436.54 - \$897.07 = -\$460.53$$

Calculator Instructions

	FC	VC	P	PFT	Q
a.	638.03	6.43	10	Answer: 897.07	430
b.	√	3195.3 ÷ 430 =	√	Answer: 466.67	√
c.	√	6.43	√	Answer:436.54	301*

*First use $\Delta\%$ function to calculate this number:

	OLD	NEW	%CH	#PD
c.	430	Answer: 301	-30	1

Under the initial scenario, a net income of $897.07 on sales of 430 units is forecast. If you pay yourself a higher wage of $35.38 per hour, your net income decreases by $430.40 to $466.67. If your forecast is inaccurate and is lower by 30%, then your net income is reduced by $460.53, resulting in a lower net income of $436.54.

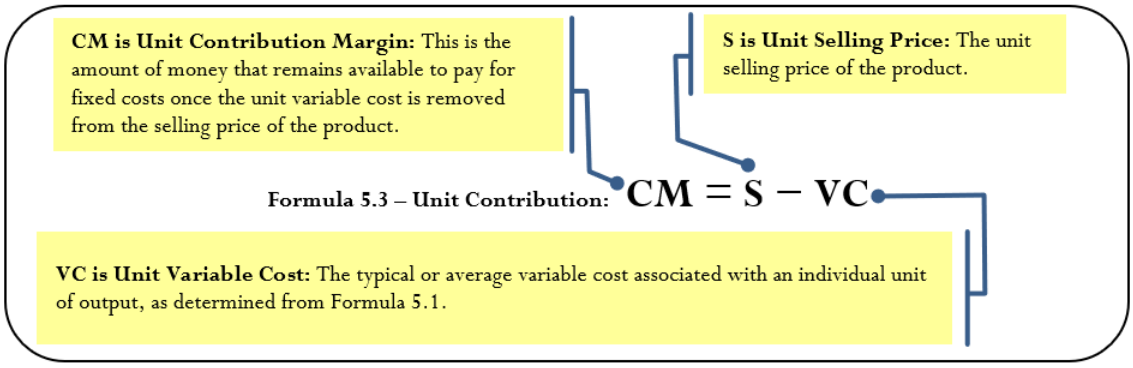
Net Income Using a Total Contribution Margin Approach

In Example 2.3.2.2 you learned that if you sell the projected 430 units of product for your Internet business, the total net income is $897.07. What if you sold 431 units of the product? How would your net income change? Clearly it should rise, but by how much? One approach to answering this question is to rerun the numbers through Formula 5.2, revising the total revenues and total variable costs to calculate a new net income. This new net income can then be compared against the existing net income to determine how it changed. This is a multi-step approach and involves a lot of work. An alternative approach explored in this section involves using a unit contribution margin to calculate the net income. The benefit of this approach is that with minimal calculations you can easily assess the impact of any change in the level of output.

In accounting and marketing, the contribution margin is the amount that each unit sold adds to the net income of the business. This approach allows you to understand the impact on net income of each unit sold, and Section 5.2 will explain its further benefit of allowing for a straightforward calculation of a break-even point. The contribution margin determines on a per-unit basis how much money is left over after unit variable costs are removed from the price of the product. This leftover money is then available to pay for the fixed costs. Ultimately, when all fixed costs have been paid for, the leftover money becomes the profits of the business. If not enough money is left over to pay for the fixed costs, then the business has a negative net income and loses money.

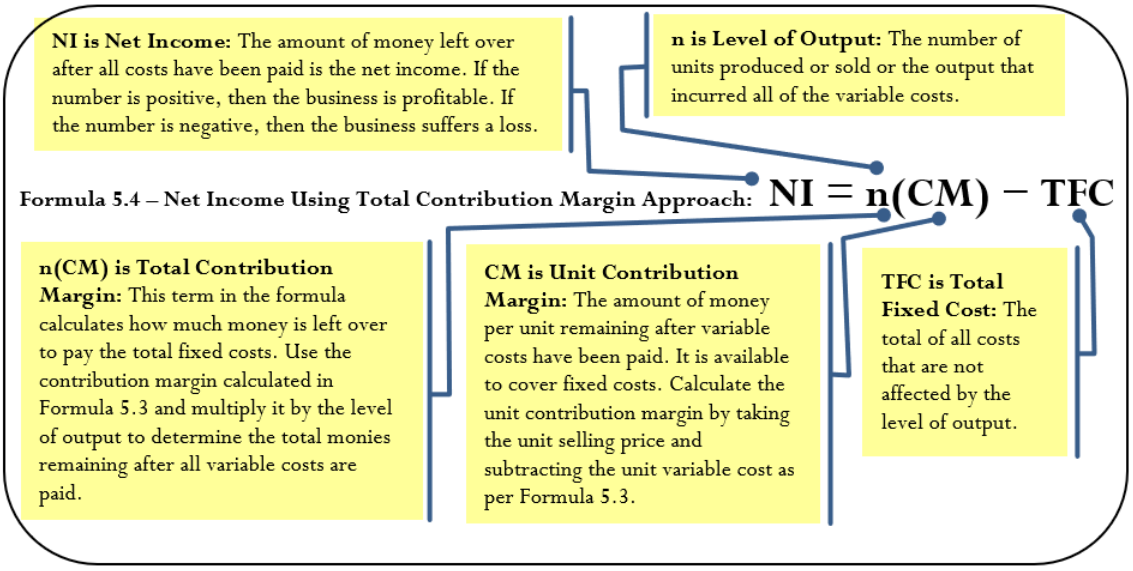
The Formula

The difference between the increase in the total revenue and the increase in the total variable costs is how much the sale of an individual product contributes toward the change in your net income. Formula 5.3 expresses this relationship.



Formula 5.3

If you have no units sold, your net income is negative and equal to the total fixed costs associated with your business, since there is no offsetting revenue to pay for those costs. With each unit sold, the contribution margin of each product is available to pay off the fixed costs. Formula 5.4 expresses this relationship when calculating net income.



Formula 5.4

How It Works

Follow these steps to calculate the net income using a contribution margin approach:

Step 1: If unit information is known, apply Formula 5.3 and calculate the unit contribution margin by subtracting the unit variable cost from the selling price. This may or may not require you to use Formula 5.1 to calculate the unit variable cost.

Step 2: Calculate the total contribution margin by multiplying the contribution margin with the associated level of output.

Step 3: Calculate the total fixed costs by adding all fixed costs.

Step 4: Based on the level of output, calculate the net income by applying Formula 5.4.

For example, using the contribution margin approach, calculate the net income for a product that sells for $75, has unit variable costs of $31, total fixed costs of $23,000, and total sales of 800 units.

Step 1: The unit contribution margin is calculated from Formula 5.3. If the product sells for $75 and has unit variable costs of $31, then $CM = \$75 - \$31 = \$44$. This means that every unit sold has $44 left over to contribute toward fixed costs.

Step 2: Now convert that into a total contribution margin. The first part of Formula 5.4 calculates total contribution margin through $n(CM)$. With sales of 800 units, the total contribution margin is $800(\$44) = \$35,200$

Step 3: Total fixed costs are known: $TFC = \$23,000$

Step 4: Apply Formula 5.4, which translates to $\text{Net Income} = \text{Total contribution margin - Total fixed costs}$, or $\$35,200 - \$23,000 = \$12,200$

Paths To Success

When you work with Formula 5.4, sometimes unit information may not be known. Instead, you might just have a single aggregate number representing the total contribution margin for which somebody has already taken the total revenue and subtracted the total variable costs. In this case, skip step 1 and take the provided total contribution margin as the answer for step 2 with no calculations necessary.

Example 2.3.2.3: The Contribution Margin for Your Internet Business

Continuing with Examples 2.3.2.1 and 2.3.2.2, calculate the unit contribution margin and net income using the contribution margin approach. From the previous examples, recall the unit variable cost is \$6.43, unit selling price is \$10, total fixed costs are \$638.03, and the projected sales are 430 units.

Solution

Calculate the unit contribution margin (CM) followed by the net income (NI) using the contribution margin approach.

What You Already Know

The cost, price, and sales information are known:

$VC = \$6.43, S = \$10$

$TFC = \$638.03, n = 430$

How You Will Get There

Step 1:

Calculate the unit contribution margin using Formula 5.3.

Step 2:

Calculate the total contribution margin by multiplying the unit contribution margin times the level of output.

Step 3:

Determine the total fixed costs.

Step 4:

To calculate net income, apply Formula 5.4.

Perform

Step 1:

$$CM = \$10.00 - \$6.43 = \$3.57$$

Step 2:

$$n(CM) = 430(\$3.57) = \$1,535.10$$

Step 3:

$$TFC = \$638.03$$

Step 4:

$$NI = \$1,535.10 - \$638.03 = \$897.07$$

Calculator Instructions

FC	VC	P	PFT	Q
638.03	6.43	10	Answer: 897.07	430

When a product sells for $10, $6.43 goes toward paying for the variable costs of your business, leaving $3.57 as your unit contribution margin. This means that for every unit increase in sales, your net income rises by $3.57. Thus, if you sell 430 units you have a total contribution margin of $1,535.10, which results in a net income of $897.07 after removing fixed costs of $638.03.

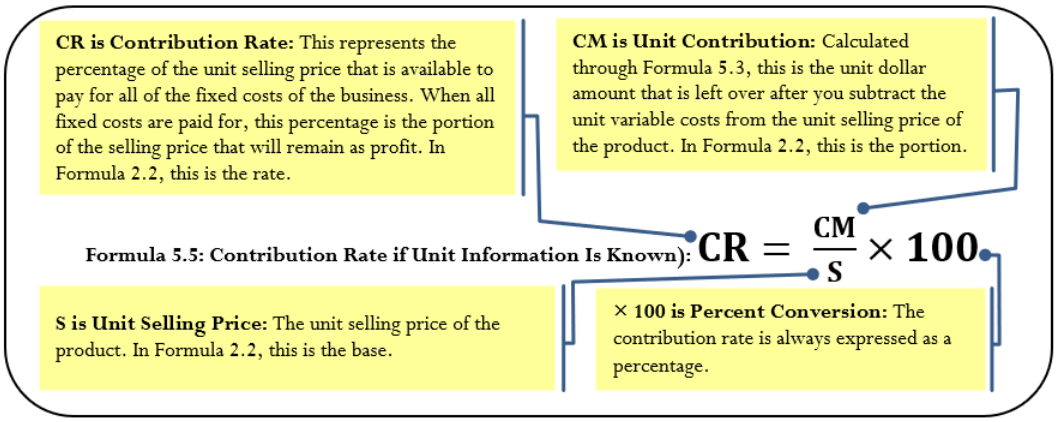
Contribution Rates

It is difficult to compare different products and their respective dollar amount contribution margins if their selling prices and costs vary widely. For example, how do you compare a unit contribution margin of $1,390 (selling price of $2,599.99) on a big screen television to a unit contribution margin of $0.33 on a chocolate bar (selling price of $0.67)? On a per-unit basis, which contributes relatively more to fixed costs? To facilitate these comparisons, the products must be placed on equal terms, requiring you to convert all dollar amount contribution margins into percentages. A contribution rate is a contribution margin expressed as a percentage of the selling price.

The Formula

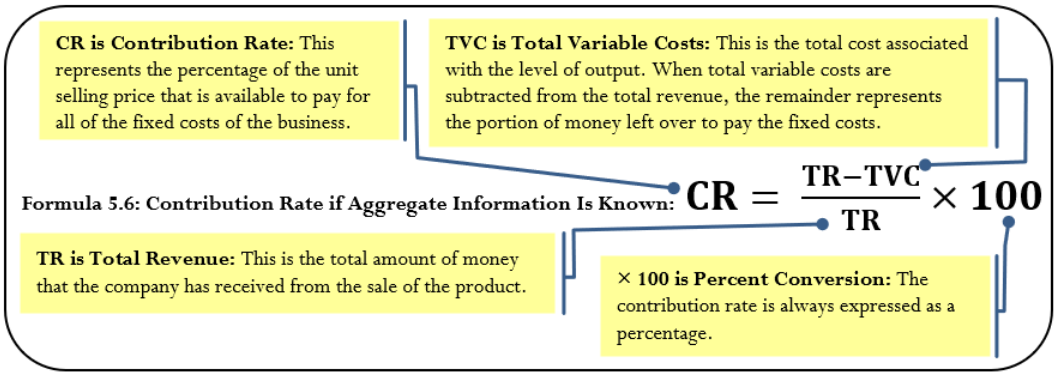
Your choice between two formulas for calculating a contribution rate depends on whether you have unit information or only aggregate information. Both formulas are adaptations of Formula 2.2 (Rate, Portion, Base).

If unit information is known, including the unit variable cost and unit selling price, then calculate the contribution rate using unit information as expressed in Formula 5.5.



Formula 5.5

If any unit information, including the unit variable cost or unit selling price, is unknown or unavailable, then you cannot apply Formula 5.5. Sometimes only aggregate information is known. When total revenue and total variable costs for any product are known or can at least be calculated, then you must calculate the contribution rate from the aggregate information as expressed in Formula 5.6.



Formula 5.6

How It Works

Follow these steps to calculate a contribution rate:

Step 1: Identify the required variables and calculate the margin, if needed.

- If unit information is known, this requires you to calculate the unit contribution margin. Otherwise, calculate the unit contribution margin by applying Formula 5.3. You must identify the unit selling price.
- If aggregate information is known, you need to identify total revenue and total variable costs.

Step 2: Calculate the contribution rate.

- If unit information is known, apply Formula 5.5.
- If only aggregate information is known, apply Formula 5.6.

As an example of these steps, recall earlier that you wanted to compare the relative contributions of the big screen television and the chocolate bar. The television sells for $2,599.99 and has a unit contribution margin of $1,390. The chocolate bar sells for $0.67 and has a unit contribution margin of $0.33. Notice that the information being provided is on a per-unit basis. Calculate the contribution rate of each product.

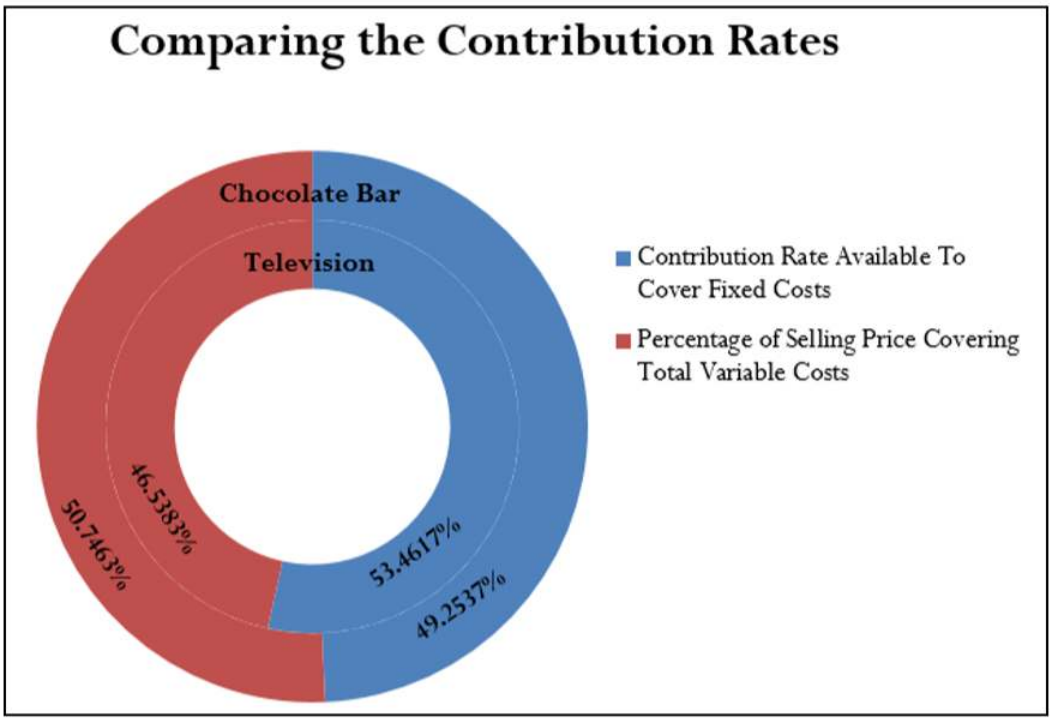
Step 1: The contribution margins are known. For the television, CM = $1,390, and for the chocolate bar, CM = $0.33.

Step 2: Applying Formula 5.5, the television $CR = \$1,390.00 \div \$2,599.99 \times 100 = 53.4617\%$ while the chocolate bar $CR = \$0.33 \div \$0.67 \times 100 = 49.2537\%$ It is now evident from the contribution rate that 4.208% more of the television's selling price is available to pay for fixed costs as compared to the chocolate bar's price.

Now change the facts. This time, assume there is no unit information. Instead, the television's total revenue is $129,999.50 and associated total variable costs are $60,499.50. The chocolate bar has total revenue of $3,886.00 with total variable costs of $1,972.00. Based on this information, you are to determine the product with the higher contribution rate.

Step 1: The aggregate numbers are known for both products. For the television, $TR = \$129,999.50$ and $TVC = \$60,499.50$ For the chocolate bar, $TR = \$3,886.00$ and $TVC = \$1,972.00$

Step 2: Applying Formula 5.6, the television $CR = \dfrac{\$129,999.50 - \$60,499.50}{\$129,999.50} \times 100 = 53.4617\%$ while the chocolate bar $CR = \dfrac{\$3,886 - \$1,972}{\$3,886} \times 100 = 49.2537\%$ You have reached the same conclusion as above.



Exercise 2.3.2.2: Give It Some Thought

In each of the following situations, what would happen to the contribution rate and why?

a. The selling price is raised.
b. The hourly wages of production workers are increased to match the increase in the consumer price index.
c. The level of output decreases.
d. Your insurance company lowers your insurance premiums because your company has had no claims in the past year.

Answer

a. The contribution rate increases, since raising the price increases the unit contribution margin.
b. The contribution rate decreases, since total variable costs rise, eroding some of the unit contribution margin.
c. There is no effect on the contribution rate. The level of output is not a factor in calculating the contribution rate.
d. There is no effect on the contribution rate since the insurance premiums are a fixed cost. Fixed costs are not a factor in calculating contribution rates

Example 2.3.2.4: Understanding the Contribution Rate of Your Internet Business

Continuing with your ongoing Internet business from the three previous examples, calculate the contribution rate using both the unit and aggregate methods and show how they arrive at the same number. Remember that you are selling products for $10 each, your unit contribution margin is $3.57, total revenue is projected at $4,300, and total variable costs are $2,764.90.

Solution

Your goal is to calculate the contribution rate (CR) using both the unit and aggregate methods.

What You Already Know

Step 1:

Prices, margins, revenue, and costs are known:

$S = \$10, CM = \$3.57$

$TR = \$4,300, TVC = \$2,764.90$

How You Will Get There

Step 2a:

Using the unit method, apply Formula 5.5.

Step 2b:

Using the aggregate method, apply Formula 5.6.

Perform

Step 2a:

$$CR = \frac{\$3.57}{\$10.00} \times 100 = 35.7\%$$

Step 2b:

$$CR = \frac{54,300.00 - \$2,764.90}{\$4,300.00} \times 100 = 35.7\%$$

Under both the unit and aggregate method, your contribution rate equals 35.7%. This means that every time you sell a $10 product, 35.7% of the revenue remains after recovering the cost of the product.

Putting It All Together

You have studied costs, volume, and net income in this section. So far, you have considered each of these concepts separately while you worked through the various applications. It is time to put the types of costs, unit variable cost, net income, sales, contribution margin, and contribution rate together so that you can see all of these concepts in a single scenario. Look at the following example.

Example 2.3.2.5: Running a Pizza Delivery Business

In the commercial section of the newspaper you come across an ad for a pizza delivery business for sale. Upon inquiry, you discover that the owner, who wants to sell the business and then retire, has four salaried employees and owns two delivery vehicles. He invites you to look through his books, where you acquire the following information:

Employees: Owner's salary Employee salaries and premiums	 $5,000 per month $2,000 per month each	**Business:** Phone Business insurance Building lease	 $85 per month $3,600 per year $2,000 per month
Delivery Vehicles: Car insurance Fuel Oil changes Vehicle maintenance and repairs	 $1,200 per year per vehicle $1,125 per month per vehicle $225 per month per vehicle $562.50 per month per vehicle	**Operations:** Pizza ingredients, materials, and packaging Selling price of pizza Average number of pizzas sold	 $19,125 per month $9 per pizza 4,500 per month

Assume that this is all of the key information. You need to understand this business and therefore want to determine:

a. A unit variable cost
b. A typical monthly net income for the business
c. The contribution margin per pizza
d. The contribution rate

Solution

All calculations require the sorting of costs into fixed and variable components, which will be your first step. You then calculate a unit variable cost (VC) using the historical operations information. Once you obtain this, you can calculate the net income (NI), contribution margin (CM), and contribution rate (CR).

What You Already Know

You have unit information on costs, volume, and revenue, as listed above.

How You Will Get There

a. Unit variable cost:

Step 1:

Sort the costs into fixed and variable categories. Total the monthly fixed costs.

Step 2:

Total the monthly variable costs.

Step 3:

Calculate the unit variable cost by applying Formula 5.1.

b. Net income:

Step 1:

Calculate total revenue at the level of output.

Step 2:

Calculate total costs at the level of output.

Step 3:

Calculate monthly net income by applying Formula 5.2.

c. and d. Contribution margin and rate:

Step 1:

Calculate the contribution margin by applying Formula 5.3.

Step 2:

Calculate the contribution rate by applying Formula 5.5.

Perform

a. Unit variable cost:

Steps 1 and 2:

Fixed Costs per Month

Owner's salary	$5,000
Employee salaries and premiums	$2,000 × 4 = $8,000
Car insurance	$1,200 × 2 cars ÷ 12 months = $200
Phone	$85
Business insurance	$3,600 ÷ 12 months = $300
Building lease	$2,000
TOTAL FIXED COSTS	TFC = $15,585 per month

Variable Costs per Month

Fuel	$1,125 × 2 vehicles = $2,250
Oil changes	$225 × 2 vehicles =$450
Vehicle maintenance and repairs	$562.50 × 2 vehicles = $1,125

Pizza ingredients, materials, and packaging	$19,125
TOTAL VARIABLE COSTS	TVC = $22,950 per month

Step 3:

$$VC = \frac{\$2,250 + \$450 + \$1,125 + \$19,125}{4,500 \text{ pizzas}} = \$5.10 \text{ per pizza}$$

b. Net income:

Step 1:

$$\text{Total Revenue } = 4,500(\$9) = \$40,500$$

Step 2:

$$TFC = \$15,585; TVC = 4,500(\$5.10) = \$22,950$$

Step 3:

$$NI = \$40,500 - (\$15,585 + \$22,950) = \$40,500 - \$38,535 = \$1,965$$

c. and d. Contribution margin and rate:

Step 1:

$$CM = \$9.00 - \$5.10 = \$3.90$$

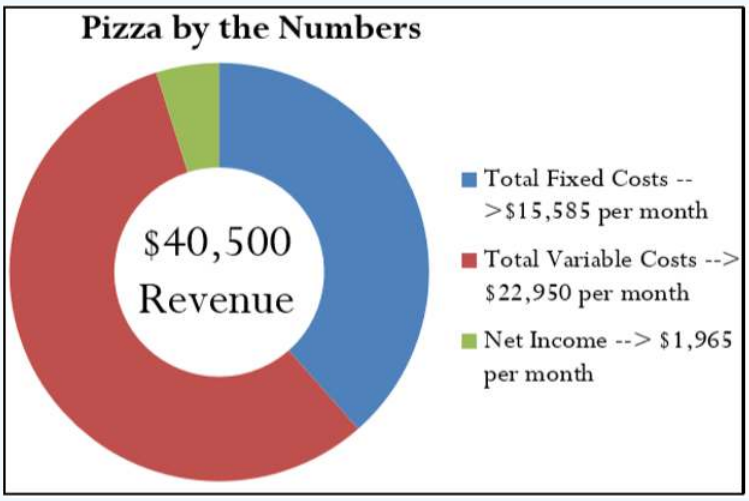
Step 2:

$$CR = \frac{53.90}{59.00} \times 100 = 43.\bar{3}\%$$

Calculator Instructions

FC	VC	P	PFT	Q
15585	5.1	9	Answer: 1,965	4500

This is a profitable business. For an average month, the business incurs $15,585 in fixed costs along with $22,950 in variable costs, which works out to a unit variable cost of $5.10 for each of the 4,500 pizzas sold. With total revenue of $40,500, after removing both fixed and variable costs, net income is $1,965 per month. Every pizza, after paying for the unit variable costs, has $3.90, or $43.\bar{3}\%$, left over.



Contributors and Attributions

- Template:ContribBusiMathOlivier

This page titled 2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know) is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Jean-Paul Olivier via source content that was edited to the style and standards of the LibreTexts platform.

- **5.1: Cost-Revenue-Net Income Analysis (Need to Be in the Know)** by Jean-Paul Olivier is licensed CC BY-NC-SA 4.0. Original source: https://open.bccampus.ca/browse-our-collection/find-open-textbooks/?uuid=16301119-8ec4-4241-b0f7-cc87ffc942d6.

2.4: Graphing Linear Functions

2.4: Graphing Linear Functions is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

2.4.1: Graphing Linear Equations

In general, a linear equation in two variables is of the form

$$Ax + By = C$$

in which A, B, C are real numbers and A and B cannot both be 0. This is called the **standard form of a linear equation**. Here are a few examples of linear equations in standard form:

$$3x + 2y = 4, \quad -0.25x + \frac{2}{3}y = 1, \quad -3x = 2, \quad 4y = 3$$

Some linear equations are not written in standard form, but they can be easily transformed to a standard form. For example, the linear equation $y = -2x - 3$ can be written in a standard form of $2x + y = -3$ by adding $2x$ to both sides of the equation.

The graph of a linear equation is a straight line in the rectangular coordinate system. In the last chapter, we learned how to plot a straight line by plotting some points of a given linear equation. In this chapter, we will learn some important concepts of linear equations and use them to plot straight lines.

Graphing Linear Equations Using Intercepts

On the graph of a straight line in the rectangular coordinate system, there are two special points, the x -intercept and the y -intercept.

From geometry, we know that a straight line is fully determined by two points. Therefore, after identifying the intercepts of a given linear equation, we can plot its graph by using these two points. It is a good practice to find a third point. If the third point doesn't lie on the straight line determined by the intercepts, this indicates that we may have made an error in our calculations.

The x -intercept

The x -intercept is the point where the graph intersects the x -axis.

To find the x -intercept, evaluate the equation at $y = 0$.

The x -intercept will look like $(x$ -intercept $, 0)$.

The y -intercept

The y -intercept is the point where the graph intersects y -axis.

To find the y -intercept, evaluate the equation at $x = 0$.

The y -intercept will look like $(0, y$ -intercept $)$.

Example 26.1

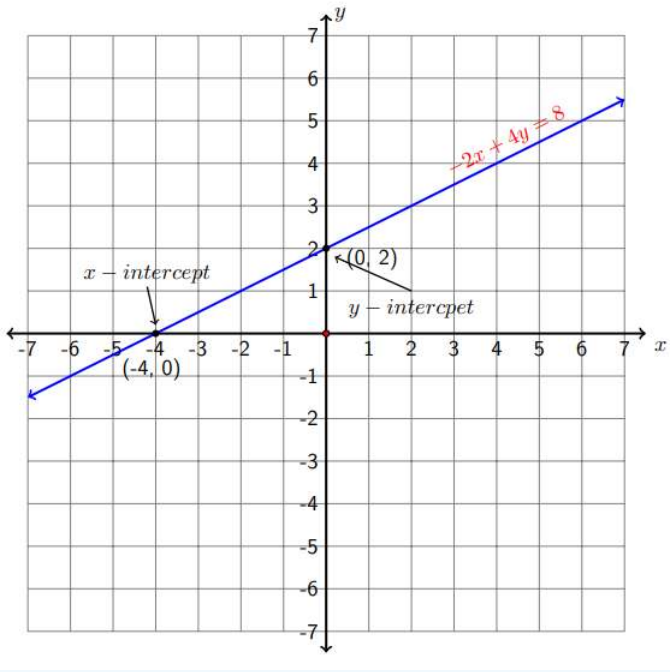
Find the intercepts of $-2x + 4y = 8$ and use them to graph the linear equation.

$$\text{Let } y = 0 \quad \Rightarrow \quad -2x + 4(0) = 8 \quad \Rightarrow \quad -2x = 8 \quad \Longrightarrow \quad x = -4$$

Therefore, the $x-$ intercept is (-4,0).

$$\text{Let } x = 0 \quad \Longrightarrow \quad -2(0) + 4y = 8 \quad \Longrightarrow \quad 4y = 8 \quad \Rightarrow \quad y = 2$$

Therefore, the $y-$ intercept is (0,2)



Example 26.2

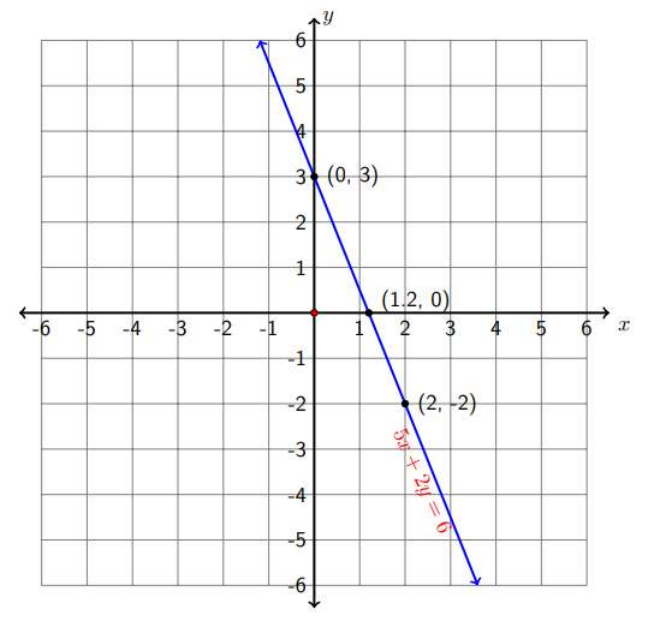
Find the intercepts of $5x+2y=6$ and use them to graph the equation. We first find the two intercepts (and then find a third point).

$$\text{Let } y=0 \implies 5x+2(0)=6 \implies 5x=6 \implies x=\frac{6}{5}=1.2$$

$$\text{Let } x=0 \implies 5(0)+2y=6 \Rightarrow 2y=6 \Rightarrow y=3$$

$$\text{Let } x=2 \implies 5(2)+2y=6 \Rightarrow 10+2y=6 \Rightarrow y=-2$$

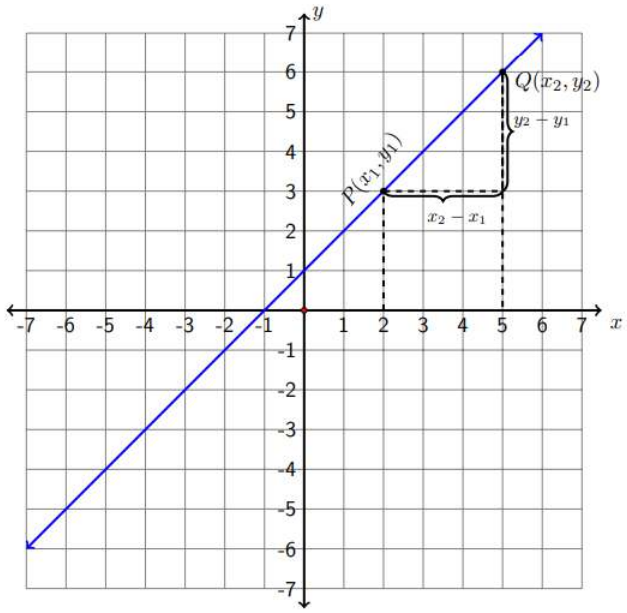
Now we have three solution pairs: (1.2,0),(0,3) and $(2,-2)$. The following figure is the graph of the equation $5x+2y=6$.



The Slope of a Line

Different straight lines have different steepness. Mathematically, there is a number that can measure the steepness of a given line. This number is called the **slope** of a line, and is denoted by the letter m. The following figure helps us to determine the formula for calculating the slope m.

In the figure below, P and Q are two arbitrary points on the given line. Point P has the coordinates (x_1, y_1), and point Q has the coordinates (x_2, y_2). Here we use subscripts to distinguish coordinates for different points. The slope m is defined as the ratio of the change of the two points in the y direction to the change of the two points in the x direction.



since the y direction is vertical, the change in y coordinates referred to as the "rise." Similarly, the x direction is horizontal so the change in x coordinates is referred to as the "run." So another way of remembering the formula for finding the slope is to remember that it is the "rise over run."

The Slope of a Line

The slope of a line passing through (x_1, y_1) and (x_2, y_2) is

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{\text{rise}}{\text{run}}$$

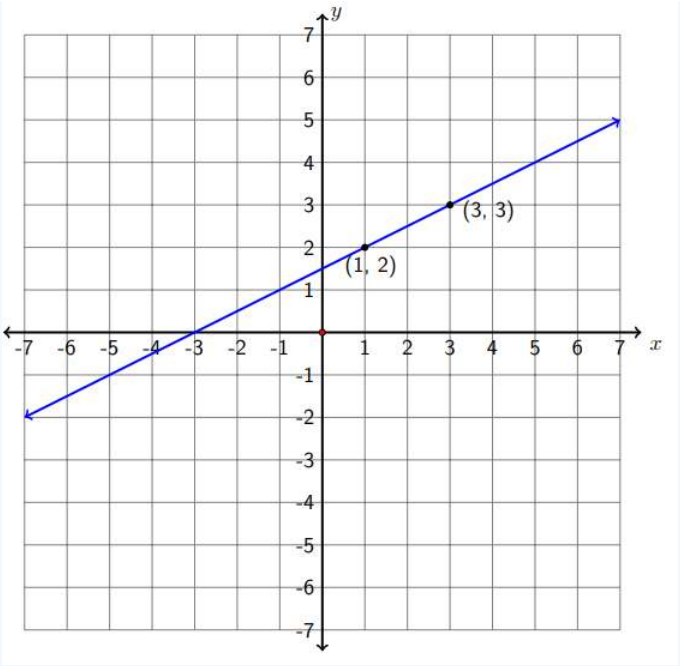
Example 26.3

Find the slope of the line passing through the following points and use the given points to graph the line.

a) (1,2) and (3,3)

We choose $(x_1, y_1) = (1, 2)$ and $(x_2, y_2) = (3, 3)$. Actually it doesn't matter which point we assign as the first and the second point. We will always get the same slope.

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{3 - 2}{3 - 1} = \frac{1}{2}$$

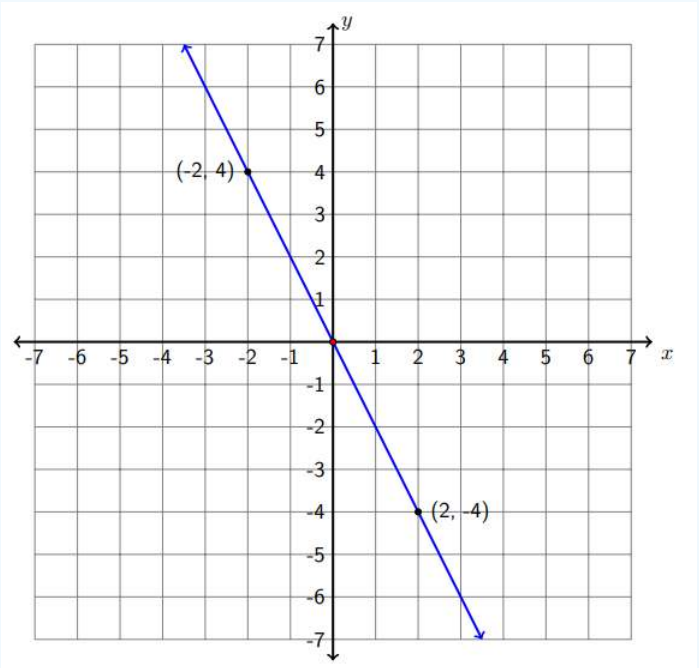


The slope is positive. The graph shows it is a rising line from left to right.

b) (-2,4) and (2,-4)

Let $(x_1, y_1) = (-2, 4)$ and $(x_2, y_2) = (2, -4)$. Then

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-4 - 4}{2 - (-2)} = \frac{-8}{4} = -2$$



The slope is negative. The graph shows it is a falling line from left to right, and this line is steeper than the line in part a).

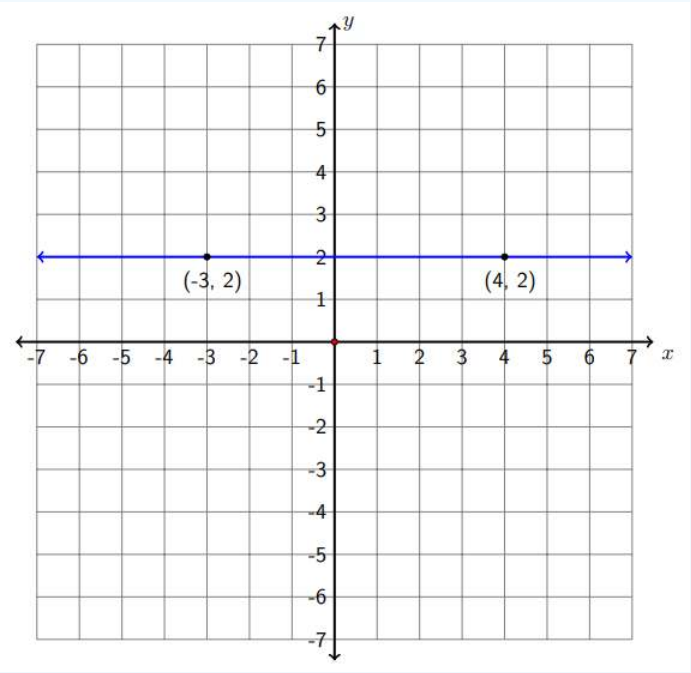
c) (-3,2) and (4,2)

Let $(x_1, y_1) = (-3, 2)$ and $(x_2, y_2) = (4, 2)$. Then

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{2-2}{4-(-3)} = \frac{0}{7} = 0$$

The slope is 0. The graph shows it is a horizontal line.

Note: All horizontal lines are characterized by points that have the same y value, so they all have a slope 0.



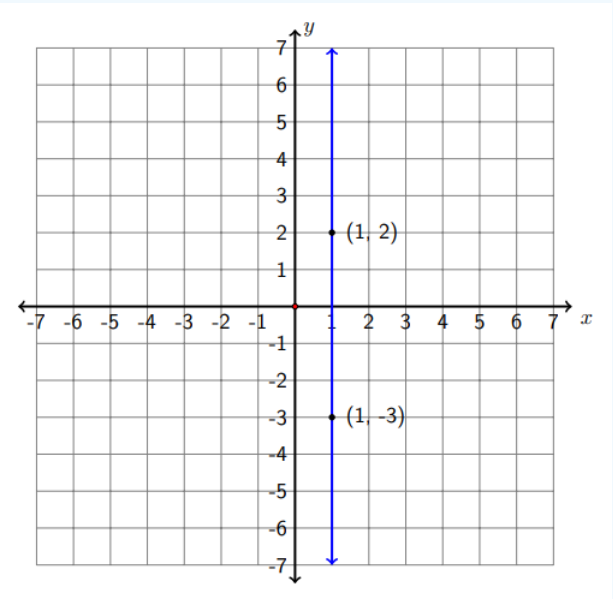
d) (1,2) and (1,-3)

Let $(x_1, y_1) = (1, 2)$ and $(x_2, y_2) = (1, -3)$. Then

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-3-2}{1-1} = \frac{-5}{0} = \text{undefined}$$

The slope is undefined. The graph shows it is a vertical line.

Note: All vertical lines are characterized by points that have the same x value, so they all have an undefined slope.



Properties of the Slope

1. If the slope is positive, the line goes upward from left to right.
2. If the slope is negative, the lines goes downward from left to right.
3. If the slope is 0, the line is horizontal, that is, there is no rise in the y direction.
4. If the slope is undefined, the line is vertical, that is, there is no run in the x direction.
5. The larger the absolute value of the slope, the steeper the straight line.

Slope-Intercept Form of the Equation of a Line $\mathbf{y = mx + b}$

In the previous section, learned that to find the slope of a line, we need two points on the line. Given an equation of a line, we can easily find two points on the line by finding two solution pairs of the equation. However, there is another way to identify the slope.

Let's find the slope of the line given by the equation: $2x - 3y = 4$. We can choose any two points on the line, so let's choose to work with the x and y -intercepts.

$$\text{Let } x = 0 \implies 2(0) - 3y = 4 \implies -3y = 4 \implies y = -\frac{4}{3}$$

$$\text{Let } y = 0 \implies 2x - 3(0) = 4 \implies 2x = 4 \implies x = 2$$

Therefore the intercepts are: $\left(0, -\frac{4}{3}\right)$ and (2,0) and so the slope is :

$$m = \frac{0 - \left(-\frac{4}{3}\right)}{2 - 0} = \frac{\frac{4}{3}}{2} = \frac{4}{3} \cdot \frac{1}{2} = \frac{2}{3}$$

Suppose we return to our original equation and solve it for y :

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

Rewriting the equation in this way helps us to identify two numerical values $\frac{2}{3}$ and $-\frac{4}{3}$. The first value is exactly the same as the slope and the second is exactly the same as the y -intercept! This is true in general. If the equation of a line is written in the form $y = mx + b,$ then the coefficient of the x term is the slope of the line and the constant term is the y -coordinate of the $y-$ intercept (so we write $(0, b)$).

The slope-intercept form of the equation of a line

$y = mx + b$

is called the slope-intercept form of the equation of a line with slope m and y -interecpt $(0, b)$.

Example 26.4

Find the slope and $y-$ intercept of the line by rewriting the line equation $x + 4y = 8$ in the slope-intercept form.

$$\begin{aligned} x + 4y = 8 &\implies 4y = -x + 8 \\ &\implies y = \frac{-x + 8}{4} \\ &\implies y = -\frac{1}{4}x + 2 \end{aligned}$$

The slope is $-\frac{1}{4}$ and the y -intercept is (0,2) . Reminder: The slope is a **numerical value**. Don't let the x variable tag along when you write it down!

Graphing Linear Equations Using the Slope and the y intercept

Using the y -intercept and the slope of a line, we can easily sketch the graph of a straight line. Remember that we only need two points to graph a line. We use the y -intercept as the first point, then we use the slope and the y -intercept to get another point.

Example 26.5

Graph the given linear equation by using the slope and the y -intercept:

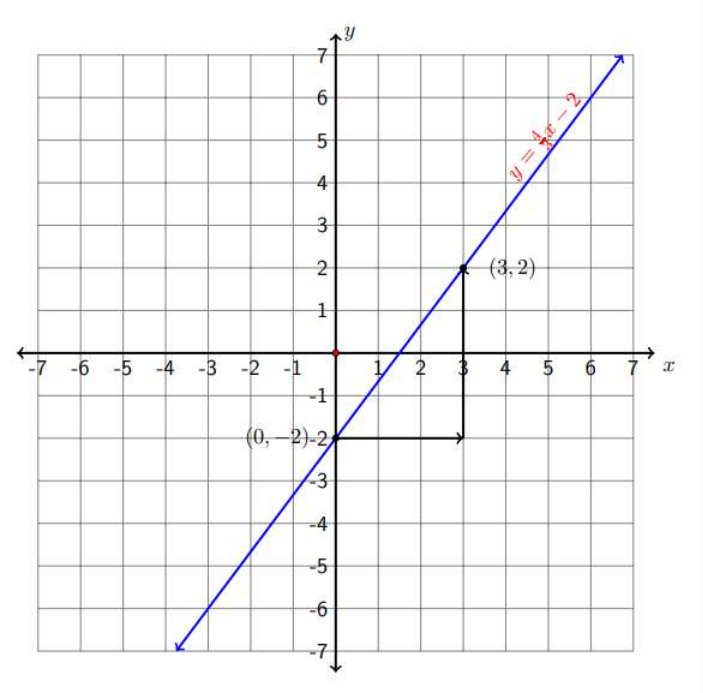
a) $y = \frac{4}{3}x - 2$

Step 1 Write the linear equation in the slope-intercept form. In this problem the equation given is already in $y = mx + b$ form, so we can identify that the slope is $m = \frac{4}{3}$ and the $y-$ intercept is (0,-2).

Step 2 Plot the $y-$ intercept (0,-2) on the $y-$ axis of the coordinate system.

Step 3 Use the slope $m = \frac{4}{3} = \frac{\text{rise}}{\text{run}}$ to locate the second point, that is, from the y -intercept we move to the right (run) three units and up (rise) 4 units to plot a second point at (3,2).

Step 4 Draw a line passing though the two points.



b) $x + 4y = 8$

Step 1 Write the linear equation in the slope-intercept form. (Refer to Example 26.4 to see how this is done):

$$x + 4y = 8 \quad \Longrightarrow \quad y = -\frac{1}{4}x + 2$$

Step 2 Plot the y -intercept (0,2) on the y -axis of the coordinate system.

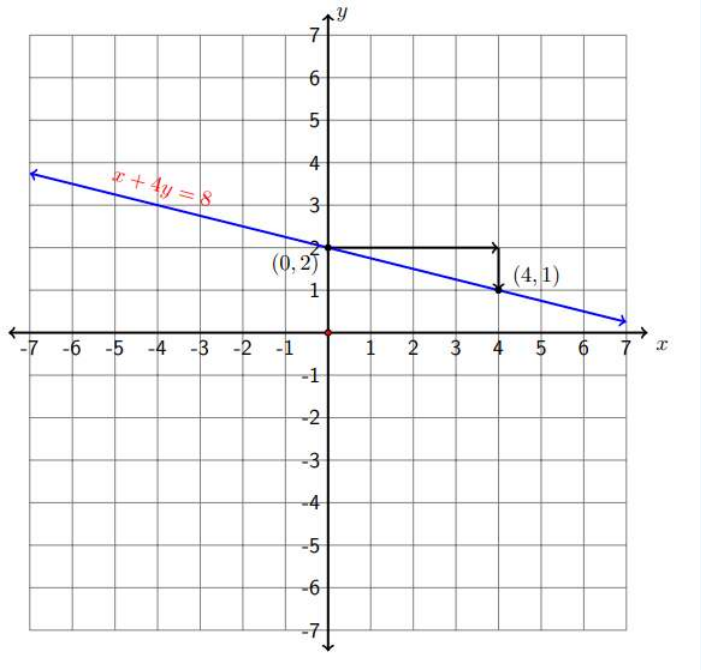
Step 3 Using the slope $m = -\frac{1}{4} = \frac{-1}{4} = \frac{\text{rise}}{\text{run}}$ to locate the second point, that is, from the y -intercept we move to the right (run) four units and down (rise) one unit (because the change in y is negative, we move down) and plot a second point at $(4, 1)$.

Here we chose to put the negative sign in the numerator, but we could have easily assigned the negative sign to the denominator. This can be done because

$$-\frac{1}{4} = \frac{-1}{4} = \frac{1}{-4}$$

If we decided to work with the slope written as $m = \frac{1}{-4}$, from the y -intercept (0,2) we would move four units left (run) and one unit up (rise). This would land us at the point $(-4, 3)$.

Step 4 Draw a line passing though the two points. Whichever two points we end up working with either $(0, 2)$ and $(4, 1)$ or $(0, 2)$ and $(-4, 3)-$ we get the same line, try it! The figure below shows the line as drawn from the points $(0, 2)$ and $(4, 1)$.



c) $2x - y = 0$

Step 1 Write the linear equation in the slope-intercept form.

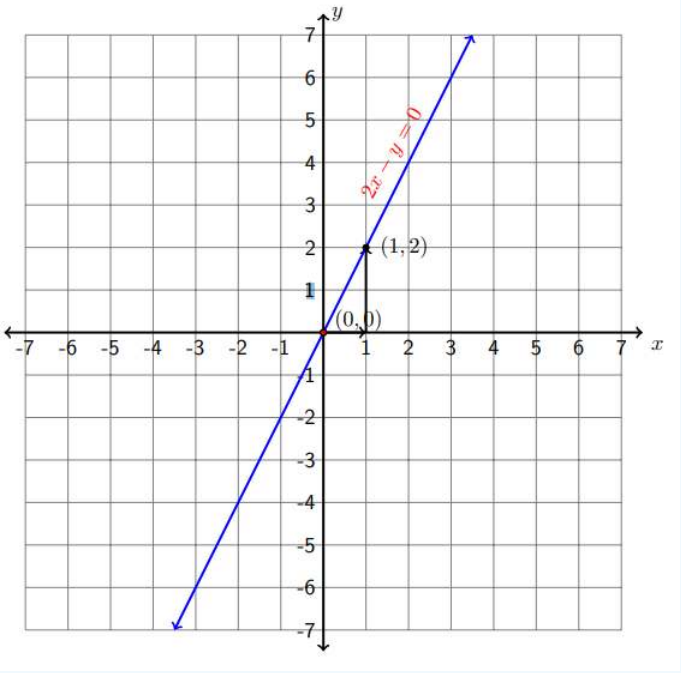
$$2x - y = 0 \quad \Longrightarrow \quad -y = -2x \quad \Longrightarrow \quad y = 2x$$

Here the slope is 2 and the $y-$ intercept is $(0, 0)$.

Step 2 Plot the y -intercept (0,0) on the y -axis, which is the origin of the coordinate system.

Step 3 Using the slope $m = 2 = \frac{2}{1} = \frac{\text{rise}}{\text{run}}$ to locate the second point, that is, from the y -intercept we move to the right (run) one unit and up (rise) two units, plot a second point at $(1, 2)$.

Step 4 Draw a line passing through the two points.



Point-Slope Form of the Equation of a Line $y - y_0 = m(x - x_0)$

Suppose we are given only the slope of a line and one point on the line and are asked to find its equation. Even though we do not know the y -intercept offhand, we can still find the equation of the line by using what's called the point-slope formula (or form).

The point-slope form of the equation of a line

$$y - y_0 = m(x - x_0)$$

is called the point-slope form of the equation of a line of slope m passing through the point (x_0, y_0).

Example 26.6

Find the equation of the line passing through the point (1,-2) with slope $\frac{1}{5}$.

We use the point-slope formula here, with $m = \frac{1}{5}, x_0 = 1$ and $y_0 = -2$ So, the equation of the line will be given by $y - (-2) = \frac{1}{5}(x - 1)$ $y + 2 = \frac{1}{5}x - \frac{1}{5}$, so $y = \frac{1}{5}x - \frac{1}{5} - 2$, and

$$y = \frac{1}{5}x - \frac{11}{5}$$

Note: The equation $y = \frac{1}{5}x - \frac{11}{5}$ is written in the slope-intercept form, therefore we can read off the y -intercept if we wish: $\left(0, -\frac{11}{5}\right)$.

Example 26.7

Find the equation of the line passing through the points $(3, 2)$ and $(-4, 3)$.

Notice that here we are *not* given the slope. We need to find it before we can proceed to find the equation of the line.

The slope is $m = \frac{3-2}{-4-3} = -\frac{1}{7}$.

Now, we can choose any of the given points as our one point and use the point-slope form to find the equation of the line. Let's use $(3, 2)$.

So, with $m = -\frac{1}{7}, x_0 = 3$ and $y_0 = 2$, the equation of the line will be given by

$$y-2=-\frac{1}{7}(x-3)$$

$$y-2=-\frac{1}{7}x+\frac{3}{7}=-\frac{1}{7}x+\frac{3}{7}+2$$

and, finally,

$$y=-\frac{1}{7}x+\frac{17}{7}$$

Note: Again, if we wish to read off the $y-$ intercept, it is $\left(0,\ \frac{17}{7}\right)$.

Equations of Horizontal and Vertical Lines

Recall that all horizontal lines have slope 0.

Equation of a Horizontal Line

The equation of a horizontal line passing through the point (a,b) is $y=b$

Recall that all vertical lines have an undefined slope.

Equation of a Vertical Line

The equation of a vertical line passing through the point (a,b) is $x=a$

Example 26.8

a) The equation of the horizontal line passing through the point (3,-7) is $y=-7$.

b) The equation of the vertical line passing through the point $\left(\frac{2}{5},-1\right)$ is $x=\frac{2}{5}$.

c) The equation of the horizontal line passing through the point (4,0) is $y=0$.

d) The equation of the vertical line passing through the point $\left(-\frac{1}{2},0\right)$ is $x=-\frac{1}{2}$.

Exit Problem

Indicate the slope and the y -intercept, and graph the equation: $6x+4y=6$

2.4.1: Graphing Linear Equations is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **1.28: Graphing Linear Equations** is licensed CC BY-NC-ND 4.0.

2.5: Linear Regression

2.5: Linear Regression is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

CHAPTER OVERVIEW

3: Systems and Matrices

- 3.1: Systems of Linear Equations
  - 3.1.1: Solving a System of Equations Algebraically
- 3.2: Matrices
  - 3.2.1: Gaussian Elimination
- 3.3: Solving Using Matrices
  - 3.3.1: Gaussian Elimination
- 3.4: Matrix Algebra
  - 3.4.1: Matrix Arithmetic
  - 3.4.2: Multiplication of Matrices
  - 3.4.3: The ijth Entry of a Product
- 3.5: Matrix Inverses
  - 3.5.1: The Identity and Inverses
  - 3.5.2: Finding the Inverse of a Matrix
  - 3.5.3: The Matrix Inverse
    - 3.5.3.1: Exercises 2.6
- 3.6: Linear Inequalities
  - 3.6.1: Solving Systems of Linear Inequalities (Two Variables)
- 3.7: Linear Programming
  - 3.7.1: Maximization Applications
    - 3.7.1.1: Maximization Applications (Exercises)
  - 3.7.2: Minimization Applications

3: Systems and Matrices is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

3.1: Systems of Linear Equations

3.1: Systems of Linear Equations is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

3.1.1: Solving a System of Equations Algebraically

Suppose that Adam has 7 bills, all fives and tens, and that their total value is $\$40$. How many of each bill does he have? In order to solve such a problem we must first define variables.

Let x be the number of five dollar bills.

Let y be the number of ten dollar bills.

Next, we write equations that describe the situation:

$5x+10y=40$: The combined value of the bills is $\$40$.

That is, we must solve the following system of two linear equations in two variables (unknowns):

$x+y=7$: Adam has 7 bills.

$5x+10y=40$: The combined value of the bills is $\$40$.

That is, we must solve the following system of two linear equations in two variables (unknowns):

$$\begin{pmatrix} x+y=7 \\ 5x+10y=40 \end{pmatrix}$$

This chapter deals with solving systems of two linear equations with two variable, such as the one above. We will consider two different algebraic methods: the substitution method and the elimination method.

The Substitution Method

In this section we solve systems of two linear equations in two variables using the **substitution method**. To illustrate, we will solve the system above with this method. We begin by solving the first equation for one variable in terms of the other. In this case we will solve for the variable y in terms of x:

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

Next, we **substitute** $y=7-x$ into the second equation $5x+10y=40$:

$$5x+10(7-x)=40$$

The equation above can now be solved for x since it only involves one variable:

$$\begin{aligned} 5x+10(7-x) &= 40 \\ 5x+70-10x &= 40 \quad \text{distribute 10 into the parentheses} \\ -5x+70 &= 40 \quad \text{collect like terms} \\ -5x &= -30 \quad \text{subtract 70 from both sides} \\ x &= 6 \quad \text{divide both sides by } -5 \end{aligned}$$

Hence, we get $x=6$. To find y, we substitute $x=6$ into the first equation of the system and solve for y (Note: We may substitute $x=6$ into either of the two original equations or the equation $y=7-x$):

$$\begin{aligned} 6+y&=7 \\ y&=1 \text{subtract 6 from both sides} \end{aligned}$$

Therefore the solution to the system of linear equations is

$$x=6, \quad y=1$$

Before we are truly finished, we should **check** our solution. The solution of a system of equations are the values of its variables which, when substituted into the two original equations, give us true statements. So to check, we substitute $x=6$ and $y=1$ into each equation of the system:

$$\begin{aligned} &x+y=7 \Longrightarrow 6+1=7 \Longrightarrow 7=7 \text{ true!} \\ &5x+10y=40 \Longrightarrow 5(6)+10(1)=40 \Longrightarrow 30+10=40 \Longrightarrow 40=40 \text{ true!} \end{aligned}$$

Hence, our solution is correct. To answer the original word problem - recalling that x is the number of five dollar bills and y is the number of ten dollar bills we have that:

$$Adam\ has\ 6\ five\ dollar\ bills\ and\ 1\ ten\ dollar\ bill.$$

Example 27.1

$$\begin{pmatrix} 6x+2y=72 \\ 3x+8y=78 \end{pmatrix}$$

Again, here we solve the system of equations using substitution. First, solve the first equation $6x+2y=72$ for y :

$$\begin{aligned} & 6x+2y=72 \\ \Longrightarrow \quad & 2y=-6x+72 \quad \text{subtract 6x from both sides} \\ \Longrightarrow \quad & y=-3x+36 \quad\quad \text{divide both sides by 2} \end{aligned}$$

Substitute $y=-3x+36$ into the second equation $3x+8y=78$:

$$\begin{aligned} & 3x+8y=78 \\ \Longrightarrow & 3x+8(-3x+36)=78 \\ \Longrightarrow & x=10 \end{aligned}$$

Hence $x=10$. Now substituting $x=10$ into the equation $y=-3x+36$ yields $y=6$, so the solution to the system of equations is $x=10, y=6$. The final step is left for the reader. It must be checked that $x=10$ and $y=6$ give true statements when substituted into the original system of equations.

To summarize the steps we followed to solve a system of linear equations in two variables using the algebraic method of substitution, we have:

Solving a System of Two Linear Equations in Two Variables using Substitution

1. Solve one equation for one variable.
2. Substitute the expression found in step 1 into the other equation.
3. Solve the resulting equation.
4. Substitute the value from step 3 back into the equation in step 1 to find the value of the remaining variable. 5
5. Check your solution!
6. Answer the question if it is a word problem.

A system of two linear equations in two variables may have one solution, no solutions, or infinitely many solutions. In Example 27.2 we will see a system with no solution.

Example 27.2

Solve the following system of equations by substitution.

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

The second equation is already solved for y in terms of x so we can substitute it directly into $x+y=1$:

$$x+(-x+2)=1 \Longrightarrow 2=1 \quad \text{False!}$$

Since we get the **false** statement $2=1$, the system of equations has **no** solution.

The Elimination Method

A second algebraic method for solving a system of linear equations is **the elimination method**. The basic idea of the method is to get the coefficients of one of the variables in the two equations to be additive inverses, such as -3 and 3, so that after the two

equations are added, this variable is eliminated. Let's use one of the systems we solved in the previous section in order to illustrate the method:

$$\begin{pmatrix} x+y=7 \\ 5x+10y=40 \end{pmatrix}$$

The coefficients of the x variable in our two equations are 1 and 5. We can make the coefficients of x to be additive inverses by multiplying the first equation by -5 and keeping the second equation untouched:

$$\begin{pmatrix} x & + & y & = & 7 \\ 5x & + & 10y & = & 40 \end{pmatrix} \Longrightarrow \begin{pmatrix} (-5)(x & + & y) & = & (-5)7 \\ 5x & + & 10y & = & 40 \end{pmatrix}$$

Using the distributive property, we rewrite the first equation as:

$$-5x-5y=-35$$

Now we are ready to add the two equations to eliminate the variable x and solve the resulting equation for y :

$$\begin{array}{rrrrrr} & -5x & - & 5y & = & -35 \\ + & 5x & + & 10y & = & 40 \\ \hline & & & 5y & = & 5 \\ & & \Longrightarrow & y & = & 1 \end{array}$$

To find x, we can substitute $y=1$ into either equation of the original system to solve for x :

$$x+1=7 \quad \Longrightarrow \quad x=6$$

Hence, we get the same solution as we obtained using the substitution method in the previous section:

$$x=6, \quad y=1$$

In this example, we only need to multiply the first equation by a number to make the coefficients of the variable x additive inverses. Sometimes, we need to multiply both equations by two different numbers to make the coefficients of one of the variables additive inverses. To illustrate this, let's look at Example 27.3.

Example 27.3

Lets aim to eliminate the y variable here. Since the least common multiple of 2 and 3 is 6, we can multiply the first equation by 3 and the second equation by 2, so that the coefficients of y are additive inverses:

$$\begin{pmatrix} -3x & + & 2y & = & 3 \\ 4x & - & 3y & = & -6 \end{pmatrix} \Longrightarrow \begin{pmatrix} (3)(-3x & + & 2y & = & (3)3 \\ (2)(4x & - & 3y & = & (2)(-6) \end{pmatrix}$$

Using the distributive property, we rewrite the two equations as:

$$\begin{pmatrix} -9x & + & 6y & = & 9 \\ 8x & - & 6y & = & -12 \end{pmatrix}$$

Adding them together gives:

$$-1x=-3 \quad \Longrightarrow \quad x=3$$

To find y, we can substitute $x=3$ into the first equation (or the second equation) of the original system to solve for y :

$$-3(3)+2y=3 \Longrightarrow -9+2y=3 \Longrightarrow 2y=12 \Longrightarrow y=6$$

Hence, the answer to the problem is:

$$x=3 \quad y=6$$

We can check the answer by substituting both numbers into the original system and see if both equations are correct.

The following steps summarize how to solve a system of equations by the elimination method:

Solving a System of Two Linear Equations in Two Variables using Elimination

1. Multiply one or both equations by a nonzero number so that the coefficients of one of the variables are additive inverses.
2. Add the equations to eliminate the variable.
3. Solve the resulting equation.
4. Substitute the value from step 3 back into either of the original equations to find the value of the remaining variable.
5. Check your solution!
6. Answer the question if it is a word problem.

Exit Problem

Solve algebraically:

$$\begin{aligned} 8x &- 4y = 4 \\ 3x &- 2y = 3 \end{aligned}$$

This page titled 3.1.1: Solving a System of Equations Algebraically is shared under a CC BY-NC-ND 4.0 license and was authored, remixed, and/or curated by Samar ElHitti, Marianna Bonanome, Holly Carley, Thomas Tradler, & Lin Zhou (New York City College of Technology at CUNY Academic Works) via source content that was edited to the style and standards of the LibreTexts platform.

- **1.29: Solving a System of Equations Algebraically** by Samar ElHitti, Marianna Bonanome, Holly Carley, Thomas Tradler, & Lin Zhou is licensed CC BY-NC-ND 4.0. Original source: https://academicworks.cuny.edu/ny_oers/5/.

3.2: Matrices

3.2: Matrices is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

3.2.1: Gaussian Elimination

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

For example, consider the linear system in Example 1.2.3

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

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

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

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

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

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

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

Consider the following definition.

Definition 3.2.1.1: Augmented Matrix of a Linear System

For a linear system of the form

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

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

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

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

Definition 3.2.1.2: Elementary Row Operations

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

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

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

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

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

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

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

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

This augmented matrix corresponds to the system

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

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

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

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

Definition 3.2.1.3: Row-Echelon Form

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

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

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

Definition 3.2.1.4: Reduced Row-Echelon Form

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

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

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

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

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

Example 3.2.1.1: Not in Row-Echelon Form

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

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

Example 3.2.1.2: Matrices in Row-Echelon Form

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

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

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

Example 3.2.1.3: Matrices in Reduced Row-Echelon Form

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

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

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

Definition 3.2.1.5: Pivot Position and Pivot Column

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

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

For example consider the following.

Example 3.2.1.4: Pivot Position

Let

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

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

Solution

The row-echelon form of this matrix is

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

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

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

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

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

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

Algorithm 3.2.1.1: Reduced Row-Echelon Form Algorithm

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

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

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

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

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

Consider the following example of Algorithm 3.2.1.1.

Example 3.2.1.5: Finding Row-Echelon Form and Reduced Row-Echelon Form of a Matrix

Let

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

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

Solution

In working through this example, we will use the steps outlined in Algorithm 3.2.1.1.

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

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

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

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

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

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

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

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

Our original matrix now looks like

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

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

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

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

This matrix is now in row-echelon form.

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

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

This matrix is now in reduced row-echelon form.

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

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

Example 3.2.1.6: Finding the Solution to a System

Give the complete solution to the following system of equations

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

Solution

The augmented matrix for this system is

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

In order to find the solution to this system, we wish to carry the augmented matrix to reduced row-echelon form. We will do so using Algorithm 3.2.1.1. Notice that the first column is nonzero, so this is our first pivot column. The first entry in the first row, 2, is the first leading entry and it is in the first pivot position. We will use row operations to create zeros in the entries below the 2. First, replace the second row with -5 times the first row plus 2 times the second row. This yields

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

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

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

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

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

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

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

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

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

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

Example 3.2.1.7: An Infinite Set of Solutions

Give the complete solution to the system of equations

$$\begin{aligned} 3x-y-5z&=9\\ y-10z&=0\\ -2x+y&=-6 \end{aligned} \tag{3.2.1.1}$$

Solution

The augmented matrix of this system is

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

In order to find the solution to this system, we will carry the augmented matrix to reduced row-echelon form, using Algorithm 3.2.1.1. The first column is the first pivot column. We want to use row operations to create zeros beneath the first entry in this column, which is in the first pivot position. Replace the third row with 2 times the first row added to 3 times the third row. This gives

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

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

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

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

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

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

This matrix is now in row-echelon form.

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

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

This is in reduced row-echelon form, which you should verify using Definition 3.2.1.4. The equations corresponding to this reduced row-echelon form are

$$\begin{aligned} x - 5z &= 3 \\ y - 10z &= 0 \end{aligned}$$

or

$$\begin{aligned} x &= 3 + 5z \\ y &= 10z \end{aligned}$$

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

$$\begin{aligned} x &= 3 + 5t \\ y &= 10t \\ z &= t \end{aligned}$$

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

$$\begin{aligned} x &= 3 + 5(4) = 23 \\ y &= 10(4) = 40 \\ z &= 4 \end{aligned}$$

In Example 3.2.1.7 the solution involved one parameter. It may happen that the solution to a system involves more than one parameter, as shown in the following example.

Example 3.2.1.8: A Two Parameter Set of Solutions

Find the solution to the system

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

Solution

The augmented matrix is

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

We wish to carry this matrix to row-echelon form. Here, we will outline the row operations used. However, make sure that you understand the steps in terms of Algorithm 3.2.1.1.

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

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

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

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{3.2.1.2}$$

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

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

It is customary to write this solution in the form

$$\left[\begin{array}{c} x \\ y \\ z \\ w \end{array}\right] = \left[\begin{array}{c} -1+s-t \\ 2 \\ s \\ t \end{array}\right] \tag{3.2.1.3}$$

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

In the above steps, this means we don't stop with the row-echelon form in equation (3.2.1.2). Instead we first place it in reduced row-echelon form as follows.

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

Then the solution is $y=2$ from the second row and $x=-1+z-w$ from the first. Thus letting $z=s$ and $w=t$, the solution is given by (3.2.1.3).

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

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

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

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

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

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

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

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

or

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

This page titled 3.2.1: Gaussian Elimination is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) via source content that was edited to the style and standards of the LibreTexts platform.

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

3.3: Solving Using Matrices

3.3: Solving Using Matrices is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

3.3.1: Gaussian Elimination

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

For example, consider the linear system in Example 1.2.3

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

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

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

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

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

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

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

Consider the following definition.

Definition 3.3.1.1: Augmented Matrix of a Linear System

For a linear system of the form

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

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

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

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

Definition 3.3.1.2: Elementary Row Operations

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

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

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

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

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

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

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

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

This augmented matrix corresponds to the system

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

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

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

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

Definition 3.3.1.3: Row-Echelon Form

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

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

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

Definition 3.3.1.4: Reduced Row-Echelon Form

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

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

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

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

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

Example 3.3.1.1: Not in Row-Echelon Form

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

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

Example 3.3.1.2: Matrices in Row-Echelon Form

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

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

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

Example 3.3.1.3: Matrices in Reduced Row-Echelon Form

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

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

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

Definition 3.3.1.5: Pivot Position and Pivot Column

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

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

For example consider the following.

Example 3.3.1.4: Pivot Position

Let

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

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

Solution

The row-echelon form of this matrix is

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

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

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

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

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

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

Algorithm 3.3.1.1: Reduced Row-Echelon Form Algorithm

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

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

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

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

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

Consider the following example of Algorithm 3.3.1.1.

Example 3.3.1.5: Finding Row-Echelon Form and Reduced Row-Echelon Form of a Matrix

Let

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

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

Solution

In working through this example, we will use the steps outlined in Algorithm 3.3.1.1.

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

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

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

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

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

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

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

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

Our original matrix now looks like

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

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

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

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

This matrix is now in row-echelon form.

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

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

This matrix is now in reduced row-echelon form.

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

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

Example 3.3.1.6: Finding the Solution to a System

Give the complete solution to the following system of equations

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

Solution

The augmented matrix for this system is

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

In order to find the solution to this system, we wish to carry the augmented matrix to reduced row-echelon form. We will do so using Algorithm 3.3.1.1. Notice that the first column is nonzero, so this is our first pivot column. The first entry in the first row, 2, is the first leading entry and it is in the first pivot position. We will use row operations to create zeros in the entries below the 2. First, replace the second row with -5 times the first row plus 2 times the second row. This yields

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

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

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

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

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

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

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

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

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

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

Example 3.3.1.7: An Infinite Set of Solutions

Give the complete solution to the system of equations

$$\begin{aligned} 3x-y-5z&=9\\ y-10z&=0\\ -2x+y&=-6 \end{aligned} \tag{3.3.1.1}$$

Solution

The augmented matrix of this system is

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

In order to find the solution to this system, we will carry the augmented matrix to reduced row-echelon form, using Algorithm 3.3.1.1. The first column is the first pivot column. We want to use row operations to create zeros beneath the first entry in this column, which is in the first pivot position. Replace the third row with 2 times the first row added to 3 times the third row. This gives

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

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

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

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

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

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

This matrix is now in row-echelon form.

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

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

This is in reduced row-echelon form, which you should verify using Definition 3.3.1.4. The equations corresponding to this reduced row-echelon form are

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

or

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

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

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

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

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

In Example 3.3.1.7 the solution involved one parameter. It may happen that the solution to a system involves more than one parameter, as shown in the following example.

Example 3.3.1.8: A Two Parameter Set of Solutions

Find the solution to the system

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

Solution

The augmented matrix is

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

We wish to carry this matrix to row-echelon form. Here, we will outline the row operations used. However, make sure that you understand the steps in terms of Algorithm 3.3.1.1.

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

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

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

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 3 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \tag{3.3.1.2}$$

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

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

It is customary to write this solution in the form

$$\left[\begin{array}{c} x \\ y \\ z \\ w \end{array}\right] = \left[\begin{array}{c} -1+s-t \\ 2 \\ s \\ t \end{array}\right] \tag{3.3.1.3}$$

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

In the above steps, this means we don't stop with the row-echelon form in equation (3.3.1.2). Instead we first place it in reduced row-echelon form as follows.

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

Then the solution is $y=2$ from the second row and $x=-1+z-w$ from the first. Thus letting $z=s$ and $w=t$, the solution is given by (3.3.1.3).

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

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

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

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

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

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

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

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

or

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

This page titled 3.3.1: Gaussian Elimination is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) via source content that was edited to the style and standards of the LibreTexts platform.

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

3.4: Matrix Algebra

3.4: Matrix Algebra is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

3.4.1: Matrix Arithmetic

Outcomes

A. Perform the matrix operations of matrix addition, scalar multiplication, transposition and matrix multiplication. Identify when these operations are not defined. Represent these operations in terms of the entries of a matrix.
B. Prove algebraic properties for matrix addition, scalar multiplication, transposition, and matrix multiplication. Apply these properties to manipulate an algebraic expression involving matrices.
C. Compute the inverse of a matrix using row operations, and prove identities involving matrix inverses.
D. Solve a linear system using matrix algebra.
E. Use multiplication by an elementary matrix to apply row operations.
F. Write a matrix as a product of elementary matrices.

You have now solved systems of equations by writing them in terms of an augmented matrix and then doing row operations on this augmented matrix. It turns out that matrices are important not only for systems of equations but also in many applications.

Recall that a **matrix** is a rectangular array of numbers. Several of them are referred to as **matrices**. For example, here is a matrix.

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 2 & 8 & 7 \\ 6 & -9 & 1 & 2 \end{bmatrix} \tag{3.4.1.1}$$

Recall that the size or dimension of a matrix is defined as $m \times n$ where m is the number of rows and n is the number of columns. The above matrix is a 3×4 matrix because there are three rows and four columns. You can remember the columns are like columns in a Greek temple. They stand upright while the rows lay flat like rows made by a tractor in a plowed field.

When specifying the size of a matrix, you always list the number of rows before the number of columns.You might remember that you always list the rows before the columns by using the phrase **Row**man **C**atholic.

Consider the following definition.

A matrix A which has size $n \times n$ is called a **square matrix** . In other words, A is a square matrix if it has the same number of rows and columns.

There is some notation specific to matrices which we now introduce. We denote the columns of a matrix A by A_j as follows

$$A = \begin{bmatrix} A_1 & A_2 & \cdots & A_n \end{bmatrix}$$

Therefore, A_j is the jth column of A, when counted from left to right.

The individual elements of the matrix are called **entries** or **components** of A. Elements of the matrix are identified according to their position. The $(\mathbf{i,j})$**-entry** of a matrix is the entry in the ith row and j^{th} column. For example, in the matrix (3.4.1.1)above, 8 is in position $(2,3)$ (and is called the $(2,3)$-entry) because it is in the second row and the third column.

In order to remember which matrix we are speaking of, we will denote the entry in the ith row and the jth column of matrix A by a_{ij}. Then, we can write A in terms of its entries, as $A = [a_{ij}]$. Using this notation on the matrix in (3.4.1.1), $a_{23} = 8, a_{32} = -9, a_{12} = 2,$ etc.

There are various operations which are done on matrices of appropriate sizes. Matrices can be added to and subtracted from other matrices, multiplied by a scalar, and multiplied by other matrices. We will never divide a matrix by another matrix, but we will see later how matrix inverses play a similar role.

In doing arithmetic with matrices, we often define the action by what happens in terms of the entries (or components) of the matrices. Before looking at these operations in depth, consider a few general definitions.

The $m \times n$ **zero matrix** is the $m \times n$ matrix having every entry equal to zero. It is denoted by 0.

One possible zero matrix is shown in the following example.

Example 3.4.1.1: The Zero Matrix

The 2×3 zero matrix is $0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Note there is a 2×3 zero matrix, a 3×4 zero matrix, etc. In fact there is a zero matrix for every size!

Let A and B be two $m \times n$ matrices. Then $A = B$ means that for $A = [a_{ij}]$ and $B = [b_{ij}]$, $a_{ij} = b_{ij}$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$.

In other words, two matrices are equal exactly when they are the same size and the corresponding entries are identical. Thus

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

because they are different sizes. Also,

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

because, although they are the same size, their corresponding entries are not identical.

In the following section, we explore addition of matrices.

Addition of Matrices

When adding matrices, all matrices in the sum need have the same size. For example,

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

and

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

cannot be added, as one has size 3×2 while the other has size 2×3.

However, the addition

$$\begin{bmatrix} 4 & 6 & 3 \\ 5 & 0 & 4 \\ 11 & -2 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 5 & 0 \\ 4 & -4 & 14 \\ 1 & 2 & 6 \end{bmatrix}$$

is possible.

The formal definition is as follows.

Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be two $m \times n$ matrices. Then $A + B = C$ where C is the $m \times n$ matrix $C = [c_{ij}]$ defined by

$$c_{ij} = a_{ij} + b_{ij}$$

This definition tells us that when adding matrices, we simply add corresponding entries of the matrices. This is demonstrated in the next example.

Add the following matrices, if possible.

$$A=\begin{bmatrix}1 & 2 & 3\\ 1 & 0 & 4\end{bmatrix}, B=\begin{bmatrix}5 & 2 & 3\\ -6 & 2 & 1\end{bmatrix}$$

Solution

Notice that both A and B are of size 2×3. Since A and B are of the same size, the addition is possible. Using Definition 3.4.1.4, the addition is done as follows.

$$A+B=\begin{bmatrix}1 & 2 & 3\\ 1 & 0 & 4\end{bmatrix}+\begin{bmatrix}5 & 2 & 3\\ -6 & 2 & 1\end{bmatrix}=\begin{bmatrix}1+5 & 2+2 & 3+3\\ 1+-6 & 0+2 & 4+1\end{bmatrix}=\begin{bmatrix}6 & 4 & 6\\ -5 & 2 & 5\end{bmatrix}$$

Addition of matrices obeys very much the same properties as normal addition with numbers. Note that when we write for example $A+B$ then we assume that both matrices are of equal size so that the operation is indeed possible.

Proposition 3.4.1.1: Properties of Matrix Addition

Let A, B and C be matrices. Then, the following properties hold.

- Commutative Law of Addition

$$A+B=B+A \tag{3.4.1.2}$$

- Associative Law of Addition

$$(A+B)+C=A+(B+C) \tag{3.4.1.3}$$

- Existence of an Additive Identity

$$\begin{gathered}\text{There exists a zero matrix } 0 \text{ such that}\\ A+0=A\end{gathered} \tag{3.4.1.4}$$

- Existence of an Additive Inverse

$$\begin{gathered}\text{There exists a matrix } -A \text{ such that}\\ A+(-A)=0\end{gathered} \tag{3.4.1.5}$$

Proof

Consider the Commutative Law of Addition given in (3.4.1.2) Let A, B, C, and D be matrices such that $A+B=C$ and $B+A=D$. We want to show that $D=C$. To do so, we will use the definition of matrix addition given in Definition 3.4.1.4 Now,

$$c_{ij}=a_{ij}+b_{ij}=b_{ij}+a_{ij}=d_{ij}$$

Therefore, $C=D$ because the ij^{th} entries are the same for all i and j. Note that the conclusion follows from the commutative law of addition of numbers, which says that if a and b are two numbers, then $a+b=b+a$. The proof of the other results are similar, and are left as an exercise.

We call the zero matrix in (3.4.1.4) the **additive identity**. Similarly, we call the matrix $-A$ in (3.4.1.5) the **additive inverse**. $-A$ is defined to equal $(-1)A=[-a_{ij}]$. In other words, every entry of A is multiplied by -1.

In the next section we will study scalar multiplication in more depth to understand what is meant by $(-1)A$.

Scalar Multiplication of Matrices

Recall that we use the word *scalar* when referring to numbers. Therefore, *scalar multiplication of a matrix* is the multiplication of a matrix by a number. To illustrate this concept, consider the following example in which a matrix is multiplied by the scalar 3.

$$3\begin{bmatrix}1 & 2 & 3 & 4\\ 5 & 2 & 8 & 7\\ 6 & -9 & 1 & 2\end{bmatrix}=\begin{bmatrix}3 & 6 & 9 & 12\\ 15 & 6 & 24 & 21\\ 18 & -27 & 3 & 6\end{bmatrix}$$

The new matrix is obtained by multiplying every entry of the original matrix by the given scalar.

The formal definition of scalar multiplication is as follows.

Definition 3.4.1.5: Scalar Multiplication of Matrices

If $A = [a_{ij}]$ and k is a scalar, then $kA = [ka_{ij}]$.

Consider the following example.

Example 3.4.1.3: Effect of Multiplication by a Scalar

Find the result of multiplying the following matrix A by 7.

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

Solution

By Definition 3.4.1.5, we multiply each element of A by 7. Therefore,

$$7A = 7\begin{bmatrix} 2 & 0 \\ 1 & -4 \end{bmatrix} = \begin{bmatrix} 7(2) & 7(0) \\ 7(1) & 7(-4) \end{bmatrix} = \begin{bmatrix} 14 & 0 \\ 7 & -28 \end{bmatrix}$$

Similarly to addition of matrices, there are several properties of scalar multiplication which hold.

Proposition 3.4.1.2: Properties of Scalar Multiplication

Let A, B be matrices, and k, p be scalars. Then, the following properties hold.

- Distributive Law over Matrix Addition

$$k(A+B) = kA + kB$$

- Distributive Law over Scalar Addition

$$(k+p)A = kA + pA$$

- Associative Law for Scalar Multiplication

$$k(pA) = (kp)A$$

- Rule for Multiplication by 1

$$1A = A$$

Proof

The proof of this proposition is similar to the proof of Proposition 3.4.1.1and is left an exercise to the reader.

This page titled 3.4.1: Matrix Arithmetic is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) via source content that was edited to the style and standards of the LibreTexts platform.

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

3.4.2: Multiplication of Matrices

The next important matrix operation we will explore is multiplication of matrices. The operation of matrix multiplication is one of the most important and useful of the matrix operations. Throughout this section, we will also demonstrate how matrix multiplication relates to linear systems of equations.

First, we provide a formal definition of row and column vectors.

Definition 3.4.2.1: Row and Column Vectors

Matrices of size $n \times 1$ or $1 \times n$ are called **vectors**. If X is such a matrix, then we write x_i to denote the entry of X in the ith row of a column matrix, or the ith column of a row matrix.

The $n \times 1$ matrix

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is called a **column vector.** The $1 \times n$ matrix

$$X = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

is called a **row vector**.

We may simply use the term **vector** throughout this text to refer to either a column or row vector. If we do so, the context will make it clear which we are referring to.

In this chapter, we will again use the notion of linear combination of vectors as in Definition 9.2.2. In this context, a linear combination is a sum consisting of vectors multiplied by scalars. For example,

$$\begin{bmatrix} 50 \\ 122 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + 8 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 9 \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

is a linear combination of three vectors.

It turns out that we can express any system of linear equations as a linear combination of vectors. In fact, the vectors that we will use are just the columns of the corresponding augmented matrix!

Definition 3.4.2.2: The Vector Form of a System of Linear Equations

Suppose we have a system of equations given by

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

We can express this system in **vector form** which is as follows:

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

Notice that each vector used here is one column from the corresponding augmented matrix. There is one vector for each variable in the system, along with the constant vector.

The first important form of matrix multiplication is multiplying a matrix by a vector. Consider the product given by

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

We will soon see that this equals

$$7 \begin{bmatrix} 1 \\ 4 \end{bmatrix} + 8 \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 9 \begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 50 \\ 122 \end{bmatrix}$$

In general terms,

$$\begin{aligned} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{bmatrix} \end{aligned}$$

Thus you take x_1 times the first column, add to x_2 times the second column, and finally x_3 times the third column. The above sum is a linear combination of the columns of the matrix. When you multiply a matrix on the left by a vector on the right, the numbers making up the vector are just the scalars to be used in the linear combination of the columns as illustrated above.

Here is the formal definition of how to multiply an $m \times n$ matrix by an $n \times 1$ column vector.

Definition 3.4.2.3: Multiplication of Vector by Matrix

Let $A = [a_{ij}]$ be an $m \times n$ matrix and let X be an $n \times 1$ matrix given by

$$A = [A_1 \cdots A_n], X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

Then the product AX is the $m \times 1$ column vector which equals the following linear combination of the columns of A:

$$x_1 A_1 + x_2 A_2 + \cdots + x_n A_n = \sum_{j=1}^{n} x_j A_j$$

If we write the columns of A in terms of their entries, they are of the form

$$A_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

Then, we can write the product AX as

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

Note that multiplication of an $m \times n$ matrix and an $n \times 1$ vector produces an $m \times 1$ vector.

Here is an example.

Example 3.4.2.1: A Vector Multiplied by a Matrix

Compute the product AX for

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

Solution

We will use Definition 3.4.2.3 to compute the product. Therefore, we compute the product AX as follows.

$$\begin{aligned} &1\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} + 1\begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 2 \\ 5 \end{bmatrix} \end{aligned}$$

Using the above operation, we can also write a system of linear equations in **matrix form**. In this form, we express the system as a matrix multiplied by a vector. Consider the following definition.

Definition 3.4.2.4: The Matrix Form of a System of Linear Equations

Suppose we have a system of equations given by

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

Then we can express this system in **matrix form** as follows.

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

The expression $AX = B$ is also known as the **Matrix Form** of the corresponding system of linear equations. The matrix A is simply the coefficient matrix of the system, the vector X is the column vector constructed from the variables of the system, and finally the vector B is the column vector constructed from the constants of the system. It is important to note that any system of linear equations can be written in this form.

Notice that if we write a homogeneous system of equations in matrix form, it would have the form $AX = 0$, for the zero vector 0.

You can see from this definition that a vector

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

will satisfy the equation $AX = B$ only when the entries $x_1, x_2, \cdots, x_n$ of the vector X are solutions to the original system.

Now that we have examined how to multiply a matrix by a vector, we wish to consider the case where we multiply two matrices of more general sizes, although these sizes still need to be appropriate as we will see. For example, in Example 3.4.2.1, we multiplied a 3×4 matrix by a 4×1 vector. We want to investigate how to multiply other sizes of matrices.

We have not yet given any conditions on when matrix multiplication is possible! For matrices A and B, in order to form the product AB, the number of columns of A must equal the number of rows of B. Consider a product AB where A has size $m \times n$ and B has size $n \times p$. Then, the product in terms of size of matrices is given by

$$(m \times \overbrace{\quad n)\,(n}^{\text{these must match!}} \times p \quad) = m \times p$$

Note the two outside numbers give the size of the product. One of the most important rules regarding matrix multiplication is the following. If the two middle numbers don't match, you can't multiply the matrices!

When the number of columns of A equals the number of rows of B the two matrices are said to be **conformable** and the product AB is obtained as follows.

Definition 3.4.2.5: Multiplication of Two Matrices

Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix of the form

$$B = [B_1 \cdots B_p]$$

where $B_1, \ldots, B_p$ are the $n \times 1$ columns of B. Then the $m \times p$ matrix AB is defined as follows:

$$AB = A\,[B_1 \cdots B_p] = [(AB)_1 \cdots (AB)_p]$$

where $(AB)_k$ is an $m \times 1$ matrix or column vector which gives the kth column of AB.

Consider the following example.

Example 3.4.2.2: Multiplying Two Matrices

Find AB if possible.

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

Solution

The first thing you need to verify when calculating a product is whether the multiplication is possible. The first matrix has size 2×3 and the second matrix has size 3×3. The inside numbers are equal, so A and B are conformable matrices. According to the above discussion AB will be a 2×3 matrix. Definition 3.4.2.5 gives us a way to calculate each column of AB, as follows.

$$\begin{bmatrix} \overbrace{\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}}^{\text{First column}}, & \overbrace{\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}}^{\text{Second column}}, & \overbrace{\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}}^{\text{Third column}} \end{bmatrix}$$

You know how to multiply a matrix times a vector, using Definition 3.4.2.3 for each of the three columns. Thus

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \\ -2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 9 & 3 \\ -2 & 7 & 3 \end{bmatrix}$$

Since vectors are simply $n \times 1$ or $1 \times m$ matrices, we can also multiply a vector by another vector.

Example 3.4.2.3: Vector Times Vector Multiplication

Multiply if possible $\begin{bmatrix}1\\2\\1\end{bmatrix}\begin{bmatrix}1 & 2 & 1 & 0\end{bmatrix}$.

Solution

In this case we are multiplying a matrix of size 3×1 by a matrix of size 1×4. The inside numbers match so the product is defined. Note that the product will be a matrix of size 3×4. Using Definition 3.4.2.5, we can compute this product as follows

$$\begin{bmatrix}1\\2\\1\end{bmatrix}\begin{bmatrix}1 & 2 & 1 & 0\end{bmatrix}=\begin{bmatrix}\overbrace{\begin{bmatrix}1\\2\\1\end{bmatrix}\begin{bmatrix}1\end{bmatrix}}^{\text{First column}}, & \overbrace{\begin{bmatrix}1\\2\\1\end{bmatrix}\begin{bmatrix}2\end{bmatrix}}^{\text{Second column}}, & \overbrace{\begin{bmatrix}1\\2\\1\end{bmatrix}\begin{bmatrix}1\end{bmatrix}}^{\text{Third column}}, & \overbrace{\begin{bmatrix}1\\2\\1\end{bmatrix}\begin{bmatrix}0\end{bmatrix}}^{\text{Fourth column}}\end{bmatrix}$$

You can use Definition 3.4.2.3to verify that this product is

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

Example 3.4.2.4: A Multiplication Which is Not Defined

Find BA if possible.

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

Solution

First check if it is possible. This product is of the form $(3\times 3)(2\times 3)$. The inside numbers do not match and so you can't do this multiplication.

In this case, we say that the multiplication is not defined. Notice that these are the same matrices which we used in Example 3.4.2.2 In this example, we tried to calculate BA instead of AB. This demonstrates another property of matrix multiplication. While the product AB maybe be defined, we cannot assume that the product BA will be possible. Therefore, it is important to always check that the product is defined before carrying out any calculations.

Earlier, we defined the zero matrix 0 to be the matrix (of appropriate size) containing zeros in all entries. Consider the following example for multiplication by the zero matrix.

Example 3.4.2.5: Multiplication by the Zero Matrix

Compute the product $A0$ for the matrix

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

and the 2×2 zero matrix given by

$$0=\begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$$

Solution

In this product, we compute

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

Hence, $A0 = 0$.

Notice that we could also multiply A by the 2×1 zero vector given by $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. The result would be the 2×1 zero vector. Therefore, it is always the case that $A0 = 0$, for an appropriately sized zero matrix or vector.

This page titled 3.4.2: Multiplication of Matrices is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) via source content that was edited to the style and standards of the LibreTexts platform.

- **2.2: Multiplication of Matrices** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

3.4.3: The ijth Entry of a Product

In previous sections, we used the entries of a matrix to describe the action of matrix addition and scalar multiplication. We can also study matrix multiplication using the entries of matrices.

What is the ij^{th} entry of AB? It is the entry in the ith row and the jth column of the product AB.

Now if A is $m \times n$ and B is $n \times p$, then we know that the product AB has the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nj} & \cdots & b_{np} \end{bmatrix}$$

The jth column of AB is of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix}$$

which is an $m \times 1$ column vector. It is calculated by

$$b_{1j} \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + b_{2j} \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + b_{nj} \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

Therefore, the ij^{th} entry is the entry in row i of this vector. This is computed by

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

The following is the formal definition for the ij^{th} entry of a product of matrices.

Definition 3.4.3.1: The ij^{th} Entry of a Product

Let $A = [a_{ij}]$ be an $m \times n$ matrix and let $B = [b_{ij}]$ be an $n \times p$ matrix. Then AB is an $m \times p$ matrix and the (i, j)-entry of AB is defined as

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Another way to write this is

$$(AB)_{ij} = \begin{bmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{bmatrix} \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

In other words, to find the (i, j)-entry of the product AB, or $(AB)_{ij}$, you multiply the ith row of A, on the left by the jth column of B. To express AB in terms of its entries, we write $AB = [(AB)_{ij}]$.

Consider the following example.

Example 3.4.3.1: The Entries of a Product

Compute AB if possible. If it is, find the $(3,2)$-entry of AB using Definition 3.4.3.1.

$$A=\begin{bmatrix}1&2\\3&1\\2&6\end{bmatrix}, B=\begin{bmatrix}2&3&1\\7&6&2\end{bmatrix}$$

Solution

First check if the product is possible. It is of the form $(3\times 2)(2\times 3)$ and since the inside numbers match, it is possible to do the multiplication. The result should be a 3×3 matrix. We can first compute AB:

$$\left[\begin{bmatrix}1&2\\3&1\\2&6\end{bmatrix}\begin{bmatrix}2\\7\end{bmatrix}, \begin{bmatrix}1&2\\3&1\\2&6\end{bmatrix}\begin{bmatrix}3\\6\end{bmatrix}, \begin{bmatrix}1&2\\3&1\\2&6\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}\right]$$

where the commas separate the columns in the resulting product. Thus the above product equals

$$\begin{bmatrix}16&15&5\\13&15&5\\46&42&14\end{bmatrix}$$

which is a 3×3 matrix as desired. Thus, the $(3,2)$-entry equals 42.

Now using Definition 3.4.3.1, we can find that the $(3,2)$-entry equals

$$\begin{aligned}\sum_{k=1}^{2} a_{3k}b_{k2} &= a_{31}b_{12}+a_{32}b_{22}\\ &= 2\times 3+6\times 6=42\end{aligned}$$

Consulting our result for AB above, this is correct!

You may wish to use this method to verify that the rest of the entries in AB are correct.

Here is another example.

Example 3.4.3.2: Finding the Entries of a Product

Determine if the product AB is defined. If it is, find the $(2,1)$-entry of the product.

$$A=\begin{bmatrix}2&3&1\\7&6&2\\0&0&0\end{bmatrix}, B=\begin{bmatrix}1&2\\3&1\\2&6\end{bmatrix}$$

Solution

This product is of the form $(3\times 3)(3\times 2)$. The middle numbers match so the matrices are conformable and it is possible to compute the product.

We want to find the $(2,1)$-entry of AB, that is, the entry in the second row and first column of the product. We will use Definition 3.4.3.1, which states

$$(AB)_{ij}=\sum_{k=1}^{n} a_{ik}b_{kj}$$

In this case, $n=3$, $i=2$ and $j=1$. Hence the $(2,1)$-entry is found by computing

$$(AB)_{21}=\sum_{k=1}^{3} a_{2k}b_{k1}=\begin{bmatrix}a_{21}&a_{22}&a_{23}\end{bmatrix}\begin{bmatrix}b_{11}\\b_{21}\\b_{31}\end{bmatrix}$$

Substituting in the appropriate values, this product becomes

$$\begin{bmatrix} a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} = \begin{bmatrix} 7 & 6 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = 1 \times 7 + 3 \times 6 + 2 \times 2 = 29$$

Hence, $(AB)_{21} = 29$.

You should take a moment to find a few other entries of AB. You can multiply the matrices to check that your answers are correct. The product AB is given by

$$AB = \begin{bmatrix} 13 & 13 \\ 29 & 32 \\ 0 & 0 \end{bmatrix}$$

This page titled 3.4.3: The ijth Entry of a Product is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) via source content that was edited to the style and standards of the LibreTexts platform.

- **2.3: The ijth Entry of a Product** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

3.5: Matrix Inverses

3.5: Matrix Inverses is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

3.5.1: The Identity and Inverses

There is a special matrix, denoted I, which is called to as the **identity matrix**. The identity matrix is always a square matrix, and it has the property that there are ones down the main diagonal and zeroes elsewhere. Here are some identity matrices of various sizes.

$$[1], \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The first is the 1×1 identity matrix, the second is the 2×2 identity matrix, and so on. By extension, you can likely see what the $n \times n$ identity matrix would be. When it is necessary to distinguish which size of identity matrix is being discussed, we will use the notation I_n for the $n \times n$ identity matrix.

The identity matrix is so important that there is a special symbol to denote the ij^{th} entry of the identity matrix. This symbol is given by $I_{ij} = \delta_{ij}$ where δ_{ij} is the **Kronecker symbol** defined by

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$$

I_n is called the **identity matrix** because it is a **multiplicative identity** in the following sense.

Lemma 3.5.1.1: Multiplication by the Identity Matrix

Suppose A is an $m \times n$ matrix and I_n is the $n \times n$ identity matrix**.** Then $AI_n = A$. If I_m is the $m \times m$ identity matrix, it also follows that $I_m A = A$.

Proof

The (i, j)-entry of AI_n is given by:

$$\sum_k a_{ik}\delta_{kj} = a_{ij}$$

and so $AI_n = A$. The other case is left as an exercise for you.

We now define the matrix operation which in some ways plays the role of division.

Definition 3.5.1.1: The Inverse of a Matrix

A square $n \times n$ matrix A is said to have an **inverse** A^{-1} if and only if

$$AA^{-1} = A^{-1}A = I_n$$

In this case, the matrix A is called **invertible**.

Such a matrix A^{-1} will have the same size as the matrix A. It is very important to observe that the inverse of a matrix, if it exists, is unique. Another way to think of this is that if it acts like the inverse, then it **is** the inverse.

Theorem 3.5.1.1: Uniqueness of Inverse

Suppose A is an $n \times n$ matrix such that an inverse A^{-1} exists. Then there is only one such inverse matrix. That is, given any matrix B such that $AB = BA = I$, $B = A^{-1}$.

Proof

In this proof, it is assumed that I is the $n \times n$ identity matrix. Let A, B be $n \times n$ matrices such that A^{-1} exists and $AB = BA = I$. We want to show that $A^{-1} = B$. Now using properties we have seen, we get:

$$A^{-1} = A^{-1}I = A^{-1}(AB) = \left(A^{-1}A\right)B = IB = B$$

Hence, $A^{-1} = B$ which tells us that the inverse is unique.

The next example demonstrates how to check the inverse of a matrix.

Example 3.5.1.1: Verifying the Inverse of a Matrix

Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. Show $\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$ is the inverse of A.

Solution

To check this, multiply

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

and

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

showing that this matrix is indeed the inverse of A.

Unlike ordinary multiplication of numbers, it can happen that $A \neq 0$ but A may fail to have an inverse. This is illustrated in the following example.

Example 3.5.1.2: A Nonzero Matrix With No Inverse

Let $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Show that A does not have an inverse.

Solution

One might think A would have an inverse because it does not equal zero. However, note that

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

If A^{-1} existed, we would have the following

$$\begin{aligned} \begin{bmatrix} 0 \\ 0 \end{bmatrix} &= A^{-1} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix} \right) \\ &= A^{-1} \left(A \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right) \\ &= \left(A^{-1} A \right) \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= I \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ 1 \end{bmatrix} \end{aligned}$$

This says that

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

which is impossible! Therefore, A does not have an inverse.

In the next section, we will explore how to find the inverse of a matrix, if it exists.

This page titled 3.5.1: The Identity and Inverses is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) via source content that was edited to the style and standards of the LibreTexts platform.

- **2.6: The Identity and Inverses** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

3.5.2: Finding the Inverse of a Matrix

In Example 2.6.1, we were given A^{-1} and asked to verify that this matrix was in fact the inverse of A. In this section, we explore how to find A^{-1}.

Let

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

as in Example 2.6.1. In order to find A^{-1}, we need to find a matrix $\begin{bmatrix} x & z \\ y & w \end{bmatrix}$ such that

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x & z \\ y & w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

We can multiply these two matrices, and see that in order for this equation to be true, we must find the solution to the systems of equations,

$$\begin{aligned} x + y &= 1 \\ x + 2y &= 0 \end{aligned}$$

and

$$\begin{aligned} z + w &= 0 \\ z + 2w &= 1 \end{aligned}$$

Writing the augmented matrix for these two systems gives

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

for the first system and

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 2 & 1 \end{array}\right] \tag{3.5.2.1}$$

for the second.

Let's solve the first system. Take -1 times the first row and add to the second to get

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

Now take -1 times the second row and add to the first to get

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

Writing in terms of variables, this says $x = 2$ and $y = -1$.

Now solve the second system, (3.5.2.1) to find z and w. You will find that $z = -1$ and $w = 1$.

If we take the values found for $x, y, z,$ and w and put them into our inverse matrix, we see that the inverse is

$$A^{-1} = \begin{bmatrix} x & z \\ y & w \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

After taking the time to solve the second system, you may have noticed that exactly the same row operations were used to solve both systems. In each case, the end result was something of the form $[I|X]$ where I is the identity and X gave a column of the inverse. In the above,

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

the first column of the inverse was obtained by solving the first system and then the second column

$$\begin{bmatrix} z \\ w \end{bmatrix}$$

To simplify this procedure, we could have solved both systems at once! To do so, we could have written

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

and row reduced until we obtained

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

and read off the inverse as the 2×2 matrix on the right side.

This exploration motivates the following important algorithm.

Algorithm 3.5.2.1: Matrix Inverse Algorithm

Suppose A is an $n \times n$ matrix. To find A^{-1} if it exists, form the augmented $n \times 2n$ matrix

$$[A|I]$$

If possible do row operations until you obtain an $n \times 2n$ matrix of the form

$$[I|B]$$

When this has been done, $B = A^{-1}$. In this case, we say that A is **invertible**. If it is impossible to row reduce to a matrix of the form $[I|B]$, then A has no inverse.

This algorithm shows how to find the inverse if it exists. It will also tell you if A does not have an inverse.

Consider the following example.

Example 3.5.2.1: Finding the Inverse

Let $A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 0 & 2 \\ 3 & 1 & -1 \end{bmatrix}$. Find A^{-1} if it exists.

Solution

Set up the augmented matrix

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 3 & 1 & -1 & 0 & 0 & 1 \end{array}\right]$$

Now we row reduce, with the goal of obtaining the 3×3 identity matrix on the left hand side. First, take -1 times the first row and add to the second followed by -3 times the first row added to the third row. This yields

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

Then take 5 times the second row and add to -2 times the third row.

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 2 & 1 & 0 & 0 \\ 0 & -10 & 0 & -5 & 5 & 0 \\ 0 & 0 & 14 & 1 & 5 & -2 \end{array}\right]$$

Next take the third row and add to -7 times the first row. This yields

$$\left[\begin{array}{ccc|ccc} -7 & -14 & 0 & -6 & 5 & -2 \\ 0 & -10 & 0 & -5 & 5 & 0 \\ 0 & 0 & 14 & 1 & 5 & -2 \end{array}\right]$$

Now take $-\frac{7}{5}$ times the second row and add to the first row.

$$\left[\begin{array}{ccc|ccc} -7 & 0 & 0 & 1 & -2 & -2 \\ 0 & -10 & 0 & -5 & 5 & 0 \\ 0 & 0 & 14 & 1 & 5 & -2 \end{array}\right]$$

Finally divide the first row by -7, the second row by -10 and the third row by 14 which yields

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{7} & \frac{2}{7} & \frac{2}{7} \\ 0 & 1 & 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{1}{14} & \frac{5}{14} & -\frac{1}{7} \end{array}\right]$$

Notice that the left hand side of this matrix is now the 3×3 identity matrix I_3. Therefore, the inverse is the 3×3 matrix on the right hand side, given by

$$\left[\begin{array}{ccc} -\frac{1}{7} & \frac{2}{7} & \frac{2}{7} \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{1}{14} & \frac{5}{14} & -\frac{1}{7} \end{array}\right]$$

It may happen that through this algorithm, you discover that the left hand side cannot be row reduced to the identity matrix. Consider the following example of this situation.

Example 3.5.2.2: A Matrix Which Has No Inverse

Let $A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 0 & 2 \\ 2 & 2 & 4 \end{bmatrix}$. Find A^{-1} if it exists.

Solution

Write the augmented matrix $[A|I]$

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

and proceed to do row operations attempting to obtain $\left[I|A^{-1}\right]$. Take -1 times the first row and add to the second. Then take -2 times the first row and add to the third row.

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

Next add -1 times the second row to the third row.

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

At this point, you can see there will be no way to obtain I on the left side of this augmented matrix. Hence, there is no way to complete this algorithm, and therefore the inverse of A does not exist. In this case, we say that A is not invertible.

If the algorithm provides an inverse for the original matrix, it is always possible to check your answer. To do so, use the method demonstrated in Example 2.6.1. Check that the products AA^{-1} and $A^{-1}A$ both equal the identity matrix. Through this method, you can always be sure that you have calculated A^{-1} properly!

One way in which the inverse of a matrix is useful is to find the solution of a system of linear equations. Recall from Definition 2.2.4 that we can write a system of equations in matrix form, which is of the form $AX = B$. Suppose you find the inverse of the matrix A^{-1}. Then you could multiply both sides of this equation on the left by A^{-1} and simplify to obtain

$$\begin{aligned} \left(A^{-1}\right) AX &= A^{-1}B \\ \left(A^{-1}A\right) X &= A^{-1}B \\ IX &= A^{-1}B \\ X &= A^{-1}B \end{aligned}$$

Therefore we can find X, the solution to the system, by computing $X = A^{-1}B$. Note that once you have found A^{-1}, you can easily get the solution for different right hand sides (different B). It is always just $A^{-1}B$.

We will explore this method of finding the solution to a system in the following example.

Example 3.5.2.3: Using the Inverse to Solve a System of Equations

Consider the following system of equations. Use the inverse of a suitable matrix to give the solutions to this system.

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

Solution

First, we can write the system of equations in matrix form

$$AX = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = B \tag{3.5.2.2}$$

The inverse of the matrix

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

is

$$A^{-1} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & -1 & 0 \\ 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

Verifying this inverse is left as an exercise.

From here, the solution to the given system (3.5.2.2) is found by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}B = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & -1 & 0 \\ 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ -2 \\ -\frac{3}{2} \end{bmatrix}$$

What if the right side, B, of (3.5.2.2) had been $\begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$? In other words, what would be the solution to

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}?$$

By the above discussion, the solution is given by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = A^{-1}B = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & -1 & 0 \\ 1 & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$$

This illustrates that for a system $AX = B$ where A^{-1} exists, it is easy to find the solution when the vector B is changed.

We conclude this section with some important properties of the inverse.

Theorem 3.5.2.1: Inverses of Transposes and Products

Let A, B, and A_i for $i = 1, \ldots, k$ be $n \times n$ matrices.

1. If A is an invertible matrix, then $(A^T)^{-1} = (A^{-1})^T$
2. If A and B are invertible matrices, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$
3. If $A_1, A_2, \ldots, A_k$ are invertible, then the product $A_1 A_2 \cdots A_k$ is invertible, and $(A_1 A_2 \cdots A_k)^{-1} = A_k^{-1} A_{k-1}^{-1} \cdots A_2^{-1} A_1^{-1}$

Consider the following theorem.

Theorem 3.5.2.2: Properties of the Inverse

Let A be an $n \times n$ matrix and I the usual identity matrix.

1. I is invertible and $I^{-1} = I$
2. If A is invertible then so is A^{-1}, and $(A^{-1})^{-1} = A$
3. If A is invertible then so is A^k, and $(A^k)^{-1} = (A^{-1})^k$
4. If A is invertible and p is a nonzero real number, then pA is invertible and $(pA)^{-1} = \frac{1}{p}A^{-1}$

This page titled 3.5.2: Finding the Inverse of a Matrix is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Ken Kuttler (Lyryx) via source content that was edited to the style and standards of the LibreTexts platform.

- **2.7: Finding the Inverse of a Matrix** by Ken Kuttler is licensed CC BY 4.0. Original source: https://lyryx.com/first-course-linear-algebra.

3.5.3: The Matrix Inverse

Learning Objectives

- T/F: If A and B are square matrices where $AB = I$, then $BA = I$.
- T/F: A matrix A has exactly one inverse, infinite inverses, or no inverse.
- T/F: Everyone is special.
- T/F: If A is invertible, then $A\vec{x} = \vec{0}$ has exactly 1 solution.
- What is a corollary?
- Fill in the blanks: _____ a matrix is invertible is useful; computing the inverse is _____.

Once again we visit the old algebra equation, $ax = b$. How do we solve for x? We know that, as long as $a \neq 0$,

$$x = \frac{b}{a}, \text{ or, stated in another way, } x = a^{-1}b.$$

What is a^{-1}? It is the number that, when multiplied by a, returns 1. That is,

$$a^{-1}a = 1.$$

Let us now think in terms of matrices. We have learned of the identity matrix I that "acts like the number 1." That is, if A is a square matrix, then

$$IA = AI = A.$$

If we had a matrix, which we'll call A^{-1}, where $A^{-1}A = I$, then by analogy to our algebra example above it seems like we might be able to solve the linear system $A\vec{x} = \vec{b}$ for $\vec{x}$ by multiplying both sides of the equation by A^{-1}. That is, perhaps

$$\vec{x} = A^{-1}\vec{b}.$$

Of course, there is a lot of speculation here. We don't know that such a matrix like A^{-1} exists. However, we do know how to solve the matrix equation $AX = B$, so we can use that technique to solve the equation $AX = I$ for X. This seems like it will get us close to what we want. Let's practice this once and then study our results.

Example 3.5.3.1

Let

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

Find a matrix X such that $AX = I$.

Solution

We know how to solve this from the previous section: we form the proper augmented matrix, put it into reduced row echelon form and interpret the results.

$$\begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} \xrightarrow[\text{rref}]{} \begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 1 & -1 & 2 \end{bmatrix}$$

We read from our matrix that

$$X = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}.$$

Let's check our work:

$$\begin{aligned} AX &= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= I \end{aligned} \tag{3.5.3.1}$$

Sure enough, it works.

Looking at our previous example, we are tempted to jump in and call the matrix X that we found "A^{-1}." However, there are two obstacles in the way of us doing this.

First, we know that in general $AB \neq BA$. So while we found that $AX = I$, we can't automatically assume that $XA = I$.

Secondly, we have seen examples of matrices where $AB = AC$, but $B \neq C$. So just because $AX = I$, it is possible that another matrix Y exists where $AY = I$. If this is the case, using the notation A^{-1} would be misleading, since it could refer to more than one matrix.

These obstacles that we face are not insurmountable. The first obstacle was that we know that $AX = I$ but didn't know that $XA = I$. That's easy enough to check, though. Let's look at A and X from our previous example.

$$\begin{aligned} AX &= \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= I \end{aligned} \tag{3.5.3.2}$$

Perhaps this first obstacle isn't much of an obstacle after all. Of course, we only have one example where it worked, so this doesn't mean that it always works. We have good news, though: it always does work. The only "bad" news to come with this is that this is a bit harder to prove. We won't worry about proving it always works, but state formally that it does in the following theorem.

Theorem 3.5.3.1

Special Commuting Matrix Products

Let A be an $n \times n$ matrix.

1. If there is a matrix X such that $AX = I_n$, then $XA = I_n$.
2. If there is a matrix X such that $XA = I_n$, then $AX = I_n$.

The second obstacle is easier to address. We want to know if another matrix Y exists where $AY = I = YA$. Let's suppose that it does. Consider the expression XAY. Since matrix multiplication is associative, we can group this any way we choose. We could group this as $(XA)Y$; this results in

$$\begin{aligned} (XA)Y &= IY \\ &= Y. \end{aligned} \tag{3.5.3.3}$$

We could also group XAY as $X(AY)$. This tells us

$$\begin{aligned} X(AY) &= XI \\ &= X \end{aligned} \tag{3.5.3.4}$$

Combining the two ideas above, we see that $X = XAY = Y$; that is, $X = Y$. We conclude that there is only one matrix X where $XA = I = AX$. (Even if we think we have two, we can do the above exercise and see that we really just have one.)

We have just proved the following theorem.

Theorem 3.5.3.2

Uniqueness of Solutions to $AX = I_n$

Let A be an $n \times n$ matrix and let X be a matrix where $AX = I_n$. Then X is unique; it is the only matrix that satisfies this equation.

So given a square matrix A, if we can find a matrix X where $AX = I$, then we know that $XA = I$ and that X is the only matrix that does this. This makes X special, so we give it a special name.

Definition: Invertible Matrices and the Inverse of A

Let A and X be $n \times n$ matrices where $AX = I = XA$. Then:

1. A is *invertible*.
2. X is the *inverse* of A, denoted by A^{-1}.

Let's do an example.

Example 3.5.3.2

Find the inverse of $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$.

Solution

By solving the equation $AX = I$ for X will give us the inverse of A. Forming the appropriate augmented matrix and finding its reduced row echelon form gives us

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 1 \end{bmatrix} \xrightarrow[\text{rref}]{} \begin{bmatrix} 1 & 2 & 0 & 1/2 \\ 0 & 0 & 1 & -1/2 \end{bmatrix}$$

Yikes! We were expecting to find that the reduced row echelon form of this matrix would look like

$$\begin{bmatrix} I & A^{-1} \end{bmatrix}.$$

However, we don't have the identity on the left hand side. Our conclusion: A is not invertible.

We have just seen that not all matrices are invertible.[1] With this thought in mind, let's complete the array of boxes we started before the example. We've discovered that if a matrix has an inverse, it has only one. Therefore, we gave that special matrix a name, "*the* inverse." Finally, we describe the most general way to find the inverse of a matrix, and a way to tell if it does not have one.

Key Idea 3.5.3.1: Finding A^{-1}

Let A be an $n \times n$ matrix. To find A^{-1}, put the augmented matrix

$$\begin{bmatrix} A & I_n \end{bmatrix}$$

into reduced row echelon form. If the result is of the form

$$\begin{bmatrix} I_n & X \end{bmatrix},$$

then $A^{-1} = X$. If not, (that is, if the first n columns of the reduced row echelon form are not I_n), then A is not invertible.

Let's try again.

Example 3.5.3.3

Find the inverse, if it exists, of $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$.

Solution

We'll try to solve $AX = I$ for X and see what happens.

$$\begin{bmatrix} 1 & 1 & -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 3 & 0 & 0 & 1 \end{bmatrix} \underset{\text{rref}}{\rightarrow} \begin{bmatrix} 1 & 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 1 & 0 & 0.2 & -0.4 & 0.2 \\ 0 & 0 & 1 & -0.3 & 0.1 & 0.2 \end{bmatrix}$$

We have a solution, so

$$A = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.2 & -0.4 & 0.2 \\ -0.3 & 0.1 & 0.2 \end{bmatrix}.$$

Multiply AA^{-1} to verify that it is indeed the inverse of A.

In general, given a matrix A, to find A^{-1} we need to form the augmented matrix $[\,A \quad I\,]$ and put it into reduced row echelon form and interpret the result. In the case of a 2×2 matrix, though, there is a shortcut. We give the shortcut in terms of a theorem.[2]

Theorem 3.5.3.3

The Inverse of a 2×2 Matrix

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

A is invertible if and only if $ad - bc \neq 0$.

If $ad - bc \neq 0$, then

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

We can't divide by 0, so if $ad - bc = 0$, we don't have an inverse. Recall Example 3.5.3.2 where

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

Here, $ad - bc = 1(4) - 2(2) = 0$, which is why A didn't have an inverse.

Although this idea is simple, we should practice it.

Example 3.5.3.4

Use Theorem 3.5.3.3to find the inverse of

$$A = \begin{bmatrix} 3 & 2 \\ -1 & 9 \end{bmatrix}$$

if it exists.

Solution

Since $ad - bc = 29 \neq 0$, A^{-1} exists. By the Theorem,

$$\begin{aligned} A^{-1} &= \frac{1}{3(9)-2(-1)}\begin{bmatrix} 9 & -2 \\ 1 & 3 \end{bmatrix} \\ &= \frac{1}{29}\begin{bmatrix} 9 & -2 \\ 1 & 3 \end{bmatrix} \end{aligned} \tag{3.5.3.5}$$

We can leave our answer in this form, or we could "simplify" it as

$$A^{-1} = \frac{1}{29}\begin{bmatrix} 9 & -2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 9/29 & -2/29 \\ 1/29 & 3/29 \end{bmatrix}.$$

We started this section out by speculating that just as we solved algebraic equations of the form $ax = b$ by computing $x = a^{-1}b$, we might be able to solve matrix equations of the form $A\vec{x} = \vec{b}$ by computing $\vec{x} = A^{-1}\vec{b}$. If A^{-1} does exist, then we *can* solve the equation $A\vec{x} = \vec{b}$ this way. Consider:

$$\begin{aligned} A\vec{x} &= \vec{b} && \text{(original equation)} \\ A^{-1}A\vec{x} &= A^{-1}\vec{b} && \text{(multiply both sides on the left by } A^{-1}) \\ I\vec{x} &= A^{-1}\vec{b} && \text{(since } A^{-1}A = I) \\ \vec{x} &= A^{-1}\vec{b} && \text{(since } I\vec{x} = \vec{x}) \end{aligned} \tag{3.5.3.6}$$

Let's step back and think about this for a moment. The only thing we know about the equation $A\vec{x} = \vec{b}$ is that A is invertible. We also know that solutions to $A\vec{x} = \vec{b}$ come in three forms: exactly one solution, infinite solutions, and no solution. We just showed that if A is invertible, then $A\vec{x} = \vec{b}$ has *at least* one solution. We showed that by setting $\vec{x}$ equal to $A^{-1}\vec{b}$, we have a solution. Is it possible that more solutions exist?

No. Suppose we are told that a known vector $\vec{v}$ is a solution to the equation $A\vec{x} = \vec{b}$; that is, we know that $A\vec{v} = \vec{b}$. We can repeat the above steps:

$$\begin{aligned} A\vec{v} &= \vec{b} \\ A^{-1}A\vec{v} &= A^{-1}\vec{b} \\ I\vec{v} &= A^{-1}\vec{b} \\ \vec{v} &= A^{-1}\vec{b}. \end{aligned} \tag{3.5.3.7}$$

This shows that *all* solutions to $A\vec{x} = \vec{b}$ are exactly $\vec{x} = A^{-1}\vec{b}$ when A is invertible. We have just proved the following theorem.

Theorem 3.5.3.4

Invertible Matrices and Solutions to $A\vec{x} = \vec{b}$

Let A be an invertible $n \times n$ matrix, and let $\vec{b}$ be any $n \times 1$ column vector. Then the equation $A\vec{x} = \vec{b}$ has exactly one solution, namely

$$\vec{x} = A^{-1}\vec{b}.$$

A corollary[3] to this theorem is: If A is not invertible, then $A\vec{x} = \vec{b}$ does not have exactly one solution. It may have infinite solutions and it may have no solution, and we would need to examine the reduced row echelon form of the augmented matrix $\begin{bmatrix} A & \vec{b} \end{bmatrix}$ to see which case applies.

We demonstrate our theorem with an example.

Example 3.5.3.5

Solve $A\vec{x} = \vec{b}$ by computing $\vec{x} = A^{-1}\vec{b}$, where

$$A = \begin{bmatrix} 1 & 0 & -3 \\ -3 & -4 & 10 \\ 4 & -5 & -11 \end{bmatrix} \quad \text{and} \quad \vec{b} = \begin{bmatrix} -15 \\ 57 \\ -46 \end{bmatrix}.$$

Solution

Without showing our steps, we compute

$$A^{-1} = \begin{bmatrix} 94 & 15 & -12 \\ 7 & 1 & -1 \\ 31 & 5 & -4 \end{bmatrix}.$$

We then find the solution to $A\vec{x} = \vec{b}$ by computing $A^{-1}\vec{b}$:

$$\begin{aligned} \vec{x} &= A^{-1}\vec{b} \\ &= \begin{bmatrix} 94 & 15 & -12 \\ 7 & 1 & -1 \\ 31 & 5 & -4 \end{bmatrix} \begin{bmatrix} -15 \\ 57 \\ -46 \end{bmatrix} \\ &= \begin{bmatrix} -3 \\ -2 \\ 4 \end{bmatrix}. \end{aligned} \tag{3.5.3.8}$$

We can easily check our answer:

$$\begin{bmatrix} 1 & 0 & -3 \\ -3 & -4 & 10 \\ 4 & -5 & -11 \end{bmatrix} \begin{bmatrix} -3 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} -15 \\ 57 \\ -46 \end{bmatrix}.$$

Knowing a matrix is invertible is incredibly useful.[4] Among many other reasons, if you know A is invertible, then you know for sure that $A\vec{x} = \vec{b}$ has a solution (as we just stated in Theorem 3.5.3.4). In the next section we'll demonstrate many different properties of invertible matrices, including stating several different ways in which we know that a matrix is invertible.

Footnotes

[1] Hence our previous definition; why bother calling A "invertible" if every square matrix is? If everyone is special, then no one is. Then again, everyone *is* special.

[2] We don't prove this theorem here, but it really isn't hard to do. Put the matrix

$$\begin{bmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{bmatrix}$$

into reduced row echelon form and you'll discover the result of the theorem. Alternatively, multiply A by what we propose is the inverse and see that we indeed get I.

[3] a *corollary* is an idea that follows directly from a theorem

[4] As odd as it may sound, *knowing* a matrix is invertible is useful; actually computing the inverse isn't. This is discussed at the end of the next section.

This page titled 3.5.3: The Matrix Inverse is shared under a CC BY-NC 3.0 license and was authored, remixed, and/or curated by Gregory Hartman et al. via source content that was edited to the style and standards of the LibreTexts platform.

- **2.6: The Matrix Inverse** by Gregory Hartman et al. is licensed CC BY-NC 3.0. Original source: https://github.com/APEXCalculus/Fundamentals-of-Matrix-Algebra.

3.5.3.1: Exercises 2.6

In Exercises 3.5.3.1.1- 3.5.3.1.8 a matrix A is given. Find A^{-1} using Theorem 2.6.3, if it exists.

Exercise 3.5.3.1.1

$$\begin{bmatrix} 1 & 5 \\ -5 & -24 \end{bmatrix}$$

Answer

$$\begin{bmatrix} -24 & -5 \\ 5 & 1 \end{bmatrix}$$

Exercise 3.5.3.1.2

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

Answer

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

Exercise 3.5.3.1.3

$$\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 1/3 & 0 \\ 0 & 1/7 \end{bmatrix}$$

Exercise 3.5.3.1.4

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

Answer

$$\begin{bmatrix} -4/7 & 5/7 \\ 3/7 & -2/7 \end{bmatrix}$$

Exercise 3.5.3.1.5

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

Answer

A^{-1} does not exist.

Exercise 3.5.3.1.6

$$\begin{bmatrix} 3 & 7 \\ 2 & 4 \end{bmatrix}$$

Answer

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

Exercise 3.5.3.1.7

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

Answer

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

Exercise 3.5.3.1.8

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

Answer

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

In Exercises 3.5.3.1.9- 3.5.3.1.28, a matrix A is given. Find A^{-1} using Key Idea 2.6.1, if it exists.

Exercise 3.5.3.1.9

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

Answer

$$\begin{bmatrix} -5/13 & 3/13 \\ 1/13 & 2/13 \end{bmatrix}$$

Exercise 3.5.3.1.10

$$\begin{bmatrix} -5 & -2 \\ 9 & 2 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 1/4 & 1/4 \\ -9/8 & -5/8 \end{bmatrix}$$

Exercise 3.5.3.1.11

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

Answer

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

Exercise 3.5.3.1.12

$$\begin{bmatrix} 5 & 7 \\ 5/3 & 7/3 \end{bmatrix}$$

Answer

A^{-1} does not exist.

Exercise 3.5.3.1.13

$$\begin{bmatrix} 25 & -10 & -4 \\ -18 & 7 & 3 \\ -6 & 2 & 1 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 1 & 2 & -2 \\ 0 & 1 & -3 \\ 6 & 10 & -5 \end{bmatrix}$$

Exercise 3.5.3.1.14

$$\begin{bmatrix} 2 & 3 & 4 \\ -3 & 6 & 9 \\ -1 & 9 & 13 \end{bmatrix}$$

Answer

A^{-1} does not exist.

Exercise 3.5.3.1.15

$$\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & -7 \\ 20 & 7 & -48 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 1 & 0 & 0 \\ 52 & -48 & 7 \\ 8 & -7 & 1 \end{bmatrix}$$

Exercise 3.5.3.1.16

$$\begin{bmatrix} -4 & 1 & 5 \\ -5 & 1 & 9 \\ -10 & 2 & 19 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 1 & -9 & 4 \\ 5 & -26 & 11 \\ 0 & -2 & 1 \end{bmatrix}$$

Exercise 3.5.3.1.17

$$\begin{bmatrix} 5 & -1 & 0 \\ 7 & 7 & 1 \\ -2 & -8 & -1 \end{bmatrix}$$

Answer

A^{-1} does not exist.

Exercise 3.5.3.1.18

$$\begin{bmatrix} 1 & -5 & 0 \\ -2 & 15 & 4 \\ 4 & -19 & 1 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 91 & 5 & -20 \\ 18 & 1 & -4 \\ -22 & -1 & 5 \end{bmatrix}$$

Exercise 3.5.3.1.19

$$\begin{bmatrix} 25 & -8 & 0 \\ -78 & 25 & 0 \\ 48 & -15 & 1 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 25 & 8 & 0 \\ 78 & 25 & 0 \\ -30 & -9 & 1 \end{bmatrix}$$

Exercise 3.5.3.1.20

$$\begin{bmatrix} 1 & 0 & 0 \\ 7 & 5 & 8 \\ -2 & -2 & -3 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 1 & 0 & 0 \\ 5 & -3 & -8 \\ -4 & 2 & 5 \end{bmatrix}$$

Exercise 3.5.3.1.21

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

Answer

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

? Exercise 3.5.3.1.22

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

Answer

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

? Exercise 3.5.3.1.23

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -19 & -9 & 0 & 4 \\ 33 & 4 & 1 & -7 \\ 4 & 2 & 0 & -1 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & -1 & 0 & -4 \\ -35 & -10 & 1 & -47 \\ -2 & -2 & 0 & -9 \end{bmatrix}$$

? Exercise 3.5.3.1.24

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 27 & 1 & 0 & 4 \\ 18 & 0 & 1 & 4 \\ 4 & 0 & 0 & 1 \end{bmatrix}$$

Answer

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

Exercise 3.5.3.1.25

$$\begin{bmatrix} -15 & 45 & -3 & 4 \\ 55 & -164 & 15 & -15 \\ -215 & 640 & -62 & 59 \\ -4 & 12 & 0 & 1 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 28 & 18 & 3 & -19 \\ 5 & 1 & 0 & -5 \\ 4 & 5 & 1 & 0 \\ 52 & 60 & 12 & -15 \end{bmatrix}$$

Exercise 3.5.3.1.26

$$\begin{bmatrix} 1 & 0 & 2 & 8 \\ 0 & 1 & 0 & 0 \\ 0 & -4 & -29 & -110 \\ 0 & -3 & -5 & -19 \end{bmatrix}$$

Answer

$$\begin{bmatrix} 1 & 28 & -2 & 12 \\ 0 & 1 & 0 & 0 \\ 0 & 254 & -19 & 110 \\ 0 & -67 & 5 & -29 \end{bmatrix}$$

Exercise 3.5.3.1.27

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

Answer

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

Exercise 3.5.3.1.28

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

Answer

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

In Exercises 3.5.3.1.29- 3.5.3.1.36, a matrix A and a vector $\vec{b}$ are given. Solve the equation $A\vec{x} = \vec{b}$ using Theorem 2.6.4.

? Exercise 3.5.3.1.29

$$A = \begin{bmatrix} 3 & 5 \\ 2 & 3 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 21 \\ 13 \end{bmatrix}$$

Answer

$$\vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

? Exercise 3.5.3.1.30

$$A = \begin{bmatrix} 1 & -4 \\ 4 & -15 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 21 \\ 77 \end{bmatrix}$$

Answer

$$\vec{x} = \begin{bmatrix} -7 \\ -7 \end{bmatrix}$$

? Exercise 3.5.3.1.31

$$A = \begin{bmatrix} 9 & 70 \\ -4 & -31 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Answer

$$\vec{x} = \begin{bmatrix} -8 \\ 1 \end{bmatrix}$$

? Exercise 3.5.3.1.32

$$A = \begin{bmatrix} 10 & -57 \\ 3 & -17 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -14 \\ -4 \end{bmatrix}$$

Answer

$$\vec{x} = \begin{bmatrix} 10 \\ 2 \end{bmatrix}$$

? Exercise 3.5.3.1.33

$$A = \begin{bmatrix} 1 & 2 & 12 \\ 0 & 1 & 6 \\ -3 & 0 & 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -17 \\ -5 \\ 20 \end{bmatrix}$$

Answer

$$\vec{x} = \begin{bmatrix} -7 \\ 1 \\ -1 \end{bmatrix}$$

? Exercise 3.5.3.1.34

$$A = \begin{bmatrix} 1 & 0 & -3 \\ 8 & -2 & -13 \\ 12 & -3 & -20 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -34 \\ -159 \\ -243 \end{bmatrix}$$

Answer

$$\vec{x} = \begin{bmatrix} -7 \\ -7 \\ 9 \end{bmatrix}$$

? Exercise 3.5.3.1.35

$$A = \begin{bmatrix} 5 & 0 & -2 \\ -8 & 1 & 5 \\ -2 & 0 & 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 33 \\ -70 \\ -15 \end{bmatrix}$$

Answer

$$\vec{x} = \begin{bmatrix} 3 \\ -1 \\ -9 \end{bmatrix}$$

? Exercise 3.5.3.1.36

$$A = \begin{bmatrix} 1 & -6 & 0 \\ 0 & 1 & 0 \\ 2 & -8 & 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} -69 \\ 10 \\ -102 \end{bmatrix}$$

Answer

$$\vec{x} = \begin{bmatrix} -9 \\ 10 \\ -4 \end{bmatrix}$$

3.5.3.1: Exercises 2.6 is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

- **2.6.1: Exercises 2.6** by Gregory Hartman et al. is licensed CC BY-NC 3.0.

3.6: Linear Inequalities

3.6: Linear Inequalities is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

3.6.1: SOLVING SYSTEMS OF LINEAR INEQUALITIES (TWO VARIABLES)

LEARNING OBJECTIVES

- Check solutions to systems of linear inequalities with two variables.
- Solve systems of linear inequalities.

SOLUTIONS TO SYSTEMS OF LINEAR INEQUALITIES

A system of linear inequalities consists of a set of two or more linear inequalities with the same variables. The inequalities define the conditions that are to be considered simultaneously. For example,

We know that each inequality in the set contains infinitely many ordered pair solutions defined by a region in a rectangular coordinate plane. When considering two of these inequalities together, the intersection of these sets defines the set of simultaneous ordered pair solutions. When we graph each of the above inequalities separately, we have

$y > x - 2 \;\; y \leq 2x + 2$

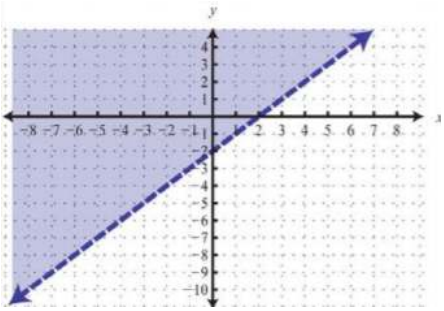


Figure 3.6.1.1

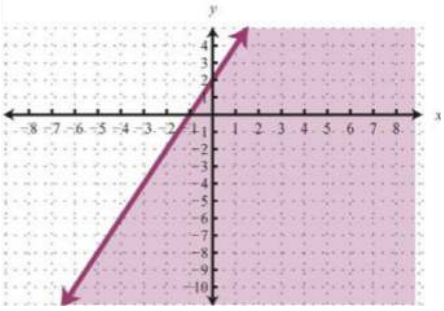


Figure 3.6.1.2

Table 3.6.1.1

When graphed on the same set of axes, the intersection can be determined.

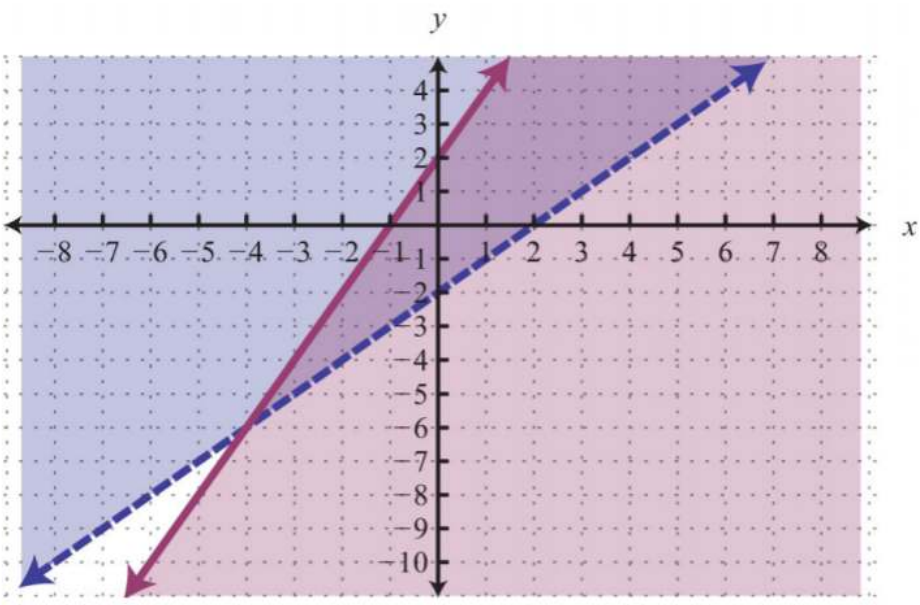


Figure 3.6.1.3

The intersection is shaded darker and the final graph of the solution set is presented as follows:

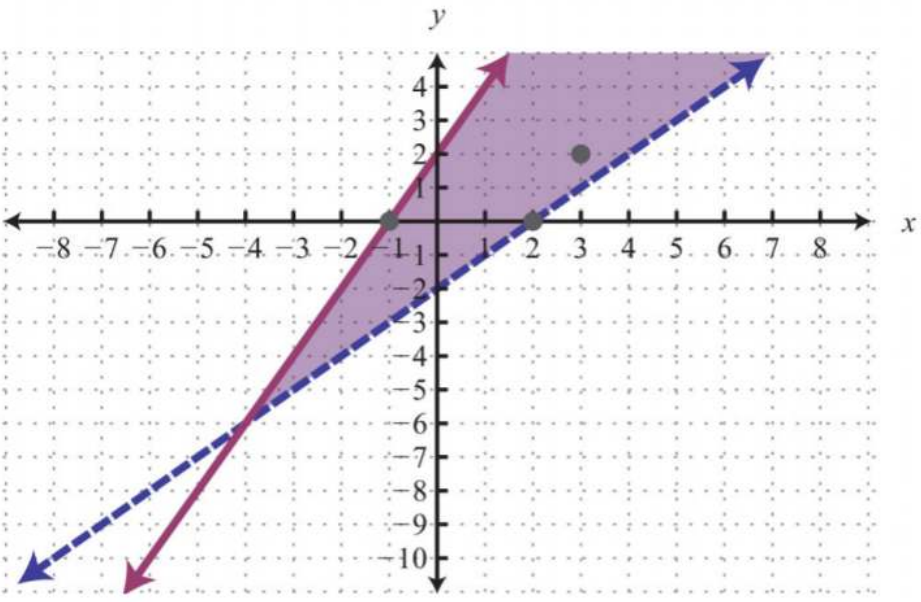


Figure 3.6.1.4

The graph suggests that $(3,2)$ is a solution because it is in the intersection. To verify this, show that it solves both of the original inequalities:

$$Check: (3,2)$$

$Inequality\,1:\quad y > x-2$	$Inequality\,2:\quad y/leq2x+2$
$2 > 3-2$	$2 \leq 2(3)+2$
$2 > 1 \quad \checkmark$	$2 \leq 8 \quad \checkmark$

Points on the solid boundary are included in the set of simultaneous solutions and points on the dashed boundary are not. Consider the point $(-1,0)$ on the solid boundary defined by $y = 2x+2$ and verify that it solves the original system:

$$Check: (-1,0)$$

$$\begin{array}{c|c} \textit{Inequality}\,1: \quad y > x-2 & \textit{Inequality}\,2: \quad y \leq 2x+2 \\ 0 > -1-2 & 0 \leq 2(-1)+2 \\ 0 > -3 \quad \checkmark & 0 \leq 0 \quad \checkmark \end{array}$$

Notice that this point satisfies both inequalities and thus is included in the solution set. Now consider the point $(2,0)$ on the dashed boundary defined by $y = x - 2$ and verify that it does not solve the original system:

$$\textit{Check}: \ (2,0)$$

$$\begin{array}{c|c} \textit{Inequality}\,1: \quad y > x-2 & \textit{Inequality}\,2: \quad y \leq 2x+2 \\ 0 > 2-2 & 0 \leq 2(2)+2 \\ 0 > 0 \quad x & 0 \leq \quad \checkmark \end{array}$$

This point does not satisfy both inequalities and thus is not included in the solution set.

SOLVING SYSTEMS OF LINEAR INEQUALITIES

Solutions to a system of linear inequalities are the ordered pairs that solve all the inequalities in the system. Therefore, to solve these systems, graph the solution sets of the inequalities on the same set of axes and determine where they intersect. This intersection, or overlap, defines the region of common ordered pair solutions.

EXAMPLE 3.6.1.1

Graph the solution set:

$$\begin{cases} -2x + y > -4 \\ 3x - 6y \geq 6 \end{cases}.$$

Solution:

To facilitate the graphing process, we first solve for y.

2x-4 \\ y&\leq \frac{1}{2}x-1 \end{aligned}\right.\)

For the first inequality, we use a dashed boundary defined by $y = 2x - 4$ and shade all points above the line. For the second inequality, we use a solid boundary defined by $y = \frac{1}{2}x - 1$ and shade all points below. The intersection is darkened.

Figure 3.6.1.5

Now we present the solution with only the intersection shaded.

Answer:

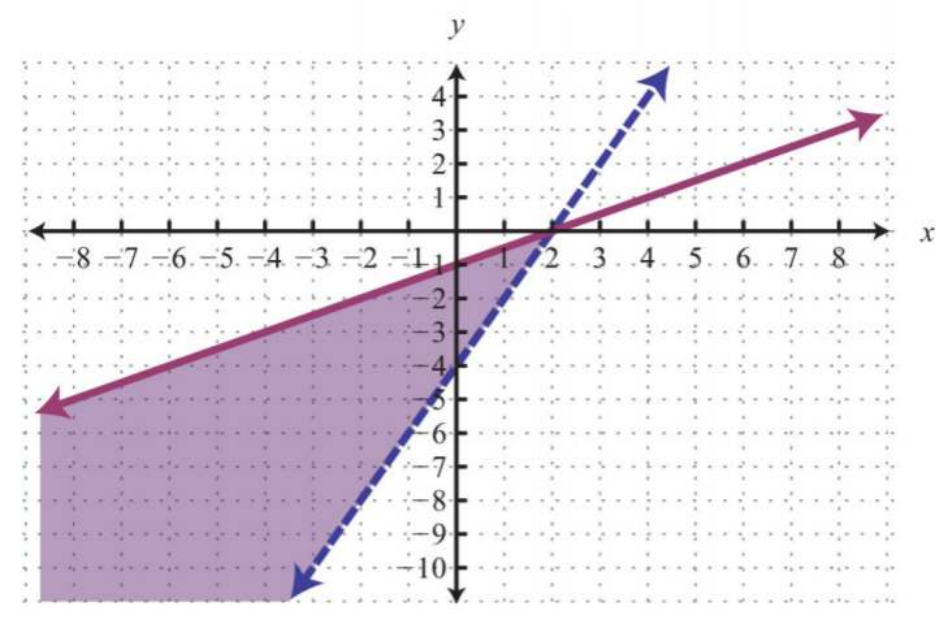


Figure 3.6.1.6

EXAMPLE 3.6.1.2

Graph the solution set:

$$\begin{cases} -2x+3y>6 \\ 4x-6y>12 \end{cases}.$$

Solution:

Begin by solving both inequalities for y.

Use a dashed line for each boundary. For the first inequality, shade all points above the boundary. For the second inequality, shade all points below the boundary.

Figure 3.6.1.7

As you can see, there is no intersection of these two shaded regions. Therefore, there are no simultaneous solutions.

Answer:

No solution, $\emptyset$

EXAMPLE 3.6.1.3

Graph the solution set:

Solution:

Figure 3.6.1.8

After graphing all three inequalities on the same set of axes, we determine that the intersection lies in the triangular region pictured.

Answer:

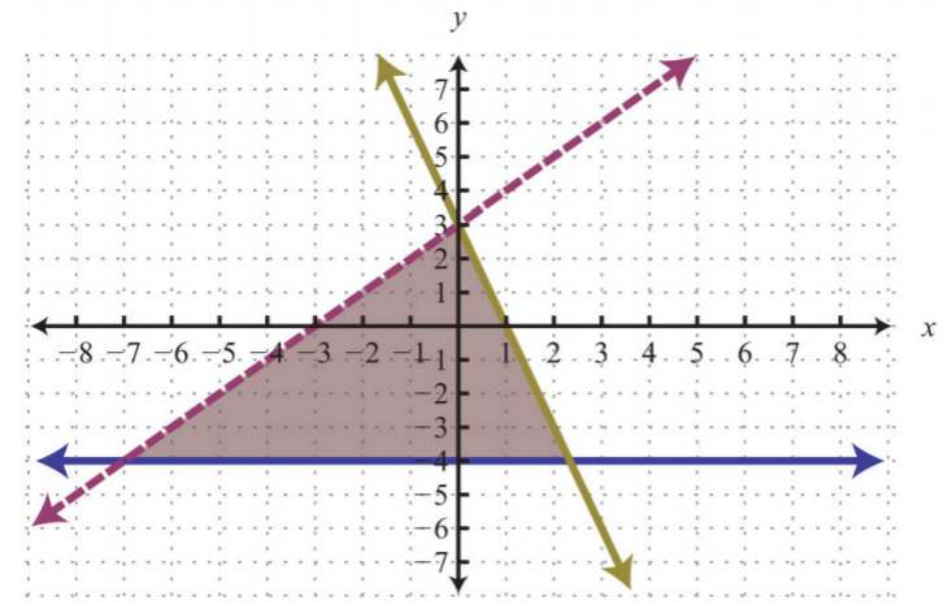


Figure 3.6.1.9

The graphic suggests that $(-1,1)$ is a common point. As a check, substitute that point into the inequalities and verify that it solves all three conditions.

$$Check: \ (-1,1)$$

$Inequality\ 1:$	$Inequality\ 2:$	$Inequality\ 3:$
$y \geq -4$	$y < x+3$	$y/leq -3x+3$
$1 \geq -4 \quad \checkmark$	$1 < -1+3$	$1 \leq -3(-1)+3$
	$1 < 2 \quad \checkmark$	$1 \leq 3+3$
		$1 \leq 6 \quad \checkmark$

KEY TAKEAWAYS

- To solve systems of linear inequalities, graph the solution sets of each inequality on the same set of axes and determine where they intersect.

EXERCISE 3.6.1.1 SOLVING SYSTEMS OF LINEAR INEQUALITIES

Determine whether the given point is a solution to the given system of linear equations.

1. $(3,2)$; $\begin{cases} y \le x+3 \\ y \ge -x+3 \end{cases}$
2. $(-3,-2)$; $\begin{cases} y < -3x+4 \\ y \ge 2x-1 \end{cases}$
3. $(5,0)$; $\begin{cases} y > -x+5 \\ y \le \frac{3}{4}x-2 \end{cases}$
4. $(0,1)$; $\begin{cases} y < \frac{2}{3}x+1 \\ y \ge \frac{5}{2}x-2 \end{cases}$
5. $(-1,\frac{8}{3})$; $\begin{cases} -4x+3y \ge -12 \\ 2x+3y < 6 \end{cases}$
6. $(-1,-2)$; $\begin{cases} -x+y < 0 \\ x+y < 0 \\ x+y < -2 \end{cases}$

Answer

1. Yes
3. No
5. No

EXERCISE 3.6.1.2 SOLVING SYSTEMS OF LINEAR INEQUALITIES

Graph the solution set.

1. $\begin{cases} y \le x+3 \\ y \ge -x+3 \end{cases}$
2. $\begin{cases} y < -3x+4 \\ y \ge 2x-1 \end{cases}$
3. $\begin{cases} y > x \\ y < -1 \end{cases}$
4. $\begin{cases} y < \frac{2}{3}x+1 \\ y \ge \frac{5}{2}x-2 \end{cases}$
5. $\begin{cases} y > -x+5 \\ y \le \frac{3}{4}x-2 \end{cases}$
6. $\begin{cases} y > \frac{3}{5}x+3 \\ y < \frac{3}{5}x-3 \end{cases}$
7. $\begin{cases} x+4y < 12 \\ -3x+12y \ge -12 \end{cases}$
8. $\begin{cases} -x+y \le 6 \\ 2x+y \ge 1 \end{cases}$
9. $\begin{cases} -2x+3y > 3 \\ 4x-3y < 15 \end{cases}$
10. $\begin{cases} -4x+3y \ge -12 \\ 2x+3y < 6 \end{cases}$

11. $\begin{cases} 5x+y \leq 4 \\ -4x+3y < -6 \end{cases}$
12. $\begin{cases} 3x+5y < 15 \\ -x+2y \leq 0 \end{cases}$
13. $\begin{cases} x \geq 0 \\ 5x+y > 5 \end{cases}$
14. $\begin{cases} x \geq -2 \\ y \geq 1 \end{cases}$
15. $\begin{cases} x-3 < 0 \\ y+2 \geq 0 \end{cases}$
16. $\begin{cases} 5y \geq 2x+5 \\ -2x < -5y-5 \end{cases}$
17. $\begin{cases} x-y \geq 0 \\ -x+y < 1 \end{cases}$
18. $\begin{cases} -x+y \geq 0 \\ y-x < 1 \end{cases}$
19. $\begin{cases} x > -2 \\ x \leq 2 \end{cases}$
20. $\begin{cases} y > -1 \\ y < 2 \end{cases}$
21. $\begin{cases} -x+2y > 8 \\ 3x-6y \geq 18 \end{cases}$
22. $\begin{cases} -3x+4y \leq 4 \\ 6x-8y > -8 \end{cases}$
23. $\begin{cases} 2x+y < 3 \\ -x \leq 12y \end{cases}$
24. $\begin{cases} 2x+6y \leq 6 \\ -13x-y \leq 3 \end{cases}$
25. $\begin{cases} y < 3 \\ y > x \\ x > -4 \end{cases}$
26. $\begin{cases} y < 1 \\ y \geq x-1 \\ y < -3x+3 \end{cases}$
27. $\begin{cases} -4x+3y > -12 \\ y \geq 2 \\ 2x+3y > 6 \end{cases}$
28. $\begin{cases} -x+y < 0 \\ x+y \leq 0 \\ x+y > -2 \end{cases}$
29. $\begin{cases} x+y < 2 \\ x < 3 \\ -x+y \leq 2 \end{cases}$
30. $\begin{cases} y+4 \geq 0 \\ \frac{1}{2}x+\frac{1}{3}y \leq 1 \\ -\frac{1}{2}x+\frac{1}{3}y \leq 1 \end{cases}$
31. Construct a system of linear inequalities that describes all points in the first quadrant.
32. Construct a system of linear inequalities that describes all points in the second quadrant.
33. Construct a system of linear inequalities that describes all points in the third quadrant.
34. Construct a system of linear inequalities that describes all points in the fourth quadrant.

Answer

1.

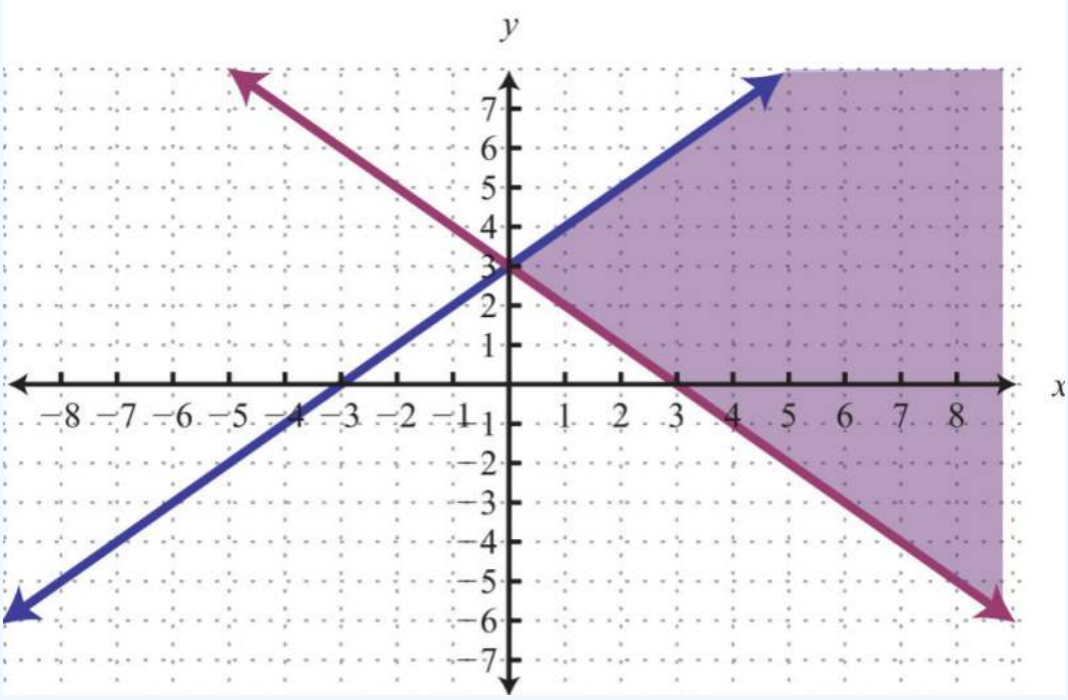


Figure 3.6.1.10

3.

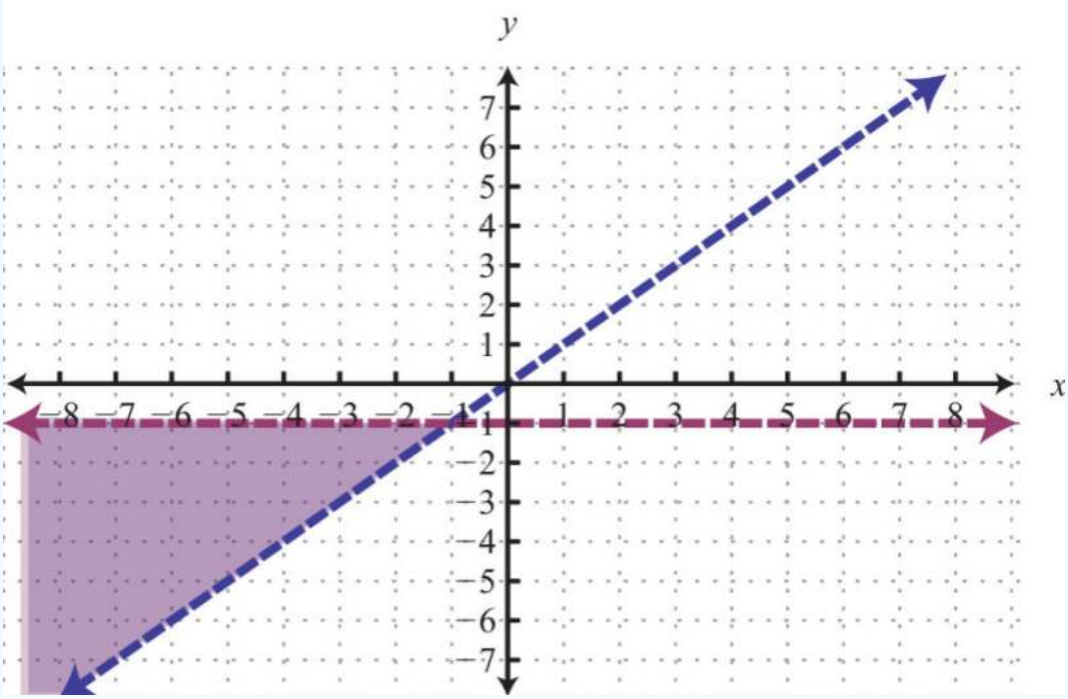


Figure 3.6.1.11

5.

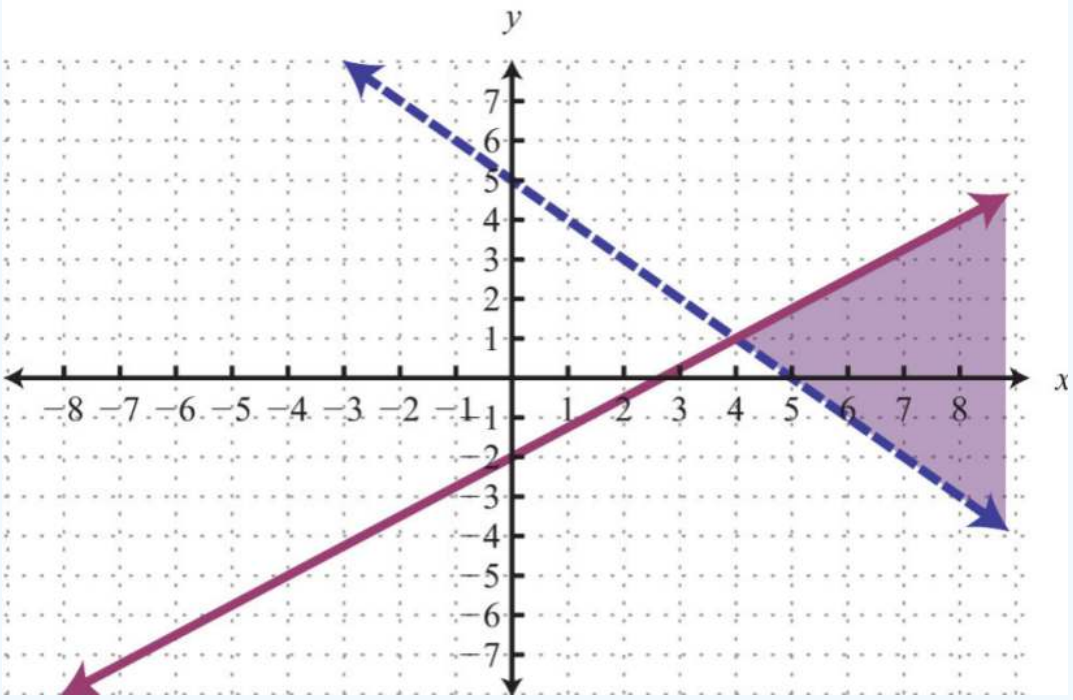


Figure 3.6.1.12

7.

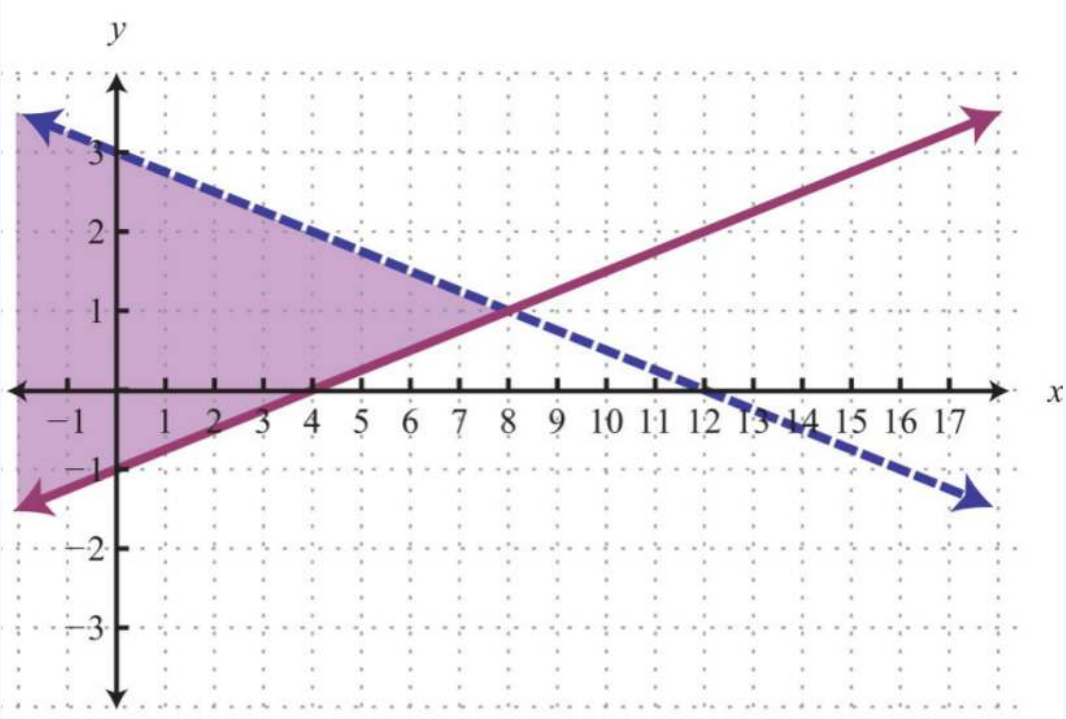


Figure 3.6.1.13

9.

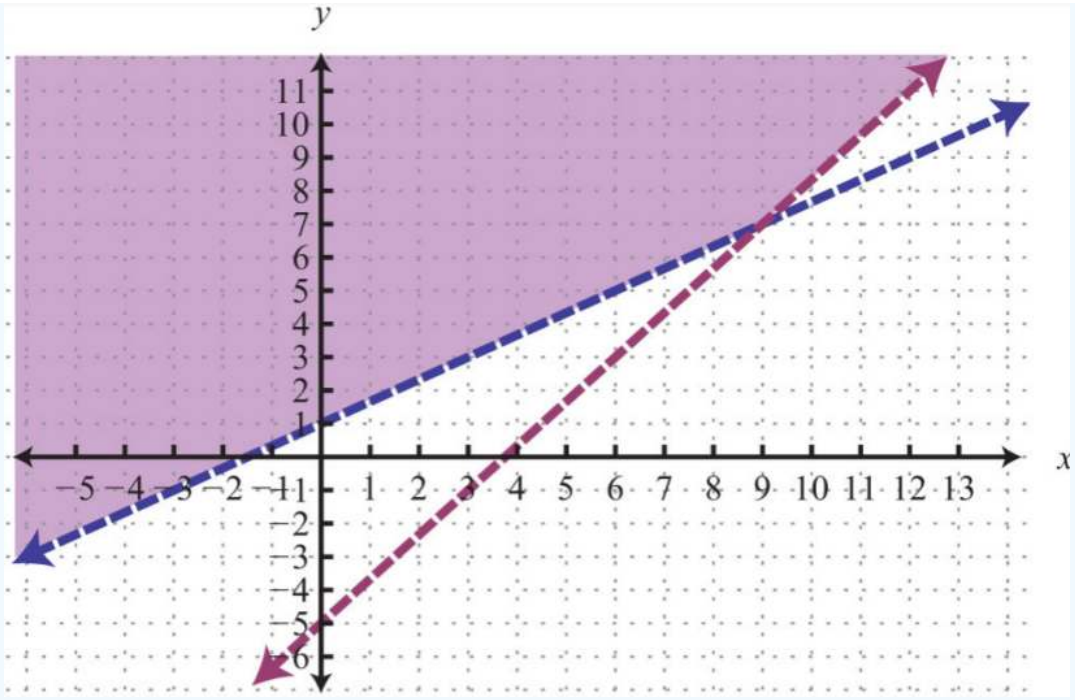


Figure 3.6.1.14

11.

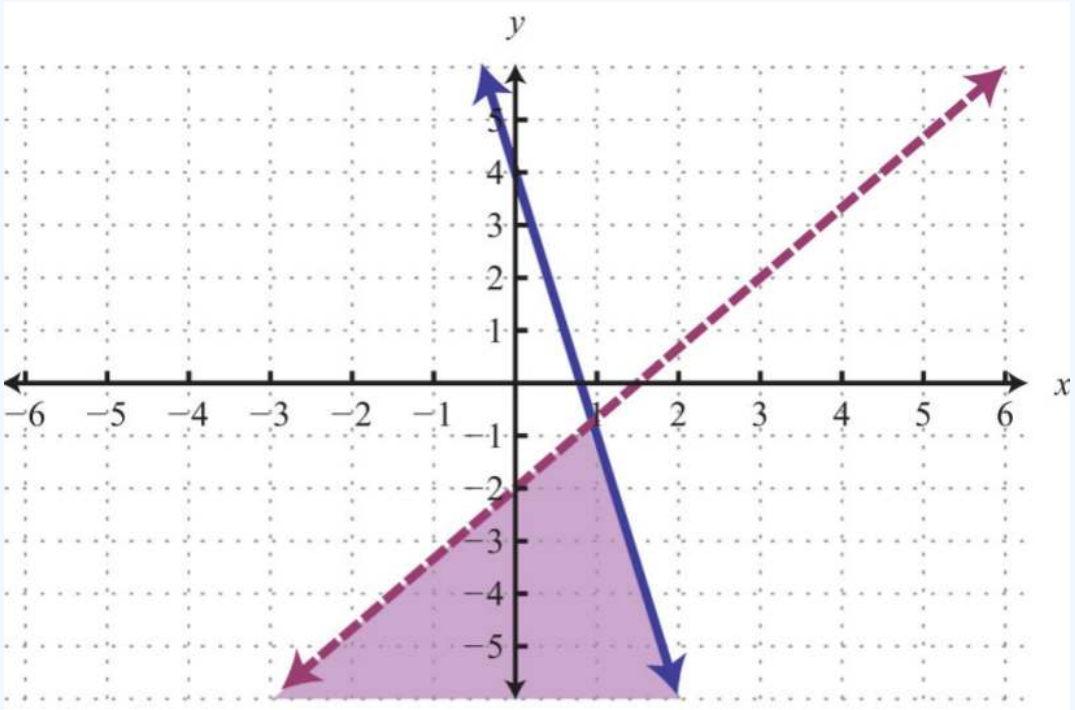


Figure 3.6.1.15

13.

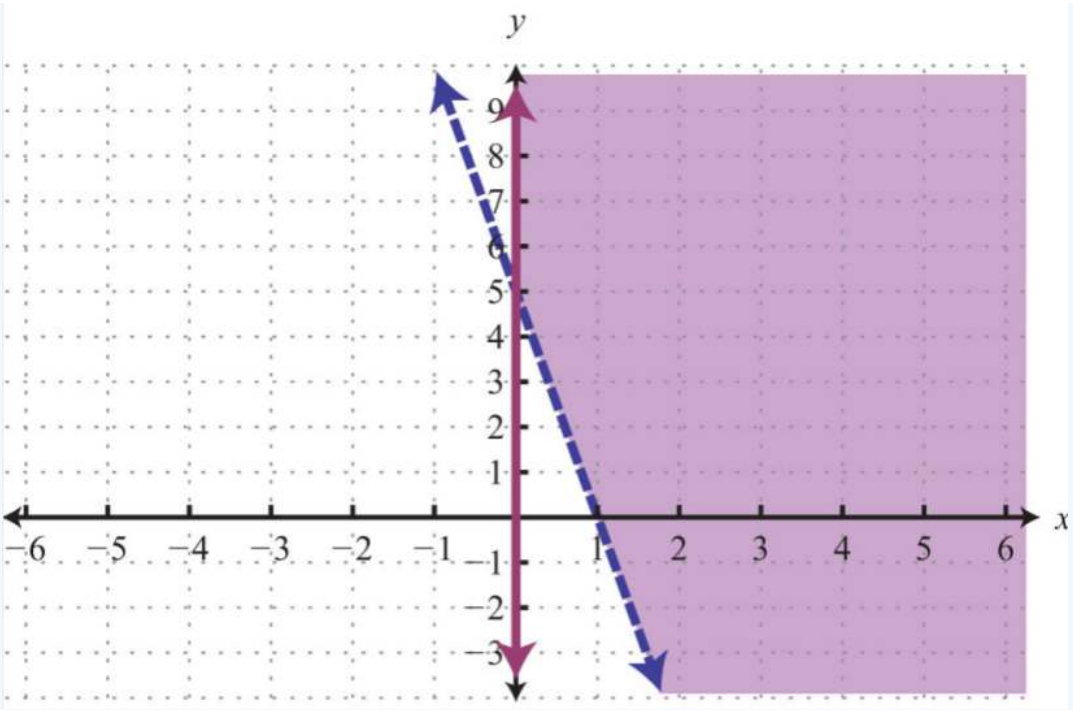


Figure 3.6.1.16

15.

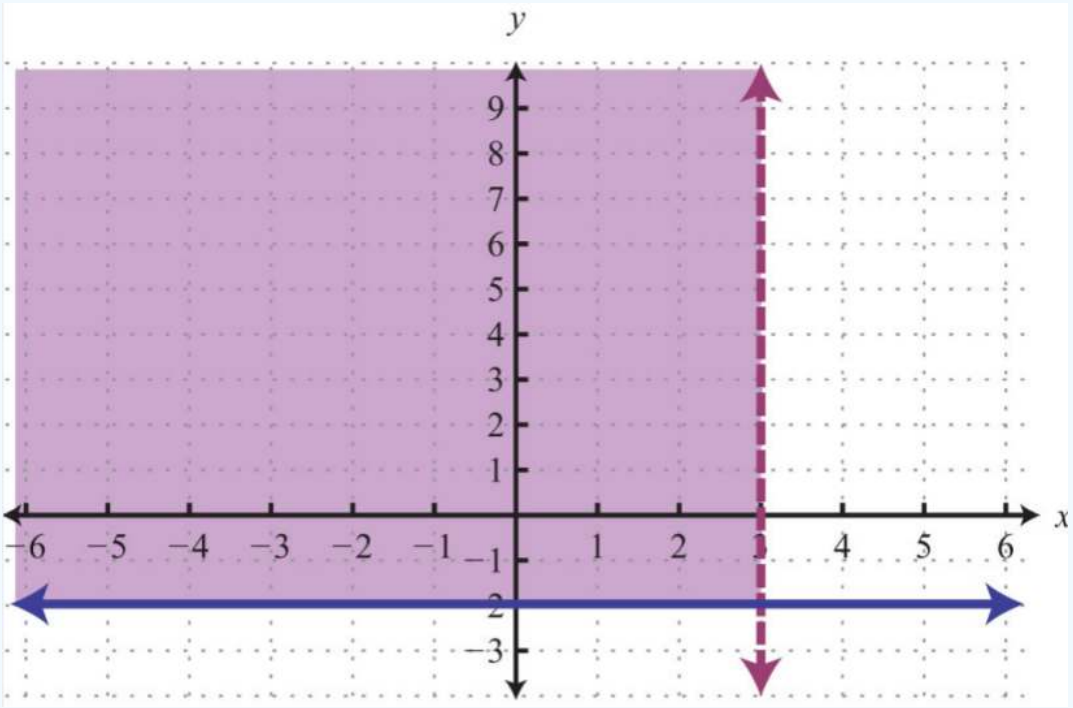


Figure 3.6.1.17

17.

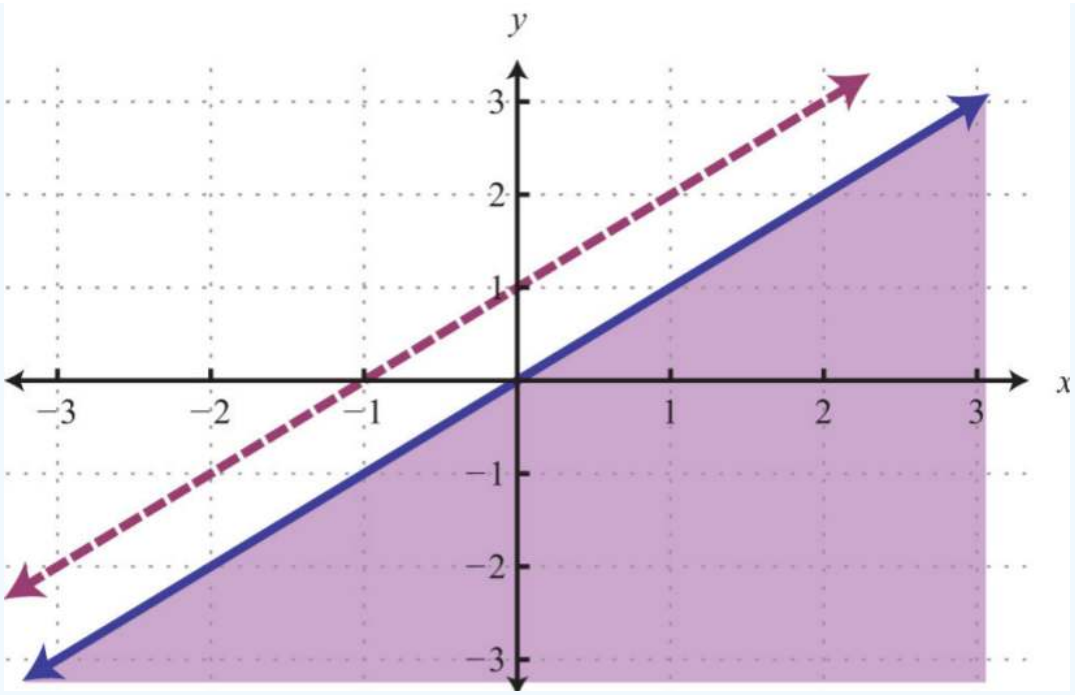


Figure 3.6.1.18

19.

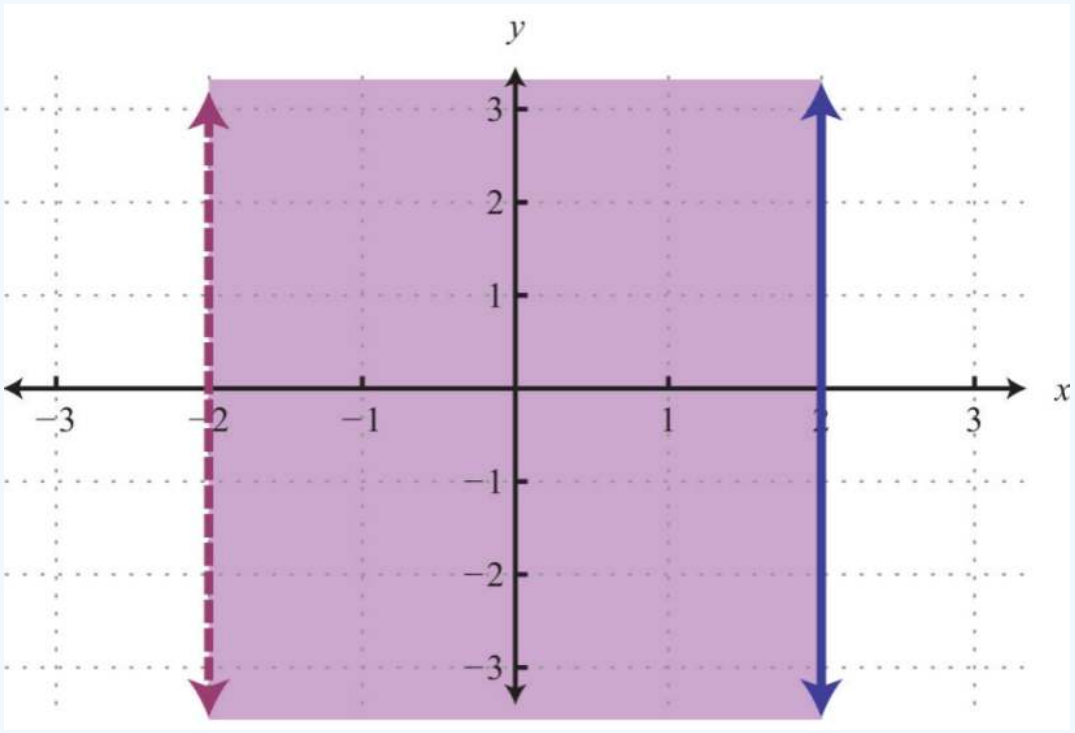


Figure 3.6.1.19

21. No solution, $\emptyset$

23.

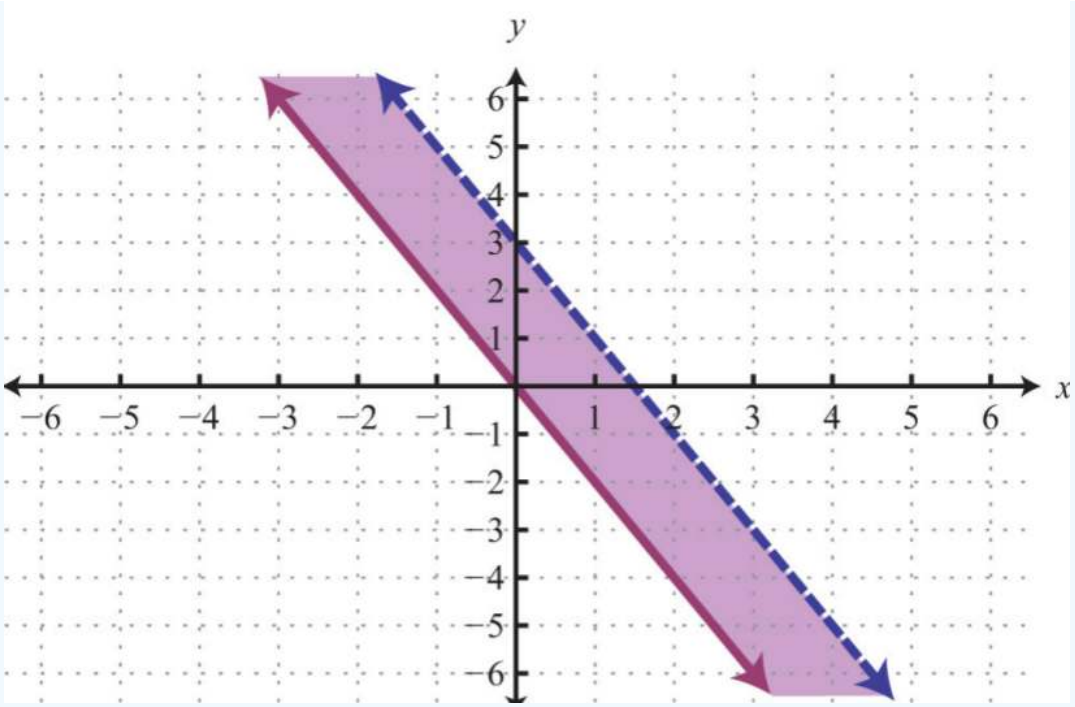


Figure 3.6.1.20

25.

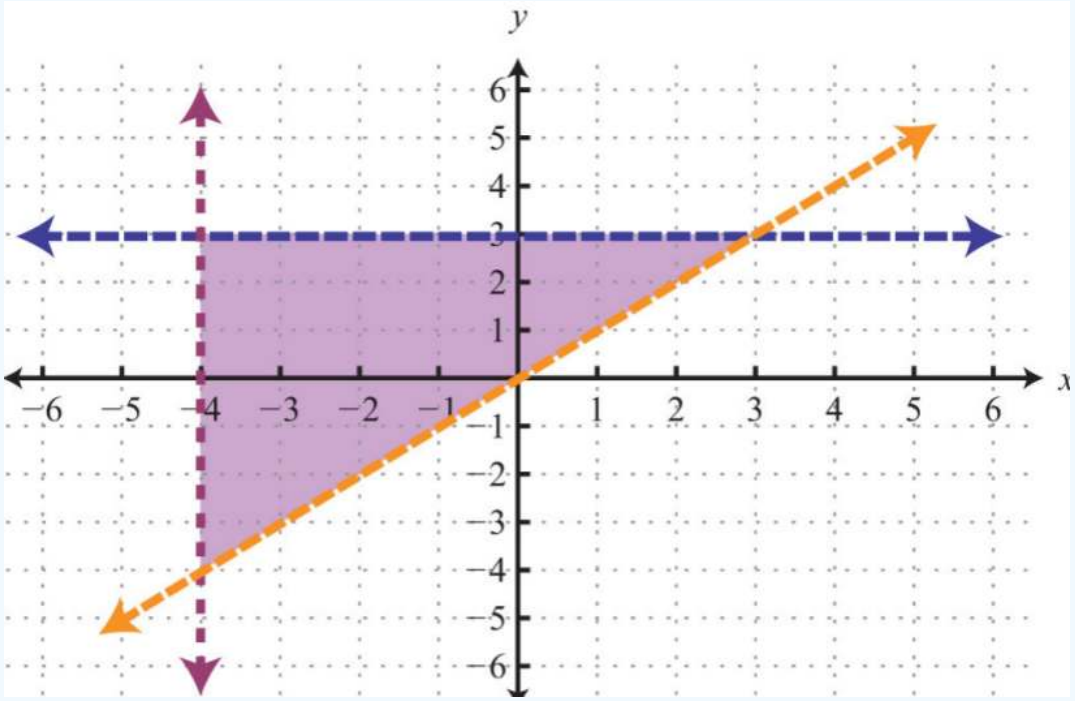


Figure 3.6.1.21

27.

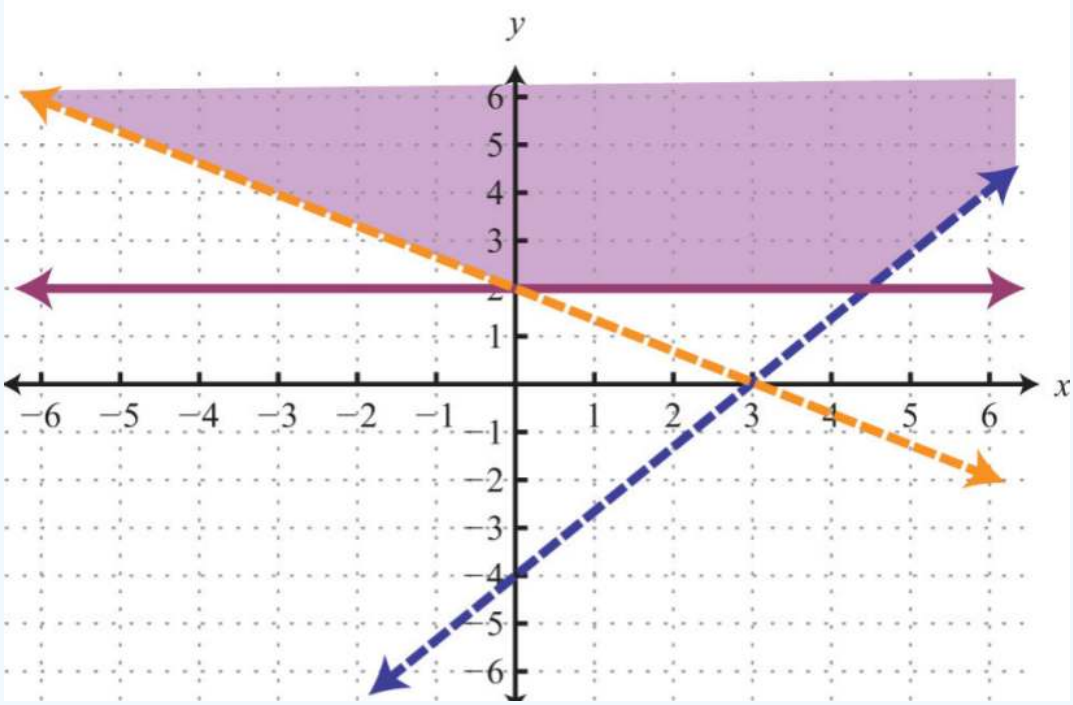


Figure 3.6.1.22

29.

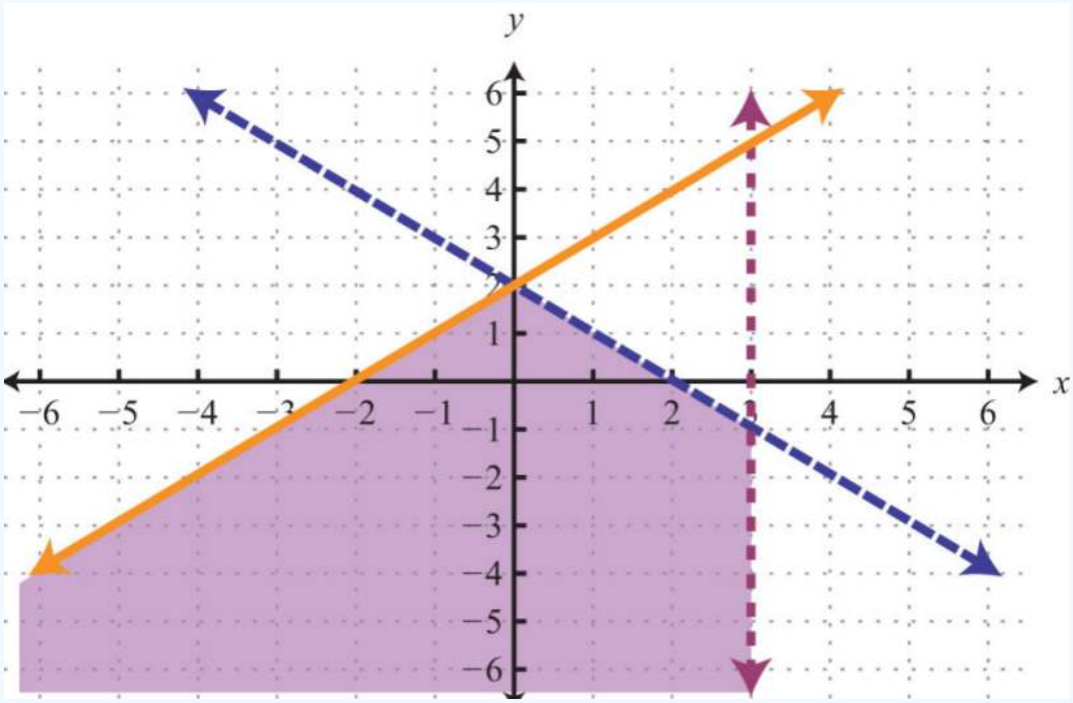


Figure 3.6.1.23

31. $\begin{cases} x > 0 \\ y > 0 \end{cases}$

33. $\begin{cases} x < 0 \\ y < 0 \end{cases}$

This page titled 3.6.1: Solving Systems of Linear Inequalities (Two Variables) is shared under a CC BY-NC-SA 3.0 license and was authored, remixed, and/or curated by Anonymous.

- **4.5: Solving Systems of Linear Inequalities (Two Variables)** by Anonymous is licensed CC BY-NC-SA 3.0. Original source: https://2012books.lardbucket.org/books/beginning-algebra.

3.7: Linear Programming

3.7: Linear Programming is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

3.7.1: Maximization Applications

Learning Objectives

In this section, you will learn to:

1. Recognize the typical form of a linear programing problem
2. Formulate maximization linear programming problems
3. Graph feasibility regions for maximization linear programming problems
4. Determine optimal solutions for maximization linear programming problems.

Application problems in business, economics, and social and life sciences often ask us to make decisions on the basis of certain conditions. The conditions or constraints often take the form of inequalities. In this section, we will begin to formulate, analyze, and solve such problems, at a simple level, to understand the many components of such a problem.

A typical **linear programming** problem consists of finding an extreme value of a linear function subject to certain constraints. We are either trying to maximize or minimize the value of this linear function, such as to maximize profit or revenue, or to minimize cost. That is why these linear programming problems are classified as **maximization** or **minimization problems**, or just **optimization problems**. The function we are trying to optimize is called an **objective function**, and the conditions that must be satisfied are called **constraints**.

A typical example is to maximize profit from producing several products, subject to limitations on materials or resources needed for producing these items; the problem requires us to determine the amount of each item produced. Another type of problem involves scheduling; we need to determine how much time to devote to each of several activities in order to maximize income from (or minimize cost of) these activities, subject to limitations on time and other resources available for each activity.

In this chapter, we will work with problems that involve only two variables, and therefore, can be solved by graphing. In the next chapter, we'll learn an algorithm to find a solution numerically. That will provide us with a tool to solve problems with more than two variables. At that time, with a little more knowledge about linear programming, we'll also explore the many ways these techniques are used in business and wide variety of other fields.

We begin by solving a maximization problem.

Example 3.7.1.1

Niki holds two part-time jobs, Job I and Job II. She never wants to work more than a total of 12 hours a week. She has determined that for every hour she works at Job I, she needs 2 hours of preparation time, and for every hour she works at Job II, she needs one hour of preparation time, and she cannot spend more than 16 hours for preparation.

If Nikki makes $40 an hour at Job I, and $30 an hour at Job II, how many hours should she work per week at each job to maximize her income?

Solution

We start by choosing our variables.

- Let x = The number of hours per week Niki will work at Job I.
- Let y = The number of hours per week Niki will work at Job II.

Now we write the objective function. Since Niki gets paid $40 an hour at Job I, and $30 an hour at Job II, her total income I is given by the following equation.

$$I = 40x + 30y$$

Our next task is to find the constraints. The second sentence in the problem states, "She never wants to work more than a total of 12 hours a week." This translates into the following constraint:

$$x + y \leq 12$$

The third sentence states, "For every hour she works at Job I, she needs 2 hours of preparation time, and for every hour she works at Job II, she needs one hour of preparation time, and she cannot spend more than 16 hours for preparation." The translation follows.

$$2x + y \leq 16$$

The fact that x and y can never be negative is represented by the following two constraints:

$$x \geq 0, \text{ and } y \geq 0.$$

Well, good news! We have formulated the problem. We restate it as

$$\begin{aligned} &\textbf{Maximize} && \mathrm{I} = 40\mathrm{x} + 30\mathrm{y} \\ &\textbf{Subject to:} && \mathrm{x} + \mathrm{y} \leq 12 \\ &&& 2\mathrm{x} + \mathrm{y} \leq 16 \\ &&& \mathrm{x} \geq 0; \mathrm{y} \geq 0 \end{aligned} \tag{3.7.1.1}$$

In order to solve the problem, we graph the constraints and shade the region that satisfies **all** the inequality constraints.

Any appropriate method can be used to graph the lines for the constraints. However often the easiest method is to graph the line by plotting the x-intercept and y-intercept.

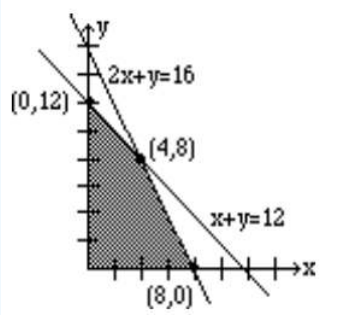
The line for a constraint will divide the plane into two region, one of which satisfies the inequality part of the constraint. A test point is used to determine which portion of the plane to shade to satisfy the inequality. Any point on the plane that is not on the line can be used as a test point.

- If the test point satisfies the inequality, then the region of the plane that satisfies the inequality is the region that contains the test point.
- If the test point does not satisfy the inequality, then the region that satisfies the inequality lies on the opposite side of the line from the test point.

In the graph below, after the lines representing the constraints were graphed using an appropriate method from Chapter 1, the point (0,0) was used as a test point to determine that

- (0,0) satisfies the constraint $x + y \leq 12$ because $0 + 0 < 12$
- (0,0) satisfies the constraint $2x + y \leq 16$ because $2(0) + 0 < 16$

Therefore, in this example, we shade the region that is below and to the left of **both** constraint lines, but also above the x axis and to the right of the y axis, in order to further satisfy the constraints $x \geq 0$ and $y \geq 0$.



The shaded region where all conditions are satisfied is called the **feasibility region** or the feasibility polygon.

The **Fundamental Theorem of Linear Programming** states that the maximum (or minimum) value of the objective function always takes place at the vertices of the feasibility region.

Therefore, we will identify all the vertices (corner points) of the feasibility region. We call these points **critical points.** They are listed as (0, 0), (0, 12), (4, 8), (8, 0). To maximize Niki's income, we will substitute these points in the objective function to see which point gives us the highest income per week. We list the results below.

Critical Points	Income

Critical Points	Income
(0, 0)	40(0) + 30(0) = $0
(0, 12)	40(0) + 30(12) = $360
(4, 8)	40(4) + 30(8) = $400
(8, 0)	40(8) + 30(0) = $320

Clearly, the point (4, 8) gives the most profit: $400.

Therefore, we conclude that Niki should work 4 hours at Job I, and 8 hours at Job II.

Example 3.7.1.2

A factory manufactures two types of gadgets, regular and premium. Each gadget requires the use of two operations, assembly and finishing, and there are at most 12 hours available for each operation. A regular gadget requires 1 hour of assembly and 2 hours of finishing, while a premium gadget needs 2 hours of assembly and 1 hour of finishing. Due to other restrictions, the company can make at most 7 gadgets a day. If a profit of $20 is realized for each regular gadget and $30 for a premium gadget, how many of each should be manufactured to maximize profit?

Solution

We choose our variables.

- Let x = The number of regular gadgets manufactured each day.
- and y = The number of premium gadgets manufactured each day.

The objective function is

$$P = 20x + 30y$$

We now write the constraints. The fourth sentence states that the company can make at most 7 gadgets a day. This translates as

$$x + y \leq 7$$

Since the regular gadget requires one hour of assembly and the premium gadget requires two hours of assembly, and there are at most 12 hours available for this operation, we get

$$x + 2y \leq 12$$

Similarly, the regular gadget requires two hours of finishing and the premium gadget one hour. Again, there are at most 12 hours available for finishing. This gives us the following constraint.

$$2x + y \leq 12$$

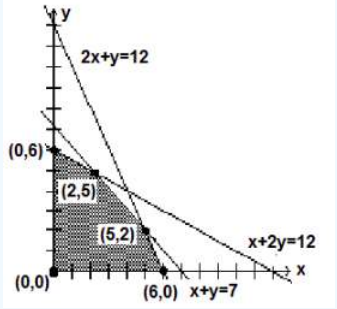
The fact that x and y can never be negative is represented by the following two constraints:

$$x \geq 0, \text{ and } y \geq 0.$$

We have formulated the problem as follows:

$$\begin{aligned} &\textbf{Maximize} && \mathrm{P = 20x + 30y} \\ &\textbf{Subject to:} && \mathrm{x + y \leq 7} \\ &&& \mathrm{x + 2y \leq 12} \\ &&& \mathrm{2x + y \leq 12} \\ &&& \mathrm{x \geq 0; y \geq 0} \end{aligned}$$

In order to solve the problem, we next graph the constraints and feasibility region.



Again, we have shaded the feasibility region, where all constraints are satisfied.

Since the extreme value of the objective function always takes place at the vertices of the feasibility region, we identify all the critical points. They are listed as (0, 0), (0, 6), (2, 5), (5, 2), and (6, 0). To maximize profit, we will substitute these points in the objective function to see which point gives us the maximum profit each day. The results are listed below.

Critical Point	Income
(0, 0)	20(0) + 30(0) = $0
(0, 6)	20(0) + 30(6) = $180
(2, 5)	20(2) + 30(5) = $190
(5, 2)	20(5) + 30(2) = $160
(6,0)	20(6) + 30(0) = $120

The point (2, 5) gives the most profit, and that profit is $190.
Therefore, we conclude that we should manufacture 2 regular gadgets and 5 premium gadgets daily to obtain the maximum profit of $190.

So far we have focused on "**standard maximization problems**" in which

1. The objective function is to be maximized
2. All constraints are of the form $ax + by \leq c$
3. All variables are constrained to by non-negative ($x \geq 0$, $y \geq 0$)

We will next consider an example where that is not the case. Our next problem is said to have "**mixed constraints**", since some of the inequality constraints are of the form $ax + by \leq c$ and some are of the form $ax + by \geq c$. The non-negativity constraints are still an important requirement in any linear program.

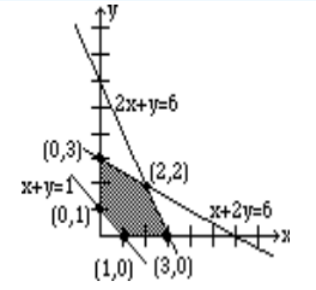
Example 3.7.1.3

Solve the following maximization problem graphically.

$$\begin{aligned} \textbf{Maximize} \quad & P = 10x + 15y \\ \textbf{Subject to:} \quad & x + y \geq 1 \\ & x + 2y \leq 6 \\ & 2x + y \leq 6 \\ & x \geq 0; y \geq 0 \end{aligned}$$

Solution

The graph is shown below.



The five critical points are listed in the above figure. The reader should observe that the first constraint $x+y \geq 1$ requires that the feasibility region must be bounded below by the line $x+y=1$; the test point (0,0) does not satisfy $x+y \geq 1$, so we shade the region on the opposite side of the line from test point (0,0).

Critical point	Income
(1, 0)	10(1) + 15(0) = $10
(3, 0)	10(3) + 15(0) = $30
(2, 2)	10(2) + 15(2) = $50
(0, 3)	10(0) + 15(3) = $45
(0,1)	10(0) + 15(1) = $15

Clearly, the point (2, 2) maximizes the objective function to a maximum value of 50.

It is important to observe that that if the point (0,0) lies on the line for a constraint, then (0,0) could not be used as a test point. We would need to select any other point we want that does not lie on the line to use as a test point in that situation.

Finally, we address an important question. Is it possible to determine the point that gives the maximum value without calculating the value at each critical point?

The answer is yes.

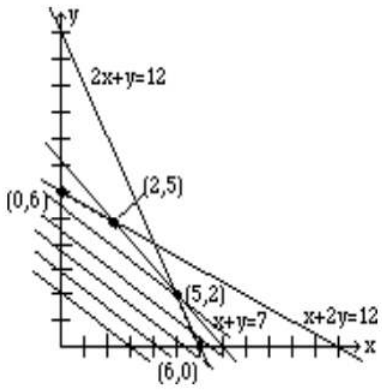
For example $3.7.1.2$ we substituted the points (0, 0), (0, 6), (2, 5), (5, 2), and (6, 0), in the objective function $P=20x+30y$, and we got the values $0, $180, $190, $160, $120, respectively.

Sometimes that is not the most efficient way of finding the optimum solution. Instead we could find the optimal value by also graphing the objective function.

To determine the largest P, we graph $P=20x+30y$ for any value P of our choice. Let us say, we choose $P=60$. We graph $20x+30y=60$.

Now we move the line parallel to itself, that is, keeping the same slope at all times. Since we are moving the line parallel to itself, the slope is kept the same, and the only thing that is changing is the P. As we move away from the origin, the value of P increases. The largest possible value of P is realized when the line touches the last corner point of the feasibility region.

The figure below shows the movements of the line, and the optimum solution is achieved at the point (2, 5). In maximization problems, as the line is being moved away from the origin, this optimum point is the farthest critical point.



We summarize:

The Maximization Linear Programming Problems

1. Write the objective function.
2. Write the constraints.
    1. For the standard maximization linear programming problems, constraints are of the form: $ax + by \leq c$
    2. Since the variables are non-negative, we include the constraints: $x \geq 0$; $y \geq 0$.
3. Graph the constraints.
4. Shade the feasibility region.
5. Find the corner points.
6. Determine the corner point that gives the maximum value.
    a. This is done by finding the value of the objective function at each corner point.
    b. This can also be done by moving the line associated with the objective function.

This page titled 3.7.1: Maximization Applications is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rupinder Sekhon and Roberta Bloom via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1: Maximization Applications** by Rupinder Sekhon and Roberta Bloom is licensed CC BY 4.0. Original source: https://www.deanza.edu/faculty/bloomroberta/math11/afm3files.html.html.

3.7.1.1: Maximization Applications (Exercises)

For the following maximization problems, choose your variables, write the objective function and the constraints, graph the constraints, shade the feasibility region, label all critical points, and determine the solution that optimizes the objective function.

1) A farmer has 100 acres of land on which she plans to grow wheat and corn. Each acre of wheat requires 4 hours of labor and $20 of capital, and each acre of corn requires 16 hours of labor and $40 of capital. The farmer has at most 800 hours of labor and $2400 of capital available. If the profit from an acre of wheat is $80 and from an acre of corn is $100, how many acres of each crop should she plant to maximize her profit?

2) Mr. Tran has $24,000 to invest, some in bonds and the rest in stocks. He has decided that the money invested in bonds must be at least twice as much as that in stocks. But the money invested in bonds must not be greater than $18,000. If the bonds earn 6%, and the stocks earn 8%, how much money should he invest in each to maximize profit?

3) A factory manufactures chairs and tables, each requiring the use of three operations: Cutting, Assembly, and Finishing. The first operation can be used at most 40 hours; the second at most 42 hours; and the third at most 25 hours. A chair requires 1 hour of cutting, 2 hours of assembly, and 1 hour of finishing; a table needs 2 hours of cutting, 1 hour of assembly, and 1 hour of finishing. If the profit is $20 per unit for a chair and $30 for a table, how many units of each should be manufactured to maximize revenue?

4) The Silly Nut Company makes two mixtures of nuts: Mixture A and Mixture B. A pound of Mixture A contains 12 oz of peanuts, 3 oz of almonds and 1 oz of cashews and sells for $4. A pound of Mixture B contains 12 oz of peanuts, 2 oz of almonds and 2 oz of cashews and sells for $5. The company has 1080 lb. of peanuts, 240 lb. of almonds, 160 lb. of cashews. How many pounds of each of mixtures A and B should the company make to maximize profit?
(Hint: Use consistent units. Work the entire problem in pounds by converting all values given in ounces into fractions of pounds).

5)

$$\begin{aligned} \text{Maximize:} \quad & Z = 4x + 10y \\ \text{Subject to:} \quad & x + y \leq 5 \\ & 2x + y \leq 8 \\ & x + 2y \leq 8 \\ & x \geq 0, y \geq 0 \end{aligned}$$

6) This maximization linear programming problem is not in “standard” form. It has mixed constraints, some involving $\leq$ inequalities and some involving $\geq$ inequalities. However with careful graphing, we can solve this using the techniques we have learned in this section.

$$\begin{aligned} \text{Maximize:} \quad & Z = 5x + 7y \\ \text{Subject to:} \quad & x + y \leq 30 \\ & 2x + y \leq 50 \\ & 4x + 3y \geq 60 \\ & 2x \geq y \\ & x \geq 0, y \geq 0 \end{aligned}$$

This page titled 3.7.1.1: Maximization Applications (Exercises) is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rupinder Sekhon and Roberta Bloom via source content that was edited to the style and standards of the LibreTexts platform.

- **3.1.1: Maximization Applications (Exercises)** by Rupinder Sekhon and Roberta Bloom is licensed CC BY 4.0. Original source: https://www.deanza.edu/faculty/bloomroberta/math11/afm3files.html.html.

3.7.2: Minimization Applications

Learning Objectives

In this section, you will learn to:

1. Formulate minimization linear programming problems
2. Graph feasibility regions for maximization linear programming problems
3. Determine optimal solutions for maximization linear programming problems.

Minimization linear programming problems are solved in much the same way as the maximization problems.

For the **standard minimization linear program**, the constraints are of the form $ax + by \geq c$, as opposed to the form $ax + by \leq c$ for the standard maximization problem. As a result, the feasible solution extends indefinitely to the upper right of the first quadrant, and is unbounded. But that is not a concern, since in order to minimize the objective function, the line associated with the objective function is moved towards the origin, and the critical point that minimizes the function is closest to the origin.

However, one should be aware that in the case of an unbounded feasibility region, the possibility of no optimal solution exists.

Example 3.7.2.1

At a university, Professor Symons wishes to employ two people, John and Mary, to grade papers for his classes. John is a graduate student and can grade 20 papers per hour; John earns $15 per hour for grading papers. Mary is an post-doctoral associate and can grade 30 papers per hour; Mary earns $25 per hour for grading papers. Each must be employed at least one hour a week to justify their employment.
If Prof. Symons has at least 110 papers to be graded each week, how many hours per week should he employ each person to minimize the cost?

Solution

We choose the variables as follows:

Let x = The number of hours per week John is employed.

and y = The number of hours per week Mary is employed.

The objective function is

$$C = 15x + 25y$$

The fact that each must work at least one hour each week results in the following two constraints:

$$x \geq 1$$
$$y \geq 1$$

Since John can grade 20 papers per hour and Mary 30 papers per hour, and there are at least 110 papers to be graded per week, we get

$$20x + 30y \geq 110$$

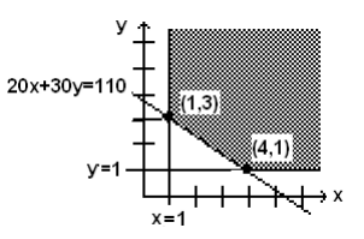
The fact that x and y are non-negative, we get

$$x \geq 0, \text{and } y \geq 0.$$

The problem has been formulated as follows.

$$\begin{aligned} &\textbf{Minimize} && \mathrm{C = 15x + 25y} \\ &\textbf{Subject to:} && \mathrm{x \geq 1} \\ &&& \mathrm{y \geq 1} \\ &&& \mathrm{20x + 30y \geq 110} \\ &&& \mathrm{x \geq 0; y \geq 0} \end{aligned}$$

To solve the problem, we graph the constraints as follows:



Again, we have shaded the feasibility region, where all constraints are satisfied.

If we used test point (0,0) that does not lie on any of the constraints, we observe that (0, 0) **does not** satisfy any of the constraints $x \geq 1$, $y \geq 1$, $20x + 30y \geq 110$. Thus all the shading for the feasibility region lies on the opposite side of the constraint lines from the point (0,0).

Alternatively we could use test point (4,6), which also does not lie on any of the constraint lines. We'd find that (4,6) **does** satisfy all of the inequality constraints. Consequently all the shading for the feasibility region lies on the same side of the constraint lines as the point (4,6).

Since the extreme value of the objective function always takes place at the vertices of the feasibility region, we identify the two critical points, (1, 3) and (4, 1). To minimize cost, we will substitute these points in the objective function to see which point gives us the minimum cost each week. The results are listed below.

Critical points	Income
(1, 3)	15(1) + 25(3) = $90
(4, 1)	15(4) + 25(1) = $85

The point (4, 1) gives the least cost, and that cost is $85. Therefore, we conclude that in order to minimize grading costs, Professor Symons should employ John 4 hours a week, and Mary 1 hour a week at a cost of $85 per week.

Example 3.7.2.2

Professor Hamer is on a low cholesterol diet. During lunch at the college cafeteria, he always chooses between two meals, Pasta or Tofu. The table below lists the amount of protein, carbohydrates, and vitamins each meal provides along with the amount of cholesterol he is trying to minimize. Mr. Hamer needs at least 200 grams of protein, 960 grams of carbohydrates, and 40 grams of vitamins for lunch each month. Over this time period, how many days should he have the Pasta meal, and how many days the Tofu meal so that he gets the adequate amount of protein, carbohydrates, and vitamins and at the same time minimizes his cholesterol intake?

	PASTA	**TOFU**
PROTEIN	8g	16g
CARBOHYDRATES	60g	40g
VITAMIN C	2g	2g
CHOLESTEROL	60mg	50mg

Solution

We choose the variables as follows.

Let x = The number of days Mr. Hamer eats Pasta.

and y = The number of days Mr. Hamer eats Tofu.

Since he is trying to minimize his cholesterol intake, our objective function represents the total amount of cholesterol C provided by both meals.

$$C = 60x + 50y$$

The constraint associated with the total amount of protein provided by both meals is

$$8x + 16y \geq 200$$

Similarly, the two constraints associated with the total amount of carbohydrates and vitamins are obtained, and they are

$$60x + 40y \geq 960$$
$$2x + 2y \geq 40$$

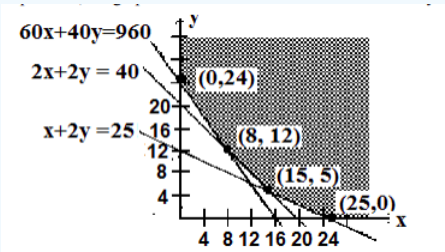
The constraints that state that x and y are non-negative are

$$x \geq 0, \text{ and } y \geq 0.$$

We summarize all information as follows:

$$\begin{array}{ll} \textbf{Minimize} & \mathrm{C} = 60\mathrm{x} + 50\mathrm{y} \\ \textbf{Subject to:} & 8\mathrm{x} + 16\mathrm{y} \geq 200 \\ & 60\mathrm{x} + 40\mathrm{y} \geq 960 \\ & 2\mathrm{x} + 2\mathrm{y} \geq 40 \\ & \mathrm{x} \geq 0; \mathrm{y} \geq 0 \end{array}$$

To solve the problem, we graph the constraints and shade the feasibility region.



We have shaded the unbounded feasibility region, where all constraints are satisfied.

To minimize the objective function, we find the vertices of the feasibility region. These vertices are (0, 24), (8, 12), (15, 5) and (25, 0). To minimize cholesterol, we will substitute these points in the objective function to see which point gives us the smallest value. The results are listed below.

Critical points	Income
(0, 24)	60(0) + 50(24) = 1200
(8, 12)	60(8) + 50(12) = 1080
(15, 5)	60(15) + 50(5) = 1150
(25, 0)	60(25) + 50(0) = 1500

The point (8, 12) gives the least cholesterol, which is 1080 mg. This states that for every 20 meals, Professor Hamer should eat Pasta 8 days, and Tofu 12 days.

We must be aware that in some cases, a linear program may not have an optimal solution.

- A linear program can fail to have an optimal solution is if there is not a feasibility region. If the inequality constraints are not compatible, there may not be a region in the graph that satisfies **all** the constraints. If the linear program does not have a feasible solution satisfying all constraints, then it can not have an optimal solution.
- A linear program can fail to have an optimal solution if the feasibility region is unbounded.

- The two minimization linear programs we examined had unbounded feasibility regions. The feasibility region was bounded by constraints on some sides but was not entirely enclosed by the constraints. Both of the minimization problems had optimal solutions.
- However, if we were to consider a maximization problem with a similar unbounded feasibility region, the linear program would have no optimal solution. No matter what values of x and y were selected, we could always find other values of x and y that would produce a higher value for the objective function. In other words, if the value of the objective function can be increased without bound in a linear program with an unbounded feasible region, there is no optimal maximum solution.

Although the method of solving minimization problems is similar to that of the maximization problems, we still feel that we should summarize the steps involved.

Minimization Linear Programming Problems

1. Write the objective function.
2. Write the constraints.
   a. For standard minimization linear programming problems, constraints are of the form: $ax + by \geq c$
   b. Since the variables are non-negative, include the constraints: $x \geq 0$; $y \geq 0$.
3. Graph the constraints.
4. Shade the feasibility region.
5. Find the corner points.
6. Determine the corner point that gives the minimum value.
   a. This can be done by finding the value of the objective function at each corner point.
   b. This can also be done by moving the line associated with the objective function.
   c. There is the possibility that the problem has no solution

---

This page titled 3.7.2: Minimization Applications is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by Rupinder Sekhon and Roberta Bloom via source content that was edited to the style and standards of the LibreTexts platform.

- **3.2: Minimization Applications** by Rupinder Sekhon and Roberta Bloom is licensed CC BY 4.0. Original source: https://www.deanza.edu/faculty/bloomroberta/math11/afm3files.html.html.

CHAPTER OVERVIEW

4: Sets and Counting

4.1: Sets and Operations

4.1.1: Basics of Sets

4.1.2: Union, Intersection, and Complement

4.1.3: Venn Diagrams

4.2: Number of Elements in a Set

4.2.1: Cardinality

4.3: Multiplication Principle

4.3.1: Tree Diagrams, Tables, and Outcomes

4.3.2: The Multiplication Rule for Counting

4.4: Permutations and Combinations

4.4.1: Permutations

4.4.2: Combinations

4: Sets and Counting is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

4.1: Sets and Operations

4.1: Sets and Operations is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

4.1.1: Basics of Sets

An art collector might own a collection of paintings, while a music lover might keep a collection of CDs. Any collection of items can form a **set**.

Set

A **set** is a collection of distinct objects, called **elements** of the set

A set can be defined by describing the contents, or by listing the elements of the set, enclosed in curly brackets.

Example 1

Some examples of sets defined by describing the contents:

a. The set of all even numbers
b. The set of all books written about travel to Chile

Some examples of sets defined by listing the elements of the set:

a. {1, 3, 9, 12}
b. {red, orange, yellow, green, blue, indigo, purple}

A set simply specifies the contents; order is not important. The set represented by {1, 2, 3} is equivalent to the set {3, 1, 2}.

Notation

Commonly, we will use a variable to represent a set, to make it easier to refer to that set later.

The symbol $\in$ means "is an element of".

A set that contains no elements, $\{\}$, is called the **empty set** and is notated $\emptyset$

Example 2

Let $A = \{1, 2, 3, 4\}$

To notate that 2 is element of the set, we'd write $2 \in A$

Sometimes a collection might not contain all the elements of a set. For example, Chris owns three Madonna albums. While Chris's collection is a set, we can also say it is a **subset** of the larger set of all Madonna albums.

Subset

A **subset** of a set A is another set that contains only elements from the set A, but may not contain all the elements of A.

If B is a subset of A, we write $B \subseteq A$

A **proper subset** is a subset that is not identical to the original set – it contains fewer elements.

If B is a proper subset of A, we write $B \subset A$

Example 3

Consider these three sets

$A =$ the set of all even numbers $\quad B = \{2, 4, 6\} \quad C = \{2, 3, 4, 6\}$

Here $B \subset A$ since every element of B is also an even number, so is an element of A.

More formally, we could say $B \subset A$ since if $x \in B$, then $x \in A$

It is also true that $B \subset C$.

C is not a subset of A, since C contains an element, 3 , that is not contained in A

Example 4

Suppose a set contains the plays “Much Ado About Nothing”, “MacBeth”, and “A Midsummer’s Night Dream”. What is a larger set this might be a subset of?

Solution

There are many possible answers here. One would be the set of plays by Shakespeare. This is also a subset of the set of all plays ever written. It is also a subset of all British literature.

Try it Now 1

The set $A = \{1, 3, 5\}$. What is a larger set this might be a subset of?

Answer

There are several answers: The set of all odd numbers less than 10. The set of all odd numbers. The set of all integers. The set of all real numbers.

This page titled 4.1.1: Basics of Sets is shared under a CC BY-SA 3.0 license and was authored, remixed, and/or curated by David Lippman (The OpenTextBookStore) via source content that was edited to the style and standards of the LibreTexts platform.

- **13.1: Basics of Sets** by David Lippman is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.

4.1.2: Union, Intersection, and Complement

Commonly sets interact. For example, you and a new roommate decide to have a house party, and you both invite your circle of friends. At this party, two sets are being combined, though it might turn out that there are some friends that were in both sets.

Union, Interection, and Complement

The **union** of two sets contains all the elements contained in either set (or both sets).

The union is notated $A \cup B$

More formally, $x \in A \cup B$ if $x \in A$ or $x \in B$ (or both)

The **intersection** of two sets contains only the elements that are in both sets.

The intersection is notated $A \cap B$

More formally, $x \in A \cap B$ if $x \in A$ and $x \in B$

The **complement** of a set *A* contains everything that is *not* in the set *A*.

The complement is notated A, or A^c, or sometimes $\sim A$.

Example 5

Consider the sets:

$A = \{ \text{red, green, blue} \}$ $B = \{ \text{red, yellow, orange} \}$ $C = \{ \text{red, orange, yellow, green, blue, purple} \}$

a. Find $A \cup B$
b. Find $A \cap B$
c. Find $A^c \cap C$

Solution

a) The union contains all the elements in either set: $A \cup B = \{ \text{red, green, blue, yellow, orange} \}$

Notice we only list red once.

b) The intersection contains all the elements in both sets: $A \cap B = \{ \text{red} \}$

c) Here we're looking for all the elements that are not in set A and are also in C.

$A^c \cap C = \{ \text{orange, yellow, purple} \}$

Try it Now 2

Using the sets from the previous example, find $A \cup C$ and $B^c \cap A$

Answer

$A \cup C = \{ \text{red, orange, yellow, green, blue purple} \}$

$B^c \cap A = \{ \text{green, blue} \}$

Notice that in the example above, it would be hard to just ask for A^c, since everything from the color fuchsia to puppies and peanut butter are included in the complement of the set. For this reason, complements are usually only used with intersections, or when we have a universal set in place.

Universal Set

A **universal set** is a set that contains all the elements we are interested in. This would have to be defined by the context.

A complement is relative to the universal set, so A^C contains all the elements in the universal set that are not in A.

Example 6

a. If we were discussing searching for books, the universal set might be all the books in the library.
b. If we were grouping your Facebook friends, the universal set would be all your Facebook friends.
c. If you were working with sets of numbers, the universal set might be all whole numbers, all integers, or all real numbers

Example 7

Suppose the universal set is $U =$ all whole numbers from 1 to 9. If $A = \{1, 2, 4\}$, then

$A^c = \{3, 5, 6, 7, 8, 9\}$

As we saw earlier with the expression $A^c \cap C$, set operations can be grouped together. Grouping symbols can be used like they are with arithmetic - to force an order of operations.

Example 8

Suppose

$H = \{ \text{cat, dog, rabbit, mouse} \}, F = \{ \text{dog, cow, duck, pig, rabbit} \} \quad W = \{ \text{duck, rabbit, deer, frog, mouse} \}$

a. Find $(H \cap P) \cup W$
b. Find $H \cap (F \cup W)$
c. Find $(H \cap P) \cap W$

Solution

a) We start with the intersection: $H \cap F = \{ \text{dog, rabbit} \}$

Now we union that result with W : $(H \cap F) \cup W = \{\text{dog, duck, rabbit, deer, frog, mouse} \}$

b) We start with the union: $F \cup W = \{\text{dog, cow, rabbit, duck, pig, deer, frog, mouse} \}$

Now we intersect that result with H : $H \cap (F \cup W) = \{ \text{dog, rabbit, mouse} \}$

c) We start with the intersection: $H \cap F = \{\text{dog, rabbit} \}$

Now we want to find the elements of W that are not in $\text{H} \cap F$

$(H \cap P)^c \cap W = \{ \text{duck, deer, frog, mouse} \}$

This page titled 4.1.2: Union, Intersection, and Complement is shared under a CC BY-SA 3.0 license and was authored, remixed, and/or curated by David Lippman (The OpenTextBookStore) via source content that was edited to the style and standards of the LibreTexts platform.

- **13.2: Union, Intersection, and Complement** by David Lippman is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.

4.1.3: Venn Diagrams

To visualize the interaction of sets, John Venn in 1880 thought to use overlapping circles, building on a similar idea used by Leonhard Euler in the 18th century. These illustrations now called **Venn Diagrams**.

Venn Diagram

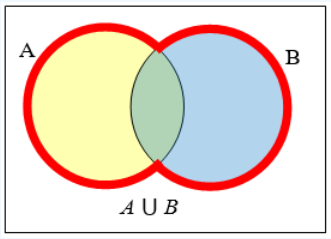
A Venn diagram represents each set by a circle, usually drawn inside of a containing box representing the universal set. Overlapping areas indicate elements common to both sets.

Basic Venn diagrams can illustrate the interaction of two or three sets.

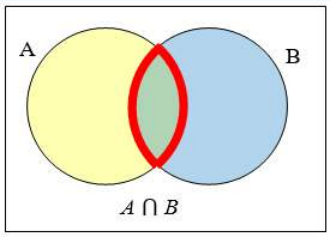
Example 9

Create Venn diagrams to illustrate $A \cup B$, $A \cap B$, and $A^c \cap B$

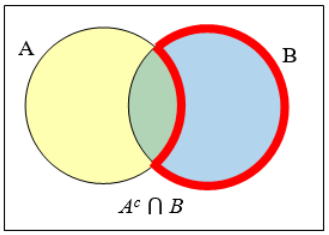
$A \cup B$ contains all elements in either set.



$A \cup B$ contains all elements in either set.



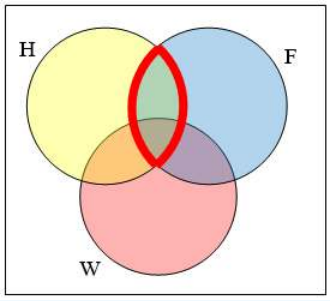
$A \cap B$ contains only those elements in both sets - in the overlap of the circles.



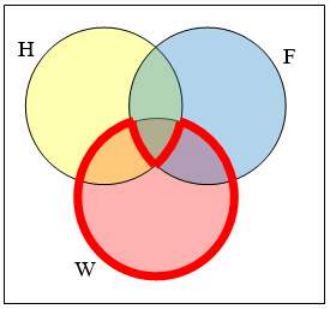
Example 10

Use a Venn diagram to illustrate $(H \cap P)^c \cap W$

We'll start by identifying everything in the set $H \cap P$



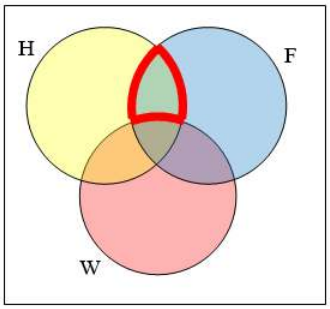
Now, $(H \cap P)^c \cap W$ will contain everything not in the set identified above that is also in set W



Example 11

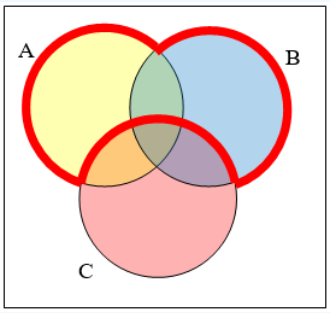
Create an expression to represent the outlined part of the Venn diagram shown.

The elements in the outlined set are in sets H and F, but are not in set W. So we could represent this set as $H \cap F \cap W$



Try it Now 3

Create an expression to represent the outlined portion of the Venn diagram shown



Answer

$A \cup B \cap C^c$

This page titled 4.1.3: Venn Diagrams is shared under a CC BY-SA 3.0 license and was authored, remixed, and/or curated by David Lippman (The OpenTextBookStore) via source content that was edited to the style and standards of the LibreTexts platform.

- **13.3: Venn Diagrams** by David Lippman is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.

4.2: Number of Elements in a Set

4.2: Number of Elements in a Set is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

4.2.1: Cardinality

Often times we are interested in the number of items in a set or subset. This is called the cardinality of the set.

Cardinality

The number of elements in a set is the cardinality of that set.

The cardinality of the set A is often notated as $|A|$ or $n(A)$

Example 12

Let $A = \{1, 2, 3, 4, 5, 6\}$and $B = \{2, 4, 6, 8\}$

What is the cardinality of B?$A \cup B$, $A \cap B$?

Solution

The cardinality of B is 4, since there are 4 elements in the set.

The cardinality of $A \cup B$ is 7, since $A \cup B = \{1, 2, 3, 4, 5, 6, 8\}$, which contains 7 elements.

The cardinality of $A \cap B$ is 3 , since $A \cap B = \{2, 4, 6\}$, which contains 3 elements.

Example 13

What is the cardinality of $P =$ the set of English names for the months of the year?

Solution

The cardinality of this set is 12, since there are 12 months in the year.

Sometimes we may be interested in the cardinality of the union or intersection of sets, but not know the actual elements of each set. This is common in surveying.

Example 14

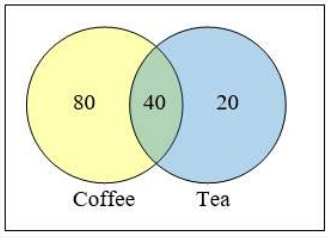
A survey asks 200 people “What beverage do you drink in the morning”, and offers choices:

- Tea only
- Coffee only
- Both coffee and tea

Suppose 20 report tea only, 80 report coffee only, 40 report both. How many people drink tea in the morning? How many people drink neither tea or coffee?

Solution

This question can most easily be answered by creating a Venn diagram. We can see that we can find the people who drink tea by adding those who drink only tea to those who drink both: 60 people.



We can also see that those who drink neither are those not contained in the any of the three other groupings, so we can count those by subtracting from the cardinality of the universal set, 200.

$200 - 20 - 80 - 40 = 60$ people who drink neither.

Example 15

A survey asks: Which online services have you used in the last month:

- Twitter
- Facebook
- Have used both

The results show 40% of those surveyed have used Twitter, 70% have used Facebook, and 20% have used both. How many people have used neither Twitter or Facebook?

Solution

Let T be the set of all people who have used Twitter, and F be the set of all people who have used Facebook. Notice that while the cardinality of F is 70% and the cardinality of T is 40%, the cardinality of $F \cup T$ is not simply $70\% + 40\%$, since that would count those who use both services twice. To find the cardinality of $F \cup T$, we can add the cardinality of F and the cardinality of T, then subtract those in intersection that we've counted twice. In symbols,

$\mathrm{n}(F \cup T) = \mathrm{n}(F) + \mathrm{n}(T) - \mathrm{n}(F \cap T)$

$\mathrm{n}(F \cup T) = 70\% + 40\% - 20\% = 90\%$

Now, to find how many people have not used either service, we're looking for the cardinality of $(F \cup T)^c$. since the universal set contains 100% of people and the cardinality of $F \cup T = 90\%$, the cardinality of $(F \cup 7)^c$ must be the other 10%

The previous example illustrated two important properties

Cardinality properties

$\mathrm{n}(A \cup B) = \mathrm{n}(A) + \mathrm{n}(B) - \mathrm{n}(A \cap B)$

$n\left(A^{\circ}\right) = n(U) - n(A)$

Notice that the first property can also be written in an equivalent form by solving for the cardinality of the intersection:

$\mathrm{n}(A \cap B) = \mathrm{n}(A) + \mathrm{n}(B) - \mathrm{n}(A \cup B)$

Example 16

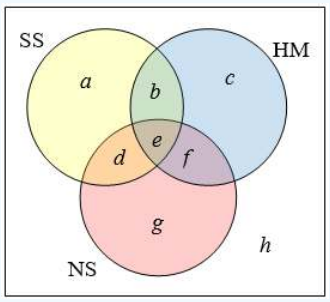
Add text here.Fifty students were surveyed, and asked if they were taking a social science (SS), humanities (HM) or a natural science (NS) course the next quarter.

21 were taking a SS course	26 were taking a HM course
19 were taking a NS course	9 were taking SS and HM
7 were taking SS and NS	10 were taking HM and NS
3 were taking all three	7 were taking none

How many students are only taking a SS course?

Solution

It might help to look at a Venn diagram.



From the given data, we know that there are 3 students in region e and 7 students in region h

since 7 students were taking a SS and NS course, we know that $n(d) + n(e) = 7$. since we know there are 3 students in region 3 , there must be
$7 - 3 = 4$ students in region d

Similarly, since there are 10 students taking HM and NS, which includes regions e and f, there must be

$10 - 3 = 7$ students in region f

Since 9 students were taking SS and HM, there must be $9-3=6$ students in region b

Now, we know that 21 students were taking a SS course. This includes students from regions a, b, d, and e. since we know the number of students in all but region a, we can determine that $21-6-4-3=8$ students are in region a

8 students are taking only a SS course.

Try it Now 4

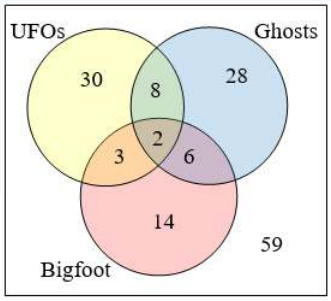
One hundred fifty people were surveyed and asked if they believed in UFOs, ghosts, and Bigfoot.

43 believed in UFOs	44 believed in ghosts
25 believed in Bigfoot	10 believed in UFOs and ghosts
8 believed in ghosts and Bigfoot	5 believed in UFOs and Bigfoot
2 believed in all three	

How many people surveyed believed in at least one of these things?

Answer

Starting with the intersection of all three circles, we work our way out. since 10 people believe in UFOs and Ghosts, and 2 believe in all three, that leaves 8 that believe in only UFOs and Ghosts. We work our way out, filling in all the regions. Once we have, we can add up all those regions, getting 91 people in the union of all three sets. This leaves $150-91=59$ who believe in none.



This page titled 4.2.1: Cardinality is shared under a CC BY-SA 3.0 license and was authored, remixed, and/or curated by David Lippman (The OpenTextBookStore) via source content that was edited to the style and standards of the LibreTexts platform.

- **13.4: Cardinality** by David Lippman is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.

4.3: Multiplication Principle

4.3: Multiplication Principle is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

4.3.1: Tree Diagrams, Tables, and Outcomes



Figure 4.3.1.1: In genetics, the characteristics of an offspring organism depends on the characteristics of its parents. (credit: "Pea Plant" by Maria Keays/Flickr, CC BY 2.0))

Learning Objectives

1. Determine the sample space of single stage experiment.
2. Use tables to list possible outcomes of a multistage experiment.
3. Use tree diagrams to list possible outcomes of a multistage experiment.

In the 19th century, an Augustinian friar and scientist named Gregor Mendel used his observations of pea plants to set out his theory of genetic propagation. In his work, he looked at the offspring that resulted from breeding plants with different characteristics together. For applications like this, it is often insufficient to only know in how many ways a process might end; we need to be able to list all of the possibilities. As we've seen, the number of possible outcomes can be very large! Thus, it's important to have a strategy that allows us to systematically list these possibilities to make sure we don't leave any out. In this section, we'll look at two of these strategies.

Single Stage Experiments

When we are talking about combinatorics or probability, the word "experiment" has a slightly different meaning than it does in the sciences. Experiments can range from very simple ("flip a coin") to very complex ("count the number of uranium atoms that undergo nuclear fission in a sample of a given size over the course of an hour"). Experiments have unknown outcomes that generally rely on something random, so that if the experiment is repeated (or replicated) the outcome might be different. No matter what the experiment, though, analysis of the experiment typically begins with identifying its sample space.

The sample space of an experiment is the set of all of the possible outcomes of the experiment, so it's often expressed as a set (i.e., as a list bound by braces; if the experiment is "randomly select a number between 1 and 4," the sample space would be written $\{1, 2, 3, 4\}$).

Example 4.3.1.1: Finding the Sample Space of an Experiment

For each of the following experiments, identify the sample space.

1. Flip a coin (which has 2 faces, typically called “heads” and “tails”) and note which face is up.
2. Flip a coin 10 times and count the number of heads.
3. Roll a 6-sided die and note the number that is on top.
4. Roll two 6-sided dice and note the sum of the numbers on top.

Answer

1. If we use “H” to denote “heads is facing up” and “T” to denote “tails is facing up”, then the sample space is {H, T}.
2. It’s possible (though unlikely) that there will be no heads flipped; the outcome in that case would be “0.” It’s also possible (more likely, but still quite unlikely) that only one flip will result in heads. Any other whole number is possible, up to the maximum: We’re flipping the coin 10 times, so we can’t get any more than 10 heads. So, the sample space is {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10}.
3. There are 6 numbers on the die: 1, 2, 3, 4, 5, and 6. So, the sample space for a single roll of the die is {1, 2, 3, 4, 5, 6}.
4. If we roll 2 dice, the smallest possible sum we could get is $1+1=2$ and the biggest is $6+6=12$. Every other whole number between those two is possible. So, the sample space is {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}.

Your Turn 4.3.1.1

Identify the sample space of each experiment.

You draw a card from a standard deck and note its suit.

You draw a card from a standard deck and note its rank.

You roll a 4-sided die and note the number on the bottom. (A 4-sided die is shaped like a pyramid, so when it comes to rest, there’s no single side facing up.)

You roll three 4-sided dice and note the sum of the numbers on the bottom.

Multistage Experiments

Some experiments have more complicated sample spaces because they occur in stages. These stages can occur in succession (like drawing cards one at a time) or simultaneously (rolling 2 dice). Sample spaces get more complicated as the complexity of the experiment increases, so it’s important to choose a systematic method for identifying all of the possible outcomes. The first method we’ll discuss is the table.

Using Tables to Find Sample Spaces

Tables are useful for finding the sample space for experiments that meet two criteria: (1) The experiment must have only two stages, and (2) the outcomes of each stage must have no effect on the outcomes of the other. When the stages do not affect each other, we say the stages are independent. Otherwise, the stages are dependent and so we can’t use tables; we’ll look at a method for analyzing dependent stages soon.

Example 4.3.1.2: Determining Independence

Decide whether the two stages in these experiments are independent or dependent.

1. You flip a coin and note the result, and then flip the coin again and note the result.
2. You draw 2 cards from a standard deck (52 cards), one at a time.

Answer

1. No matter what happens on the first flip, the second flip has the same sample space: {H, T} (You’ll sometimes hear the phrase “The coin has no memory”). So, these stages are independent.

2. Let's say that the first card you draw is A♠. The sample space for the second draw consists of all the cards *except* A♠ (since that card is no longer in the deck, you can't draw it again). If instead that first card was 2♡, the sample space for the second draw is different: it's every card except 2♡. Since the sample space for the second card changes based on the result of the first draw, these stages are dependent.

Your Turn 4.3.1.2

Decide whether the two stages in these experiments are dependent or independent.

You're getting dressed to go to a party, and you plan to wear a blouse and a skirt. You choose the blouse first, then the skirt (assume that you'd be comfortable wearing any of your skirts with any of your blouses).

On further reflection, you realize that some of your skirts clash with some of your blouses. So, you choose the blouse first, and then choose a skirt that goes with your chosen blouse.

If you have a two-stage experiment with independent stages, a table is the most straightforward way to identify the sample space. To build a table, you list the outcomes of one stage of the experiment along the top of the table and the outcomes of the other stage down the side. The cells in the interior of the table are then filled using the outcomes associated with each cell's row and column. Let's look at an example.

Example 4.3.1.3: Using Tables to Identify Sample Spaces

Identify the sample spaces of these experiments using tables.

1. You roll two dice: one 4-sided and one 6-sided.
2. You're in an ice-cream shop, and you're going to get a single scoop of ice cream with a topping. The flavors of ice cream you're considering are vanilla, chocolate, and rocky road; the toppings are fudge, whipped cream, and sprinkles.
3. The pea plants you're breeding have two possible pod colors: green and yellow. These colors are decided by a particular gene, which comes in two types: "G" for green, and "g" for yellow (In genetics, capital letters usually denote *dominant* genes, while lower-case letters denote *recessive* genes). Each plant has two genes. If you breed a Gg pea plant with a gg plant, the offspring plant will get one gene from each parent. What are the possible outcomes?

Answer

1. **Step 1:** Make the outline of a table, with the results of the 4-sided roll on one side and the results of the 6-sided roll on the other. In practice, it doesn't matter which you choose; for this example, we'll put the 4-sided results on top (labeling the columns of the chart) and the 6-sided results on the side (labeling the rows of the chart) as shown in the following table:

		4-Sided Roll			
		1	2	3	4
6-Sided Roll	**1**				
	2				
	3				
	4				
	5				
	6				

Step 2: Fill in the results in each cell of the table below. Use the notation of an ordered pair, with the row label first.

	4-Sided Roll			
	1	2	3	4

		4-Sided Roll			
		1	2	3	4
6-Sided Roll	**1**	(1,1)	(1,2)	(1,3)	(1,4)
	2	(2,1)	(2,2)	(2,3)	(2,4)
	3	(3,1)	(3,2)	(3,3)	(3,4)
	4	(4,1)	(4,2)	(4,3)	(4,4)
	5	(5,1)	(5,2)	(5,3)	(5,4)
	6	(6,1)	(6,2)	(6,3)	(6,4)

So, for example, (3,2) represents the outcome where the 6-sided roll results in a 3, and the 4-sided roll gives us a 2. Thus, the sample space of the experiment is {(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4), (4,1), (4,2), (4,3), (4,4), (5,1), (5,2), (5,3), (5,4), (6,1), (6,2), (6,3), (6,4)}.

Step 1: Let's put the flavors on the rows and the toppings on the columns of the following table:

		Toppings		
		fudge	whipped cream	sprinkles
Flavors	**vanilla**			
	chocolate			
	rocky road			

Step 2: We can fill in the cells of the table below with the resulting combinations.

		Toppings		
		fudge	whipped cream	sprinkles
Flavors	**vanilla**	vanilla with fudge	vanilla with whipped cream	vanilla with sprinkles
	chocolate	chocolate with fudge	chocolate with whipped cream	chocolate with sprinkles
	rocky road	rocky road with fudge	rocky road with whipped cream	rocky road with sprinkles

So, the sample space is {vanilla with fudge, vanilla with whipped cream, vanilla with sprinkles, chocolate with fudge, chocolate with whipped cream, chocolate with sprinkles, rocky road with fudge, rocky road with whipped cream, rocky road with sprinkles}.

Step 1: We'll put the parents' genes (P1 and P2) as labels on the rows and columns of the following table:

		P2	
		g	g
P1	**G**		
	g		

Step 2: We'll fill in the offspring's gene composition, listing parent 1's gene first in the table below.

		P2	
		g	**g**
P1	**G**	Gg	Gg
	g	gg	gg

Thus, the sample space is {Gg, Gg, gg, gg}. (Diagrams like this, which allow us to identify the genotypes of offspring, are called *Punnett squares* in honor of Reginald Punnett (1875–1967), who first used them in the context of genetics.)

Your Turn 4.3.1.3

Use a table to identify the sample space of an experiment in which you flip a coin and roll a 6-sided die.

Using Tree Diagrams to Identify Sample Spaces

In experiments where there are more than two stages, or where the stages are dependent, a tree diagram is a helpful tool for systematically identifying the sample space. Tree diagrams are built by first drawing a single point (or *node*), then from that node we draw one branch (a short line segment) for each outcome of the first stage. Each branch gets its own node at the other end (which we typically label with the corresponding outcome for that branch); from each of these, we draw another branch for each outcome of the second stage, assuming that the outcome of the first stage matches the branch we were on. If there are other stages, we can continue from there by continuing to add branches and nodes. This sounds really complicated, but it's easier to understand through an example.

Example 4.3.1.4: Using a Tree Diagram to Identify a Sample Space

Use a tree diagram to find the sample spaces of each of the following experiments:

1. You flip a coin 3 times, noting the outcome of each flip in order.
2. You flip a coin. If the result is heads, you roll a 4-sided die. If it's tails, you roll a 6-sided die.
3. You are planning to go on a hike with a group of friends. There are 3 trails to consider: Abel Trail, Borel Trail, and Condorcet Trail. One of your friends, Jess, requires a wheelchair; if she joins you, the group couldn't handle the rocky Condorcet Trail.

Answer

1. **Step 1:** Let's start by placing our first node (Figure 4.3.1.2).



Figure 4.3.1.2

Step 2: We'll add two branches, one for each outcome of the first coin flip, and label them (Figure 4.3.1.3).



Figure 4.3.1.3

Step 3: We're ready for stage two of the experiment: another coin flip. At each node, we add in branches that represent those outcomes (Figure ✔ 4.3.1.4).

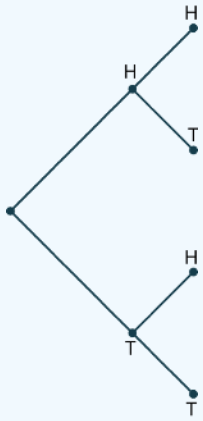


Figure ✔ 4.3.1.4

Finally, we can add another set of branches for the outcomes of the third stage (Figure ✔ 4.3.1.5).

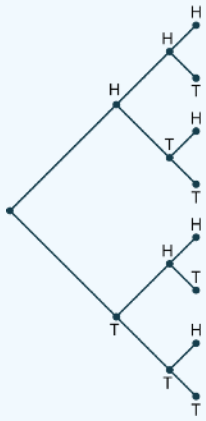


Figure ✔ 4.3.1.5

(These final nodes are called *leaves.)*

Step 4: We can write down the outcomes in the sample space by tracing the path out to each leaf, writing down the outcome at each node we pass through. For example, this leaf (Figure ✔ 4.3.1.6):

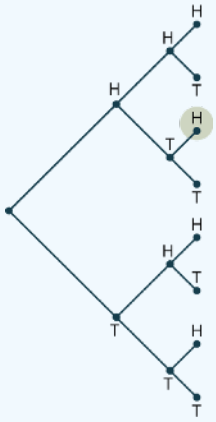


Figure ✔ 4.3.1.6

is reached via this path (Figure 7.16):

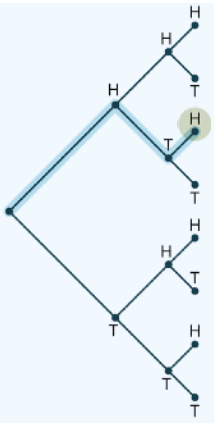


Figure ✔ 4.3.1.7

Step 5: We'll label that leaf as "HTH" (Figure ✔ 4.3.1.8), since the path passes through nodes labeled H, T, and H on its way out to our leaf.

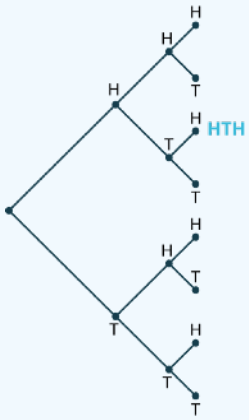


Figure ✔ 4.3.1.8

Step 6: We can label the remaining leaves using the same method (Figure ✔ 4.3.1.9).

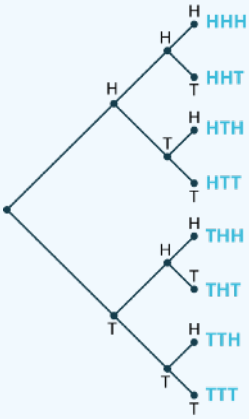


Figure ✔ 4.3.1.9

The sample space is the labels on the leaves: {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}.

Step 1: We'll start with our initial node (Figure ✔ 4.3.1.10).



Figure ✔ 4.3.1.10

Step 2: We'll add in branches for the outcomes of the first stage (Figure ✔ 4.3.1.11), which is the coin flip.



Figure ✔ 4.3.1.11

Step 3: The second stage of the experiment depends on the outcome of the first stage.

a. If the outcome of the first stage was H, then we roll a 4-sided die. So, *only* on the node for H, we'll add in the outcomes of a 4-sided die roll (Figure

✔ 4.3.1.12).

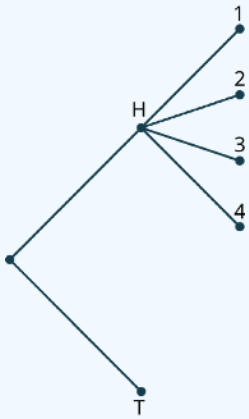


Figure ✔ 4.3.1.12

If the outcome of the first stage was T, then we roll a 6-sided die. So, we'll add those branches to the node for T (Figure ✔ 4.3.1.13).

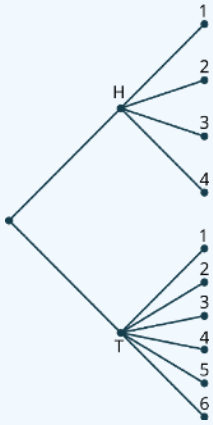


Figure ✔ 4.3.1.13

Step 4: We can label the leaves to get the sample space (Figure 4.3.1.14).

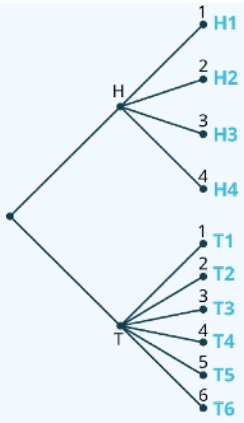


Figure ✔ 4.3.1.14

The sample space is: {H1, H2, H3, H4, T1, T2, T3, T4, T5, T6}.

c. Step 1: Let's label the trails A, B, and C for ease of labeling. Even though the trails are listed first in the exercise, we can't use the trail choice as our first stage: the trails available to us depend on whether Jess is able to join the trip. So, the first stage is whether Jess joins us (J) or not (N) (Figure ✔ 4.3.1.15).



Figure ✔ 4.3.1.15

Step 2: We list the appropriate trails on each branch (Figure ✔ 4.3.1.16).

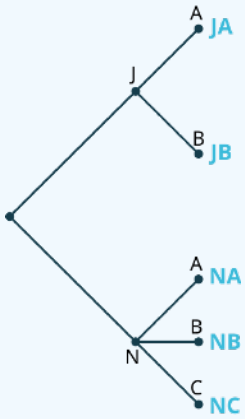


Figure ✔ 4.3.1.16

So, the sample space is {JA, JB, NA, NB, NC}

Your Turn 4.3.1.4

You have a modified deck of cards containing only J♡, Q♡, and K♡. You draw 2 cards without replacing them (where order matters). Use a tree diagram to identify the sample space.

Check Your Understanding

1. You flip a coin 6 times and note the number of heads. What is the sample space of this experiment?
2. You are ordering a combo meal at a restaurant. The meal comes with either 8 or 12 chicken nuggets, and your choice of crinkle fries, curly fries, or onion rings. Create a table to help you identify the sample space containing your combo meal possibilities.
3. You need one more class to fill out your schedule for next semester. You want to take either History 101 (H), English 220 (E), or Sociology 112 (S). There are two professors teaching the history class: Anderson (A) and Burr (B); one professor teaching the English class: Carter (C); and three people teaching sociology: Johnson (J), Kirk (K), and Lambert (L). Create a tree diagram that helps you identify all your options.
4. Identify the sample space from Exercise 18.
5. Identify the sample space from Exercise 19.

This page titled 4.3.1: Tree Diagrams, Tables, and Outcomes is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **7.4: Tree Diagrams, Tables, and Outcomes** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

4.3.2: The Multiplication Rule for Counting

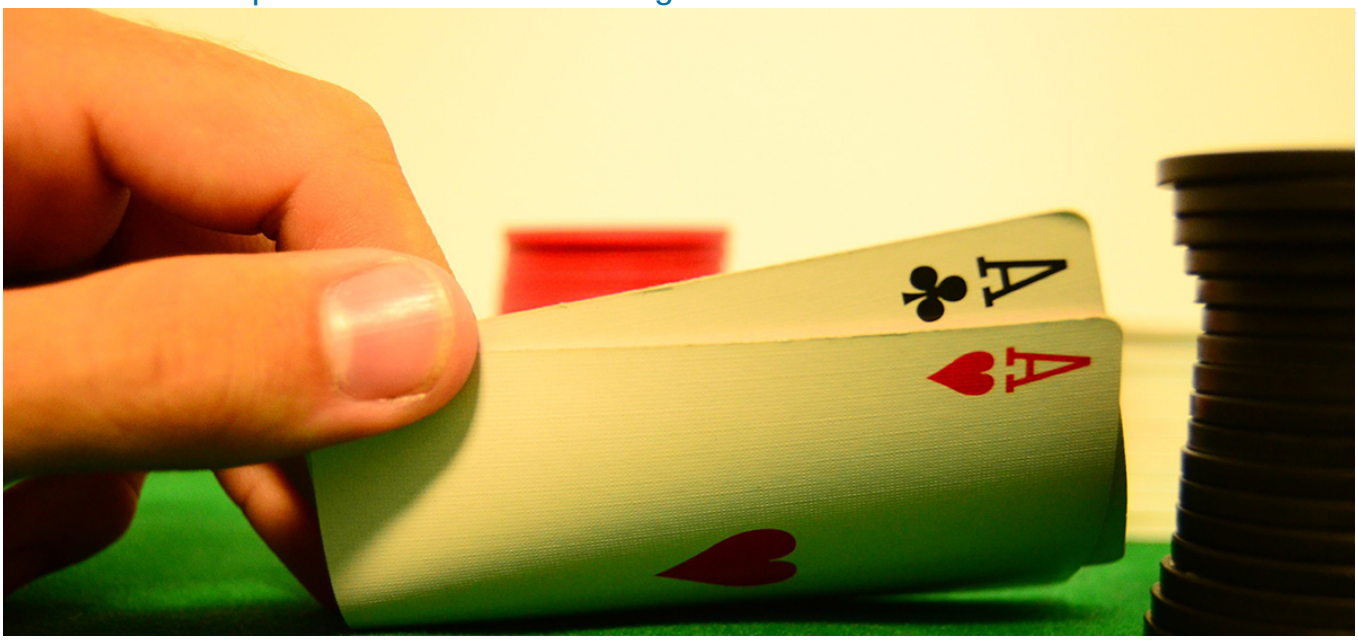


Figure 4.3.2.1 : The Multiplication Rule for Counting allows us to compute more complicated probabilities, like drawing two aces from a deck. (credit: “Pair of Aces – Poker” by Poker Photos/Flickr, CC BY 2.0)

Learning Objectives

Apply the Multiplication Rule for Counting to solve problems.

One of the first bits of mathematical knowledge children learn is how to count objects by pointing to them in turn and saying: “one, two, three, …” That’s a useful skill, but when the number of things that we need to count grows large, that method becomes onerous (or, for *very* large numbers, impossible for humans to accomplish in a typical human lifespan). So, mathematicians have developed short cuts to counting big numbers. These techniques fall under the mathematical discipline of combinatorics, which is devoted to counting.

Multiplication as a Combinatorial Short Cut

One of the first combinatorial short cuts to counting students learn in school has to do with areas of rectangles. If we have a set of objects to be counted that can be physically arranged into a rectangular shape, then we can use multiplication to do the counting for us. Consider this set of objects (Figure 4.3.2.2):



Figure 4.3.2.2

Certainly we can count them by pointing and running through the numbers, but it’s more efficient to group them (Figure 4.3.2.3).

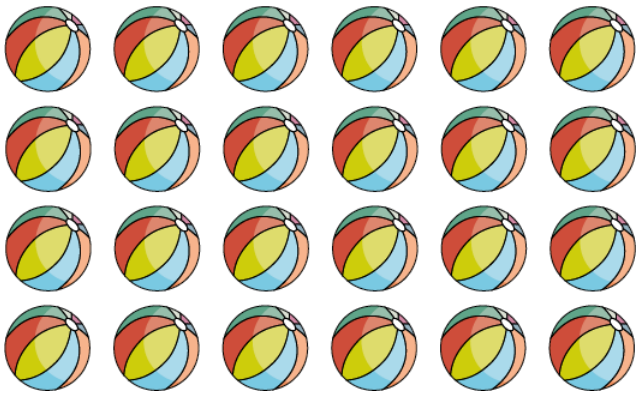


Figure 4.3.2.3

If we group the balls by 4s, we see that we have 6 groups (or, we can see this arrangement as 4 groups of 6 balls). Since multiplication is repeated addition (i.e., $6 \times 4 = 4+4+4+4+4+4$), we can use this grouping to quickly see that there are 24 balls.

Let's generalize this idea a little bit. Let's say that we're visiting a bakery that offers customized cupcakes. For the cake, we have three choices: vanilla, chocolate, and strawberry. Each cupcake can be topped with one of four types of frosting: vanilla, chocolate, lemon, and strawberry. How many different cupcake combinations are possible? We can think of laying out all the possibilities in a grid, with cake choices defining the rows and frosting choices defining the columns (Figure 4.3.2.4).

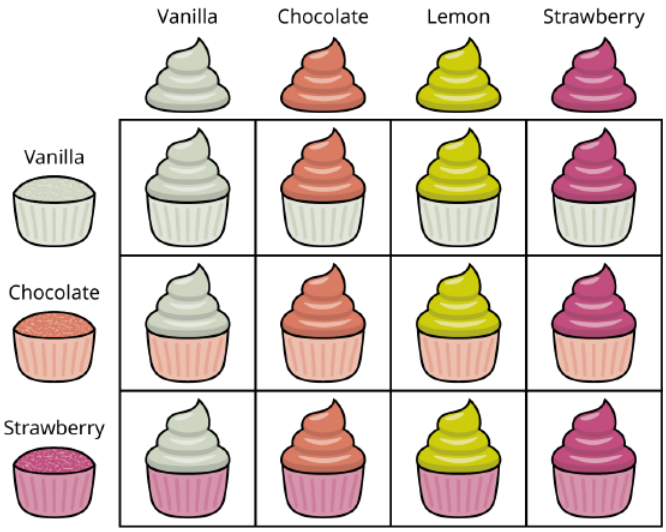


Figure 4.3.2.4

Since there are 3 rows (cakes) and 4 columns (frostings), we have $3 \times 4 = 12$ possible combinations. This is the reasoning behind the Multiplication Rule for Counting, which is also known as the Fundamental Counting Principle. This rule says that if there are n ways to accomplish one task and m ways to accomplish a second task, then there are $n \times m$ ways to accomplish both tasks. We can tack on additional tasks by multiplying the number of ways to accomplish those tasks to our previous product.

✔ Example ✔ 4.3.2.1: Using the Multiplication Rule for Counting

Every card in a standard deck of cards has two identifying characteristics: a suit (clubs, diamonds, hearts, or spades; these are indicated by these symbols, respectively: ♣, ♢, ♡, ♠) and a rank (ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, jack, queen, and king; the letters A, J, Q, and K are used to represent the words). Each possible pair of suit and rank appears exactly once in the deck. How many cards are in the standard deck?

Answer

Since there are 4 suits and 13 ranks, the number of cards must be $4 \times 13 =$

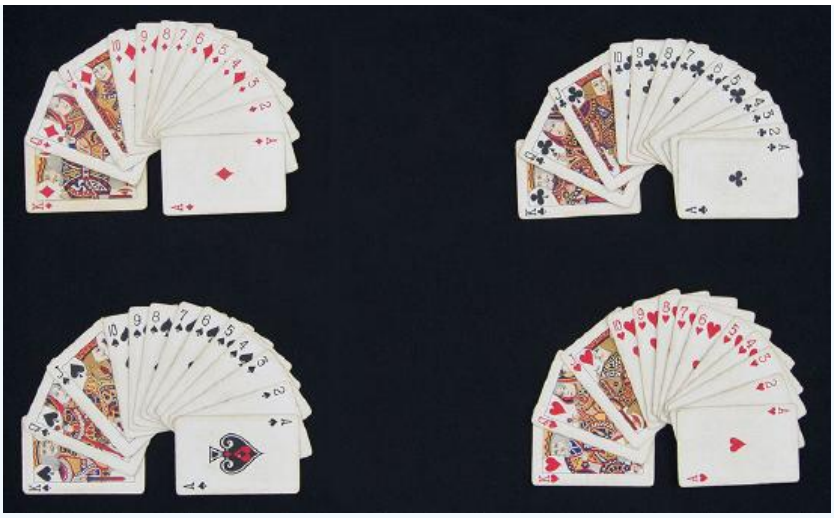


Figure 4.3.2.5 Standard Deck of Cards, Sorted by Rank and Suit (credit: "Playing Cards, USS Arkansas" by Naval History & Heritage Command/Flickr, CC BY 2.0)

Your Turn 4.3.2.1

Joe's Pizza Shack offers pizzas with 4 different types of crust and a choice of 15 toppings. How many different one-topping pizzas can be made at Joe's?

Example 4.3.2.2: Using the Multiplication Rule for Counting for 4 Groups

The University Combinatorics Club has 31 members: 8 seniors, 7 juniors, 5 sophomores, and 11 first-years. How many possible 4-person committees can be formed by selecting 1 member from each class?

Answer

Since we have 8 choices for the senior, 7 choices for the junior, 5 for the sophomore, and 11 for the first-year, there are $8 \times 7 \times 5 \times 11 = 3{,}080$ different ways to fill out the committee.

Your Turn 4.3.2.2

The menu for Joe's Pizza Shack offers pizzas with 4 different types of crust and a choice of 15 toppings. Suppose that Joe's also offers a choice of 3 sauces and 2 cheese blends. How many different one-topping pizzas can be made at Joe's now?

Example 4.3.2.3: Using the Multiplication Rule for Counting for More Groups

The standard license plates for vehicles in a certain state consist of 6 characters: 3 letters followed by 3 digits. There are 26 letters in the alphabet and 10 digits (0 through 9) to choose from. How many license plates can be made using this format?

Answer

Since there are 26 different letters and 10 different digits, the total number of possible license plates is $26 \times 26 \times 26 \times 10 \times 10 \times 10 = 17,576,000$.

Your Turn 4.3.2.3

At a certain college, ID cards are issued to all students, faculty, and staff. These cards have unique ID codes for each person: a letter to indicate the person's status (S for students, F for faculty, and E for staff), followed by 5 digits and finally 3 letters (these letters can be anything). How many different ID codes can be created using this scheme?

Check Your Understanding

1. A website that lets you build custom belts has 18 different buckles and 30 different straps. How many different belts can be made using those materials?
2. A chain of chicken restaurants offers a combo that includes your choice of 3 or 5 chicken strips, along with your choice of side dish. If there are 7 side dishes, how many different ways are there to build this combo meal?
3. When you flip a coin, there are 2 possible outcomes: heads and tails. Let's say you flip a coin 10 times, and after each you write down the result of the flip (H for heads, T for tails). How many different results (strings of 10 characters, where each is either an H or a T) are possible?
4. A T-shirt company allows shoppers to customize their shirts in several ways. There are 5 sizes, 8 shirt colors, 4 designs, and 5 design colors. How many different shirts can be made?
5. Josephine is trying to build her class schedule for next semester. Because of her work schedule, she has only 4 class periods that can work for her, and she must take 4 classes. If there are 15 classes that she could take during the first period, 18 during the second, 12 during the third, and 8 during the fourth, how many different schedules could Josephine build?

This page titled 4.3.2: The Multiplication Rule for Counting is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **7.1: The Multiplication Rule for Counting** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

4.4: Permutations and Combinations

4.4: Permutations and Combinations is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

4.4.1: Permutations

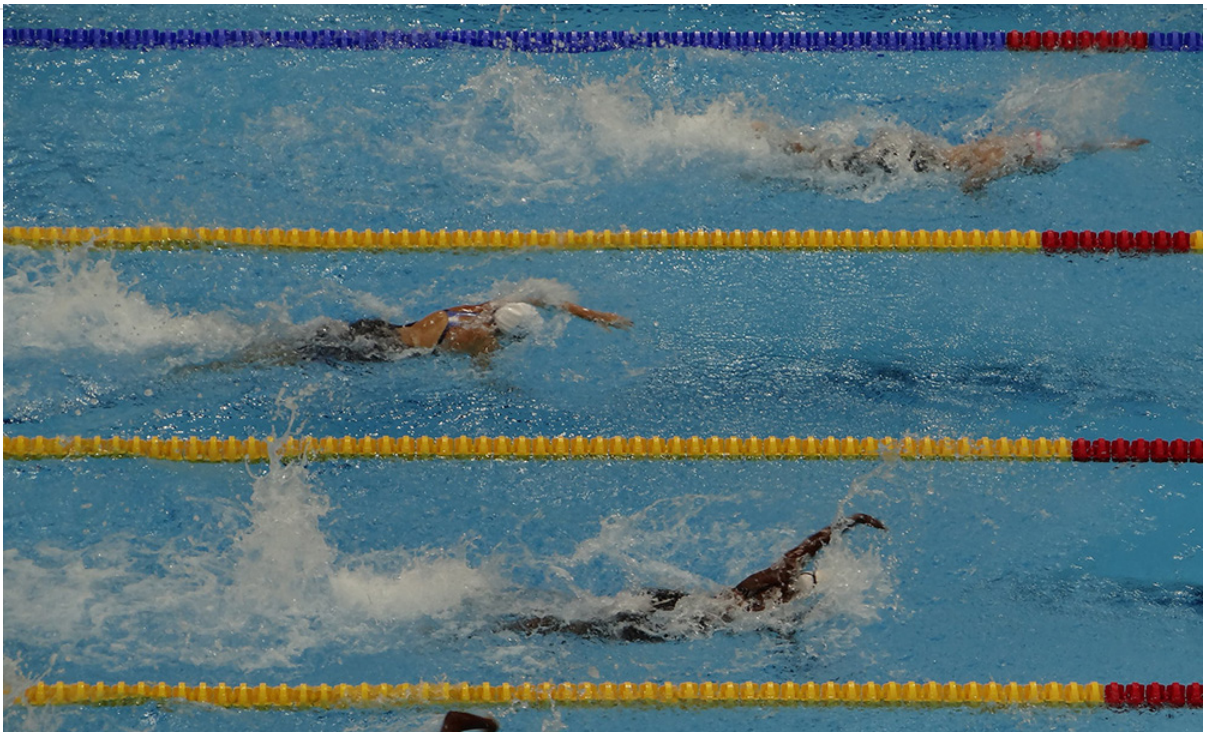


Figure 4.4.1.1: We can use permutations to calculate the number of different orders of finish in an Olympic swimming heat. (credit: “London 2012 Olympics Park Stratford London” by Gary Bembridge/Flickr, CC BY 2.0)

Learning Objectives

1. Use the Multiplication Rule for Counting to determine the number of permutations.
2. Compute expressions containing factorials.
3. Compute permutations.
4. Apply permutations to solve problems.

Swimming events are some of the most popular events at the summer Olympic Games. In the finals of each event, 8 swimmers compete at the same time, making for some exciting finishes. How many different orders of finish are possible in these events? In this section, we’ll extend the Multiplication Rule for Counting to help answer questions like this one, which relate to permutations. A permutation is an ordered list of objects taken from a given population. The length of the list is given, and the list cannot contain any repeated items.

Applying the Multiplication Rule for Counting to Permutations

In the case of the swimming finals, one possible permutation of length 3 would be the list of medal winners (first, second, and third place finishers). A permutation of length 8 would be the full order of finish (first place through eighth place). Let’s use the Multiplication Rule for Counting to figure out how many of each of these permutations there are.

Example

4.4.1.1: Using the Multiplication Rule for Counting to Find the Number of Permutations

The final heat of Olympic swimming events features 8 swimmers (or teams of swimmers).

1. How many different podium placements (first place, second place, and third place) are possible?
2. How many different complete orders of finish (first place through eighth place) are possible?

Answer

1. Let's start with the first place finisher. How many options are there? Since 8 swimmers are competing, there are 8 possibilities. Once that first swimmer completes the race, there are 7 swimmers left competing for second place. After the second finisher is decided, there are 6 swimmers remaining who could possibly finish in third place. Thus, there are 8 possibilities for first place, 7 for second place, and 6 for third place. The Multiplication Rule for Counting then tells us there are $8 \times 7 \times 6 = 336$ different ways the winners' podium can be filled out.
2. To look at the complete order of finish, we can continue the pattern we can see in part 1 of this example: There are 5 possibilities for fourth place, 4 for fifth place, 3 for sixth place, 2 for seventh place, and then just 1 swimmer is left to finish in eighth place. Using the Multiplication Rule for Counting, we see that there are $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 40,320$ possible orders of finish.

Your Turn 4.4.1.1

You have a hand of 5 cards (that happen to create what's called a royal flush in the game of poker): 10♠, J♠, Q♠, K♠, and A♠. Into how many different orders can you put those cards?

Factorials

The pattern we see in Example 7.4 occurs commonly enough that we have a name for it: factorial.
For any positive whole number n, we define the factorial of n (denoted $n!$ and read "n factorial") to be the product of every whole number less than or equal to n. We also define 0! to be equal to one. We will use factorials in a couple of different contexts, so let's get some practice doing computations with them.

Example 4.4.1.2: Computing Factorials

Compute the following:

1. $4!$
2. $\frac{8!}{6!}$
3. $\frac{9!}{3!4!}$

Answer

1. $4! = 4 \times 3 \times 2 \times 1 = 24$
2. There are two ways to approach this calculation. The first way is to compute the factorials first, then divide:

$$\frac{8!}{6!} = \frac{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{40,320}{720} = 56$$

However, there is an easier way! You may notice in the second step that there are several terms that can be canceled; that's always the case whenever we divide factorials. In this case, notice that we can rewrite the numerator like this:

$$8! = 8 \times 7 \times (6 \times 5 \times 4 \times 3 \times 2 \times 1) = 8 \times 7 \times 6!$$

With that in mind, we can proceed this way by canceling out the 6!:

$$\frac{8!}{6!} = \frac{8 \times 7 \times \mathbf{6!}}{\mathbf{6!}} = 8 \times 7 = 56$$

That's much easier!

3. Let's approach this one using our canceling technique. When we see two factorials in either the numerator or denominator, we should focus on the larger one first. So:

$$\frac{9!}{3!4!} = \frac{9 \times 8 \times 7 \times 6 \times 5 \times \mathbf{4!}}{3!\mathbf{4!}} = \frac{9 \times 8 \times 7 \times \mathbf{6} \times 5}{\mathbf{3} \times \mathbf{2} \times \mathbf{1}} = 9 \times 8 \times 7 \times 5 = 2,520$$

Your Turn 4.4.1.2

Compute the following:

$6!$

$\frac{12!}{10!}$

$\frac{8!}{4!4!}$

Permutations

As we've seen, factorials can pop up when we're computing permutations. In fact, there is a formula that we can use to make that connection explicit. Let's define some notation first. If we have a collection of n objects and we wish to create an ordered list of r of the objects (where $1 \leq r \leq n$), we'll call the number of those permutations (read "the number of permutations of n objects taken r at a time"). We formalize the formula we'll use to compute permutations below.

FORMULA

$${}_nP_r = \frac{n!}{(n-r)!}$$

If you wondered why we defined $0! = 1$ earlier, it was to make formulas like this one work; if we have n objects and want to order all of them (so, we want the number of permutations of n objects taken n at a time), we get ${}_nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = \frac{n!}{1} = n!$. Next, we'll get some practice computing these permutations.

Example 4.4.1.3: Computing Permutations

Find the following numbers:

1. The number of permutations of 12 objects taken 3 at a time
2. The number of permutations of 8 objects taken 5 at a time
3. The number of permutations of 32 objects taken 2 at a time

Answer

1. $$12P_3 = \frac{12!}{(12-3)!} = \frac{12!}{9!} = \frac{12 \times 11 \times 10 \times \mathbf{9!}}{\mathbf{9!}} = 12 \times 11 \times 10 = 1,$$
2. $$8P_5 = \frac{8!}{(8-5)!} = \frac{8 \times 7 \times 6 \times 5 \times 4 \times \mathbf{3!}}{\mathbf{3!}} = 8 \times 7 \times 6 \times 5 \times 4 = 6,$$
3. $$32P_2 = \frac{32!}{(32-2)!} = \frac{32 \times 31 \times \mathbf{30!}}{\mathbf{30!}} = 32 \times 31 =$$

Your Turn 4.4.1.3

Find the following numbers:

The number of permutations of 6 objects taken 2 at a time

The number of permutations of 14 objects taken 4 at a time

The number of permutations of 19 objects taken 3 at a time

Example 4.4.1.4: Applying Permutations

1. A high school graduating class has 312 students. The top student is declared valedictorian, and the second-best is named salutatorian. How many possible outcomes are there for the valedictorian and salutatorian?
2. In the card game blackjack, the dealer's hand of 2 cards is dealt with 1 card faceup and 1 card facedown. If the game is being played with a single deck of (52) cards, how many possible hands could the dealer get?
3. The University Combinatorics Club has 3 officers: president, vice president, and treasurer. If there are 18 members of the club, how many ways are there to fill the officer positions?

Answer

1. This is the number of permutations of 312 students taken 2 at a time, and $312P_2 = 97,$.
2. We want the number of permutations of 52 cards taken 2 at a time, and $52P_2 = 2,$.
3. Here we're looking for the number of permutations of 18 members taken 3 at a time, and $18P_3 = 4,$.

Your Turn 4.4.1.4

One of the big draws at this year's state fair is the pig race. There are 15 entrants, and prizes are given to the top three finishers. How many different combinations of top-three finishes could there be?

Who Knew?: Very Big Permutations

Permutations involving relatively small sets of objects can get very big, very quickly. A standard deck contains 52 cards. So, the number of different ways to shuffle the cards—in other words, the number of permutations of 52 objects taken 52 at a time—is $52! \approx 8 \times 10^{67}$ (written out, that's an 8 followed by 67 zeroes). The estimated age of the universe is *only* about 4×10^{17} seconds. So, if a very bored all-powerful being started shuffling cards at the instant the universe began, it would have to have averaged at least $\frac{8 \times 10^{67}}{4 \times 10^{17}} \approx 2 \times 10^{50}$ shuffles *per second since the beginning of time* to have covered every possible arrangement of a deck of cards. That means the next time you pick up a deck of cards and give it a good shuffle, it's almost certain that the particular arrangement you created has never been created before and likely never will be created again.

Check Your Understanding

1. Compute $5!$.
2. Compute $\frac{10!}{7!3!}$.
3. Compute ${}_{12}P_3$.
4. Compute ${}_8P_4$.
5. The standard American edition of the board game *Monopoly* has a deck of 15 orange Chance cards. In how many different ways could the first 4 Chance cards drawn in a game appear?

This page titled 4.4.1: Permutations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **7.2: Permutations** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

4.4.2: Combinations



Figure 4.4.2.1: Combinations help us count things like the number of possible card hands, when the order in which the cards were drawn doesn't matter. (credit: "IMG_3177" by Zanaca/Flickr, CC BY 2.0)

Learning Objectives

1. Distinguish between permutation and combination uses.
2. Compute combinations.
3. Apply combinations to solve applications.

In Permutations, we studied permutations, which we use to count the number of ways to generate an ordered list of a given length from a group of objects. An important property of permutations is that the order of the list matters: The results of a race and the selection of club officers are examples of lists where the order is important. In other situations, the order is not important. For example, in most card games where a player receives a hand of cards, the order in which the cards are received is irrelevant; in fact, players often rearrange the cards in a way that helps them keep the cards organized.

Combinations: When Order Doesn't Matter

In situations in which the order of a list of objects doesn't matter, the lists are no longer permutations. Instead, we call them combinations.

Example 4.4.2.1: Distinguishing Between Permutations and Combinations

For each of the following situations, decide whether the chosen subset is a permutation or a combination.

1. A social club selects 3 members to form a committee. Each of the members has an equal share of responsibility.
2. You are prompted to reset your email password; you select a password consisting of 10 characters without repeats.
3. At a dog show, the judge must choose first-, second-, and third-place finishers from a group of 16 dogs.
4. At a restaurant, the special of the day comes with the customer's choice of 3 sides taken from a list of 6 possibilities.

Answer

1. Since there is no distinction among the responsibilities of the 3 committee members, the order isn't important. So, this is a combination.
2. The order of the characters in a password matter, so this is a permutation.
3. The order of finish matters in a dog show, so this is a permutation.
4. A plate with mashed potatoes, peas, and broccoli is functionally the same as a plate with peas, broccoli, and mashed potatoes, so this is a combination.

Your Turn 4.4.2.1

Decide whether the following represent permutations or combinations:

On Halloween, you give each kid who comes to your door 3 pieces of candy, taken randomly from a candy dish.

Your class is going on a field trip, but there are too many people for one vehicle. Your instructor chooses half the class to take the first vehicle.

\(\frac{12!}{{10!}\) \(\frac{8!}{{4!4!}\)

Counting Combinations

Permutations and combinations are certainly related, because they both involve choosing a subset of a large group. Let's explore that connection, so that we can figure out how to use what we know about permutations to help us count combinations. We'll take a basic example. How many ways can we select 3 letters from the group A, B, C, D, and E? If order matters, that number is $5P_3 =$. That's small enough that we can list them all out in the table below.

ABC	ABD	ABE	ACB	ACD	ACE
ADB	ADC	ADE	AEB	AEC	AED
BAC	BAD	BAE	BCA	BCD	BCE
BDA	BDC	BDE	BEA	BEC	BED
CAB	CAD	CAE	CBA	CBD	CBE
CDA	CDB	CDE	CEA	CEB	CED
DCA	DAC	DAE	DBA	DBC	DBE
DCA	DCB	DCE	DEA	DEB	DEC
EAB	EAC	EAD	EBA	EBC	EBD
ECA	ECB	ECD	EDA	EDB	EDC

Now, let's look back at that list and color-code it so that groupings of the same 3 letters get the same color, as shown in Figure 7.9:

ABC	ABD	ABE	ACB	ACD	ACE
ADB	ADC	ADE	AEB	AEC	AED
BAC	BAD	BAE	BCA	BCD	BCE
BDA	BDC	BDE	BEA	BEC	BED
CAB	CAD	CAE	CBA	CBD	CBE
CDA	CDB	CDE	CEA	CEB	CED
DAB	DAC	DAE	DBA	DBC	DBE
DCA	DCB	DCE	DEA	DEB	DEC
EAB	EAC	EAD	EBA	EBC	EBD
ECA	ECB	ECD	EDA	EDB	EDC

Figure 4.4.2.2

After color-coding, we see that the 60 cells can be seen as 10 groups (colors) of 6. That's no coincidence! We've already seen how to compute the number of permutations using the formula To compute the number of combinations, let's count them another way using the Multiplication Rule for Counting. We'll do this in two steps:

Step 1: Choose 3 letters (paying no attention to order).

Step 2: Put those letters in order.

The number of ways to choose 3 letters from this group of 5 (A, B, C, D, E) is the number of combinations we're looking for; let's call that number (read "the number of combinations of 5 objects taken 3 at a time"). We can see from our chart that this is ten (the number of colors used). We can generalize our findings this way: remember that the number of permutations of n things taken r at a time is $nP_r = \frac{n!}{(n-r)!}$. That number is *also* equal to $nC_r \times r!$, and so it must be the case that $\frac{n!}{(n-r)!} = nC_r \times r!$. Dividing both sides of that equation by $r!$ gives us the formula below.

FORMULA

$${}_nC_r = \frac{n!}{r!(n-r)!}$$

nCr=n!r!(n−r)!

Example 4.4.2.2: Using the Combination Formula

Compute the following:

1. ${}_8C_3$
2. ${}_{12}C_5$
3. ${}_{15}C_9$

Answer

1. ${}_8C_3 = \frac{8!}{3!(8-3)!} = \frac{8\times7\times\mathbf{6}\times\mathbf{5!}}{\mathbf{3}\times\mathbf{2}\times\mathbf{1}\times\mathbf{5!}} = 8\times7 = 56$
2. ${}_{12}C_5 = \frac{12!}{5!(12-5)!} = \frac{12\times11\times10\times9\times8\times7!}{5\times4\times3\times2\times1\times7!} = 11\times9\times8 = 792$

3. ${}_{15}C_9 = \frac{15!}{9!(15-9)!} = \frac{\mathbf{15}\times14\times13\times\mathbf{12}\times11\times10\times\mathbf{9!}}{\mathbf{9!}\times\mathbf{6}\times\mathbf{5}\times4\times\mathbf{3}\times\mathbf{2}\times1} = \frac{14\times13\times11\times10}{4} = 5,005$

Your Turn 4.4.2.2

Compute the following:

${}_6C_4$

${}_{10}C_8$

${}_{14}C_5$

Example 4.4.2.3: Applying the Combination Formula

1. In the card game Texas Hold'em (a variation of poker), players are dealt 2 cards from a standard deck to form their hands. How many different hands are possible?
2. The board game *Clue* uses a deck of 21 cards. If 3 people are playing, each person gets 6 cards for their hand. How many different 6-card *Clue* hands are possible?
3. Palmetto Cash 5 is a game offered by the South Carolina Education Lottery. Players choose 5 numbers from the whole numbers between 1 and 38 (inclusive); the player wins the jackpot of $100,000 if the randomizer selects those numbers in any order. How many different sets of winning numbers are possible?

Answer

1. A standard deck has 52 cards, and a hand has 2 cards. Since the order doesn't matter, we use the formula for counting combinations:

$$52C_2 = \frac{52!}{2!(52-2)!} = \frac{52\times51\times\mathbf{50!}}{2\times1\times\mathbf{50!}} = \frac{52\times51}{2} = 1{,}326.$$

2. Again, the order doesn't matter, so the number of combinations is:

$$21C_6 = \frac{21!}{6!(21-6)!} = \frac{21\times\mathbf{20}\times19\times\mathbf{18}\times17\times16\times\mathbf{15!}}{\mathbf{6}\times\mathbf{5}\times\mathbf{4}\times\mathbf{3}\times2\times1\times\mathbf{15!}} = \frac{21\times19\times17\times16}{2} = 54{,}264.$$

3. There are 38 numbers to choose from, and we must pick 5. Since order doesn't matter, the number of combinations is:

$$38C_5 = \frac{38!}{5!(38-5)!} = \frac{38\times37\times36\times35\times34\times\mathbf{33!}}{5\times4\times3\times2\times1\times\mathbf{33!}} = 501{,}492.$$

Your Turn 4.4.2.3

At a charity event with 58 people in attendance, 3 raffle winners are chosen. All receive the same prize, so order doesn't matter. How many different groups of 3 winners can be chosen?

A sorority with 42 members needs to choose a committee with 4 members, each with equal responsibility. How many committees are possible?

Checkpoint

The notation and nomenclature used for the number of combinations is not standard across all sources. You'll sometimes see $\binom{n}{r}$ *instead of . Sometimes you'll hear that expression read as "*n *choose* r*" as shorthand for "the number of combinations of* n *objects taken* r *at a time."*

People in Mathematics: Early Eastern Mathematicians

Although combinations weren't really studied in Europe until around the 13th century, mathematicians of the Middle and Far East had already been working on them for hundreds of years. The Indian mathematician known as Pingala had described them by the second century BCE; Varāhamihira (fl. sixth century) and Halayudha (fl. 10th century) extended Pingala's work. In the ninth century, a Jain mathematician named Mahāvīra gave the formula for combinations that we use today.

In 10th-century Baghdad, a mathematician named Al-Karaji also knew formulas for combinations; though his work is now lost, it was known to (and repeated by) Persian mathematician Omar Khayyam, whose work survives. Khayyam is probably best remembered as a poet, with his *Rubaiyat* being his most famous work.

Meanwhile, in 11th-century China, Jia Xian also was working with combinations, as was his 13th-century successor Yang Hui.

It is not known whether the discoveries of any of these men were known in the other regions, or if the Indians, Persians, and Chinese all came to their discoveries independently. We do know that mathematical knowledge and sometimes texts did get passed along trade routes, so it can't be ruled out.

Example 4.4.2.4: Combining Combinations with the Multiplication Rule for Counting

The student government at a university consists of 10 seniors, 8 juniors, 6 sophomores, and 4 first-years.

1. How many ways are there to choose a committee of 8 people from this group?
2. How many ways are to choose a committee of 8 people if the committee must consist of 2 people from each class?

Answer

1. There are 28 people to choose from, and we need 8. So, the number of possible committees is $28C_8 = 3{,}108$.
2. Break the selection of the committee members down into a 4-step process: Choose the seniors, then choose the juniors, then the sophomores, and then the first-years, as shown in the table below:

Class	Number of Ways to Choose Committee Representatives
senior	$10C_2 =$
junior	$8C_2 =$
sophomore	$6C_2 =$
first-year	$4C_2 =$

The Multiplication Rule for Counting tells us that we can get the total number of ways to complete this task by multiplying together the number of ways to do each of the four subtasks. So, there are $45 \times 28 \times 15 \times 6 = 113{,}400$ possible committees with these restrictions.

Your Turn 4.4.2.4

How many ways are there to choose a hand of 6 cards from a standard deck with the constraint that 3 are ♠, 2 are ♡, and 1 is ♣?

Check Your Understanding

1. Suppose you want to count the number of ways that you can arrange the apps on the home screen on your phone. Should you use permutations or combinations?
2. Your little brother is packing up for a family vacation, but there's only room for 3 of his toys. If you want to know how many possible groups of toys he can bring, should you use permutations or combinations?
3. Compute ${}_{12}C_{10}$.
4. Compute ${}_{16}C_3$.
5. You're planning a road trip with some friends. Though you have 6 friends you'd consider bringing along, you only have room for 3 other people in the car. How many different possibilities are there for your road trip squad?

6. You’re packing for a trip, for which you need 3 shirts and 3 skirts. If you have 8 shirts and 5 skirts that would work for the trip, how many different ways are there to pack for the trip?

This page titled 4.4.2: Combinations is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **7.3: Combinations** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

CHAPTER OVERVIEW

5: Probability

5.1: Sample Spaces and Probability

5.1.1: Basic Concepts

5.2: Probability Rules

5.2.1: The Addition Rule for Probability

5.2.2: Conditional Probability and the Multiplication Rule

5.3: Binomial Distributions

5.3.1: The Binomial Distribution

5.4: Markov Chains

5.4.1: Stochastic Matrices

5: Probability is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

5.1: Sample Spaces and Probability

5.1: Sample Spaces and Probability is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

5.1.1: Basic Concepts

If you roll a die, pick a card from deck of playing cards, or randomly select a person and observe their hair color, we are executing an experiment or procedure. In probability, we look at the likelihood of different outcomes. We begin with some terminology.

Events and Outcomes

The result of an experiment is called an **outcome**.

An **event** is any particular outcome or group of outcomes.

A **simple event** is an event that cannot be broken down further

The **sample space** is the set of all possible simple events.

Example 1

If we roll a standard 6-sided die, describe the sample space and some simple events.

The sample space is the set of all possible simple events: $\{1, 2, 3, 4, 5, 6\}$

Some examples of simple events:

We roll a 1

We roll a 5

Some compound events:

We roll a number bigger than 4

We roll an even number



Two dice One die

Basic Probability

Given that all outcomes are equally likely, we can compute the probability of an event E using this formula:

$$P(E) = \frac{\text{Number of outcomes corresponding to the event E}}{\text{Total number of equally - likely outcomes}}$$

Example 2

If we roll a 6-sided die, calculate

a. P(rolling a 1)
b. P(rolling a number bigger than 4)

Solution

Recall that the sample space is $\{1, 2, 3, 4, 5, 6\}$

a. There is one outcome corresponding to “rolling a 1”, so the probability is $\frac{1}{6}$
b. There are two outcomes bigger than a 4, so the probability is $\frac{2}{6} = \frac{1}{3}$

Probabilities are essentially fractions, and can be reduced to lower terms like fractions.

Example 3

Let's say you have a bag with 20 cherries, 14 sweet and 6 sour. If you pick a cherry at random, what is the probability that it will be sweet?

Solution

There are 20 possible cherries that could be picked, so the number of possible outcomes is 20. Of these 20 possible outcomes, 14 are favorable (sweet), so the probability that the cherry will be sweet is $\frac{14}{20} = \frac{7}{10}$.

There is one potential complication to this example, however. It must be assumed that the probability of picking any of the cherries is the same as the probability of picking any other. This wouldn't be true if (let us imagine) the sweet cherries are smaller than the sour ones. (The sour cherries would come to hand more readily when you sampled from the bag.) Let us keep in mind, therefore, that when we assess probabilities in terms of the ratio of favorable to all potential cases, we rely heavily on the assumption of equal probability for all outcomes.

Try it Now 1

At some random moment, you look at your clock and note the minutes reading.

a. What is probability the minutes reading is 15?
b. What is the probability the minutes reading is 15 or less?

Answer

There are 60 possible readings, from 00 to 59.

a. $\frac{1}{60}$
b. $\frac{16}{60}$ (counting 00 through 15)

Cards

A standard deck of 52 playing cards consists of four **suits** (hearts, spades, diamonds and clubs). Spades and clubs are black while hearts and diamonds are red. Each suit contains 13 cards, each of a different **rank**: an Ace (which in many games functions as both a low card and a high card), cards numbered 2 through 10, a Jack, a Queen and a King.

Example 4

Compute the probability of randomly drawing one card from a deck and getting an Ace.

Solution

There are 52 cards in the deck and 4 Aces so

$$P(Ace) = \frac{4}{52} = \frac{1}{13} \approx 0.0769$$

We can also think of probabilities as percents: There is a 7.69% chance that a randomly selected card will be an Ace.

Notice that the smallest possible probability is 0 – if there are no outcomes that correspond with the event. The largest possible probability is 1 – if all possible outcomes correspond with the event.

Certain and Impossible events

An **impossible event** has a probability of 0.

A **certain event** has a probability of 1.

The probability of any event must be $0 \le P(E) \le 1$.

In the course of this chapter, *if you compute a probability and get an answer that is negative or greater than 1, you have made a mistake and should check your work.*

This page titled 5.1.1: Basic Concepts is shared under a CC BY-SA 3.0 license and was authored, remixed, and/or curated by David Lippman (The OpenTextBookStore) via source content that was edited to the style and standards of the LibreTexts platform.

- **12.2: Basic Concepts** by David Lippman is licensed CC BY-SA 3.0. Original source: http://www.opentextbookstore.com/mathinsociety.

5.2: Probability Rules

5.2: Probability Rules is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

5.2.1: The Addition Rule for Probability

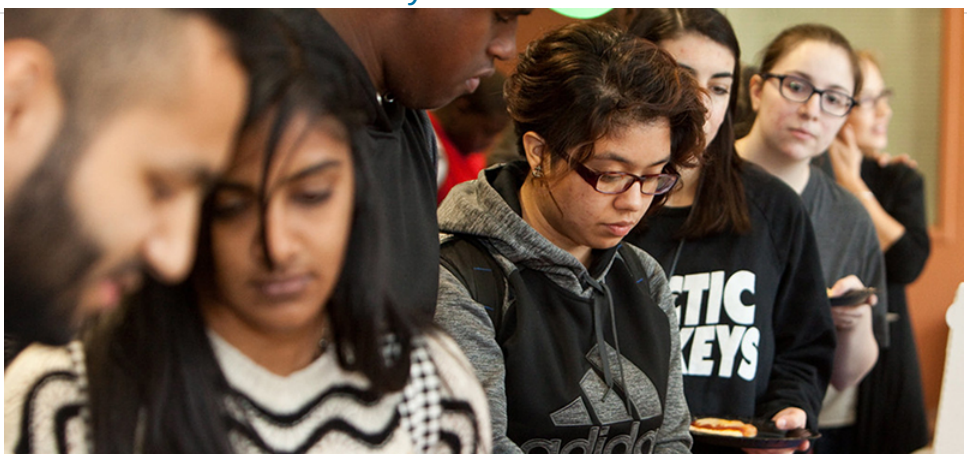


Figure 5.2.1.1: Students can be sorted using a variety of possible categories like class year, major, whether they are a varsity athlete, and so forth. (credit: "Multicultural Mashup Melds Languages, Cultures at COD 36" by COD Newsroom/Flickr, CC BY 2.0)

Learning Objectives

1. Identify mutually exclusive events.
2. Apply the Addition Rule to compute probability.
3. Use the Inclusion/Exclusion Principle to compute probability.

Up to this point, we have looked at the probabilities of *simple* events. Simple events are those with a single, simple characterization. Sometimes, though, we want to investigate more complicated situations. For example, if we are choosing a college student at random, we might want to find the probability that the chosen student is a varsity athlete *or* in a Greek organization. This is a compound event: there are two possible criteria that might be met. We might instead try to identify the probability that the chosen student is *both* a varsity athlete *and* in a Greek organization. In this section and the next, we'll cover probabilities of two types of compound events: those build using "or" and those built using "and." We'll deal with the former first.

Mutual Exclusivity

Before we get to the key techniques of this section, we must first introduce some new terminology. Let's say you're drawing a card from a standard deck. We'll consider 3 events: H is the event "the card is a ♡," T is the event "the card is a 10," and S is the event "the card is a ♠." If the card drawn is J♠, then H and T didn't occur, but S did. If the card drawn is instead 10♠, then H didn't occur, but both T and S did.

We can see from these examples that, if we are interested in several possible events, more than one of them can occur simultaneously (both T and S, for example). But, if you think about all the possible outcomes, you can see that H and S can *never* occur simultaneously; there are no cards in the deck that are both ♡ and ♠. Pairs of events that cannot both occur simultaneously are called mutually exclusive. Let's go through an example to help us better understand this concept.

Example 5.2.1.1: Identifying Mutually Exclusive Events

Decide whether the following events are mutually exclusive. If they are not mutually exclusive, identify an outcome that would result in both events occurring.

1. You are about to roll a standard 6-sided die. E is the event "the die shows an even number" and F is the event "the die shows an odd number."
2. You are about to roll a standard 6-sided die. E is the event "the die shows an even number" and S is the event "the die shows a number less than 4."
3. You are about to flip a coin 4 times. J is the event "at least 2 heads are flipped" and K is the event "fewer than 3 tails are flipped."

Answer

1. Let's look at the outcomes for each event: $E = \{2, 4, 6\}$ and $F = \{1, 3, 5\}$. There are no outcomes in common, so E and F are mutually exclusive.
2. Again, consider the outcomes in each event: $E = \{2, 4, 6\}$ and $S = \{1, 2, 3\}$. Since the outcome 2 belongs to both events, these are *not* mutually exclusive.
3. Suppose the results of the 4 flips are HTTH. Then at least 2 heads are flipped, and fewer than 3 tails are flipped. That means that both J and K occurred, and so these events are *not* mutually exclusive.

Your Turn 5.2.1.1

Suppose you're about to draw one card from a deck containing only these 10 cards: A♡, A♠, A♣, A♢, K♠, K♣, Q♡, Q♠, J♡, J♠. Decide whether these events are mutually exclusive:

E is the event "the card is an ace" and F is the event "the card is a king."

R is the event "the card is a ♡ " and E is the event "the card is an ace."

R is the event "the card is a ♡ " and F is the event "the card is a king."

The Addition Rule for Mutually Exclusive Events

If two events are mutually exclusive, then we can use addition to find the probability that one or the other event occurs.

FORMULA

If E and F are mutually exclusive events, then

$P(E \text{ or } F) = P(E) + P(F)$

Why does this formula work? Let's consider a basic example. Suppose we're about to draw a *Scrabble* tile from a bag containing A, A, B, E, E, E, R, S, S, U. What is the probability of drawing an E or an S? Since 3 of the tiles are marked with E and 2 are marked with S, there are 5 tiles that satisfy the criteria. There are ten tiles in the bag, so the probability is $\frac{5}{10} = \frac{1}{2}$. Notice that the probability of drawing an E is $\frac{3}{10}$ and the probability of drawing an S is $\frac{2}{10}$; adding those together, we get $\frac{3}{10} + \frac{2}{10} = \frac{5}{10}$. Look at the numerators in the fractions involved in the sum: the 3 represents the number of E tiles and the 2 is the number of S tiles. This is why the Addition Rule works: The total number of outcomes in one event or the other is the sum of the numbers of outcomes in each of the individual events.

Example 5.2.1.2: Using the Addition Rule

For each of the given pairs of events, decide if the Addition Rule applies. If it does, use the Addition Rule to find the probability that one or the other occurs.

1. You are rolling a standard 6-sided die. Event A is "roll an even number" and event B is "roll a 3."
2. You are drawing a card at random from a standard 52-card deck. Event R is "draw a ♡" and event S is "draw a king."
3. You are rolling a pair of standard 6-sided dice. Example 7.18 might help.

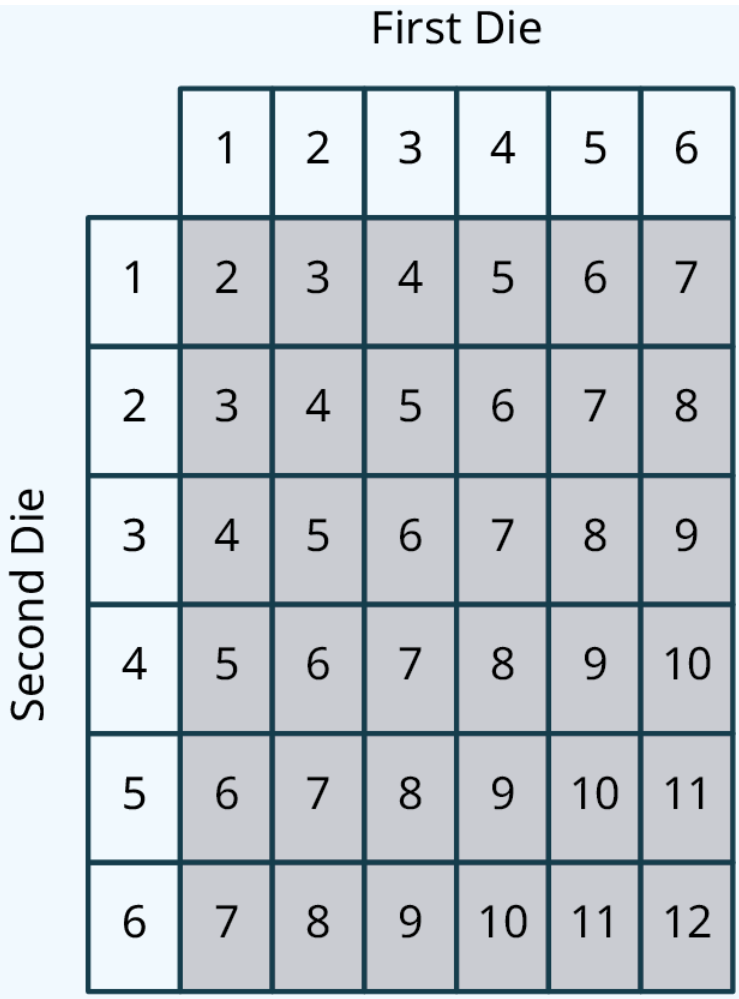


Figure ✔ 5.2.1.2

Answer

1. Since 3 is not an even number, these events are mutually exclusive. So, we can use the Addition Rule: since $P(A) =$ and $P(B) =$, we get $P(A \text{ or } B) = \frac{3}{6} + \frac{1}{6} =$.
2. If the card drawn is K♡, then both R and S occur. So, they aren't mutually exclusive, and the Addition Rule doesn't apply.
3. Since 10 is not odd, these events are mutually exclusive. Since $P(E) =$ and $P(F) =$, the Addition Rule gives us $P(E \text{ or } F) = \frac{18}{36} + \frac{3}{36} =$.

Your Turn \(\PageIndex{2\)

Suppose you're about to draw one card from a deck containing only these 10 cards:
A♡, A♠, A♣, AA♢, K♠, K♣, Q♡, Q♠, J♡, J♠. If appropriate, use the Addition Rule to find the probability that one or the other of these events occurs:

E is the event "the card is an ace" and F is the event "the card is a king."

R is the event "the card is a ♡ " and E is the event "the card is an ace."

R is the event "the card is a ♡ " and F is the event "the card is a king."

Finding Probabilities When Events Aren't Mutually Exclusive

Let's return to the example we used to explore the Addition Rule: We're about to draw a *Scrabble* tile from a bag containing A, A, B, E, E, E, R, S, S, U. Consider these events: J is "draw a vowel" and K is "draw a letter that comes after L in the alphabet." Since there are 6 vowels, $P(J) =$. There are 4 tiles with letters that come after L alphabetically, so $P(K) =$. What is $P(J \text{ or } K)$? If we blindly apply the Addition Rule, we get $\frac{6}{10} + \frac{4}{10} = 1$, which would mean that the compound event J or K is certain. However, it's possible to draw a B, in which case neither J nor K happens. Where's the error?

The events are not mutually exclusive: the outcome U belongs to both events, and so the Addition Rule doesn't apply. However, there's a way to extend the Addition Rule to allow us to find this probability anyway; it's called the Inclusion/Exclusion Principle. In this example, if we just add the two probabilities together, the outcome U is included in the sum twice: It's one of the 6 outcomes represented in the numerator of $\frac{6}{10}$, and it's one of the 4 outcomes represented in the numerator of $\frac{4}{10}$. So, that particular outcome has been "double counted." Since it has been *included* twice, we can get a true accounting by *excluding* it once: $\frac{6}{10} + \frac{4}{10} - \frac{1}{10} = \frac{9}{10}$. We can generalize this idea to a formula that we can apply to find the probability of any compound event built using "or."

FORMULA

Inclusion/Exclusion Principle: If E and F are events that contain outcomes of a single experiment, then

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

.

It's worth noting that this formula is truly an extension of the Addition Rule. Remember that the Addition Rule requires that the events E and F are mutually exclusive. In that case, the compound event $(E \text{ and } F)$ is impossible, and so $P(E \text{ and } F) = 0$. So, in cases where the events in question are mutually exclusive, the Inclusion/Exclusion Principle reduces to the Addition Rule.

Example 7.5.2.1.3: Using the Inclusion/Exclusion Principle

Suppose we have events E, F, and G, associated with these probabilities:

$$\begin{aligned} P(E) &= 0.45 \\ P(F) &= 0.6 \\ P(G) &= 0.55 \\ P(E \text{ and } F) &= 0.2 \\ P(E \text{ and } G) &= 0.2 \\ P(F \text{ and } G) &= 0.25 \end{aligned}$$

Compute the following:
1. $P(E \text{ or } F)$
2. $P(E \text{ or } G)$
3. $P(F \text{ or } G)$

Answer

1. Using the Inclusion/Exclusion Principle, we get:

$$\begin{aligned} P(E \text{ or } F) &= P(E) + P(F) - P(E \text{and } F) \\ &= 0.45 + 0.6 - 0.2 \\ &= 0.85. \end{aligned}$$

2. Again, we'll apply the Inclusion/Exclusion Principle:

$$\begin{aligned} P(E \text{or } G) &= P(E) + P(G) - P(E \text{and } G) \\ &= 0.45 + 0.55 - 0.2 \\ &= 0.8. \end{aligned}$$

3. Applying the Inclusion/Exclusion Principle one more time:

$$\begin{aligned} P(F \text{ or } G) &= P(F) + P(G) - P(F \text{and } G) \\ &= 0.6 + 0.55 - 0.25 \\ &= 0.9. \end{aligned}$$

Your Turn 5.2.1.3

You are about to roll a special 6-sided die that has both a colored letter and a colored number on each face. The faces are labeled with: a red 1 and a blue A, a red 1 and a green A, an orange 1 and a green B, an orange 2 and a red C, a purple 3 and a brown D, an orange 4 and a blue E. Find the probabilities of these events:

The number is orange or even.

The letter is green or an A.

The number is even or the letter is green.

Check Your Understanding

You are about to draw a card at random from a deck containing only these 10 cards: A♡, A♠, A♣, A♢, K♠, K♣, Q♡, Q♠, J♡, J♠. Compute the following probabilities:

1. You draw an ace or a king.
2. You draw a ♠ or a ♣.
3. You draw an ace or a ♡.
4. You draw a jack or a ♡.
5. You draw a jack or a ♣.
6. You draw a king or a ♢.

This page titled 5.2.1: The Addition Rule for Probability is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **7.8: The Addition Rule for Probability** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

5.2.2: Conditional Probability and the Multiplication Rule



Figure 5.2.2.1: If you roll two dice by throwing them one at a time, the face showing on the first die will affect the possible outcomes for the sum of the two dice. (credit: “dice” by Ciarán Archer/Flickr, CC BY 2.0)

Learning Objectives

After completing this section, you should be able to:

1. Calculate conditional probabilities.
2. Apply the Multiplication Rule for Probability to compute probabilities.

Back in Example 5.2.2.2, we constructed the following table (Figure 7.38) to help us find the probabilities associated with rolling two standard 6-sided dice:

Second Die \ First Die	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Figure 5.2.2.2

For example, 3 of these 36 equally likely outcomes correspond to rolling a sum of 10, so the probability of rolling a 10 is $\frac{3}{36} = \frac{1}{12}$. However, if you choose to roll the dice one at a time, the probability of rolling a 10 will change after the first die comes to rest. For example, if the first die shows a 5, then the probability of rolling a sum of 10 has jumped to $\frac{1}{6}$—the event will occur if the second die also shows a 5, which is 1 of 6 equally likely outcomes for the second die. If instead the first die shows a 3, then the probability of rolling a sum of 10 drops to 0—there are no outcomes for the second die that will give us a sum of 10.

Understanding how probabilities can shift as we learn new information is critical in the analysis of our second type of compound events: those built with "and." This section will explain how to compute probabilities of those compound events.

Conditional Probabilities

When we analyze experiments with multiple stages, we often update the probabilities of the possible final outcomes or the later stages of the experiment based on the results of one or more of the initial stages. These updated probabilities are called conditional probabilities.

In other words, if O is a possible outcome of the first stage in a multistage experiment, then the probability of an event E conditional on O (denoted $P(E|O)$, read "the probability of E given O") is the updated probability of E under the assumption that O occurred.

In the example that opened this section, we might consider rolling two dice as a multistage experiment: rolling one, then the other. If we define E to be the event "roll a sum of 10," O to be the event "first die shows 5," and Q to be the event "first die shows 3," then we computed $P(E) = \frac{1}{12}$, $P(E|O) = \frac{1}{6}$, and $P(E|Q) = 0$.

✔ Example ✔ 5.2.2.1: Computing Conditional Probabilities

1. April is playing a coin-flipping game with Ben. She will flip a coin 3 times. If the coin lands on heads more than tails, April wins; if it lands on tails more than heads, Ben wins. Let A be the event "April wins," H be "first flip is heads," and T be "first flip is tails." Compute $P(A)$, $P(A|H)$, and $P(A|T)$.
2. You are about to draw 2 cards without replacement from a deck containing only these 10 cards: A♡, A♠, A♣, A♢, K♠, K♣, Q♡, Q♠, J♡, J♠. We'll define the following events: F is "both cards are the same rank," A is "first card is an ace," and K is "first card is a king." Compute $P(F|A)$ and $P(F|K)$.
3. Jim's sock drawer contains 5 black socks and 3 blue socks. To avoid waking his partner, Jim doesn't want to turn the lights on, so he puts on 2 socks at random. Let M be the event "Jim's 2 socks match," let K be the event "the sock on Jim's left foot is black," and let L be the event "the sock on Jim's left foot is blue." Compute $P(M)$, $P(M|K)$, and $P(M|L)$.

Answer

1. **Step 1.** The sample space is {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}. The event A consists of the first 4 of those outcomes: HHH, HHT, HTH, and THH. Thus, $P(A) = \frac{4}{8} =$.

   Step 2. Now, let's compute $P(A|H)$. We are assuming the result of the first flip is heads. That leaves us with 4 possible outcomes: HHH, HHT, HTH, and HTT. Of those, April wins 3 (HHH, HHT, HTH) and loses one (HTT). So, $P(A|H) = \frac{3}{4}$.

   Step 3. If the result of the first flip is instead tails, the 4 possible outcomes are THH, THT, TTH, and TTT. Of those, April wins 1 (THH) and loses 3 (THT, TTH, TTT). So, $P(A|T) = \frac{1}{4}$.

2. **Step 1.** If the event A happens, then 1 of the 4 aces is drawn first; the remaining cards in the deck are 3 aces, 2 kings, 2 queens, and 2 jacks. In order for the event F to occur, the second card drawn has to be an ace. Since there are 3 aces among the remaining 9 cards, $P(F|A) = \frac{3}{9} = \frac{1}{3}$.

   Step 2. If the event K happens instead, then the first card drawn is a king. That leaves 4 aces, 1 king, 2 queens, and 2 jacks in the deck. Under the assumption that the first card is a king, the event F will occur only if the second card is also a king. Since only one of the remaining 9 cards is a king, we have $P(F|K) = \frac{1}{9}$.

3. **Step 1.** We can view the event M as a compound event using "or": both socks are blue or both socks are black. Let's compute the probability that both socks are blue using combinations. We're choosing 2 socks from a group of 8; 3 of the 8 are blue. So, $P(\text{both socks blue}) = \frac{}{{}_3C_2} =$. Similarly, $P(\text{both socks black}) = \frac{}{{}_5C_2} =$. Therefore, since these events are mutually exclusive, we can use the Addition Rule: $P(M) = P(\text{both socks blue}) + P(\text{both socks black}) = \frac{3}{28} + \frac{10}{28} =$.

   Step 2. If the sock on Jim's left foot is black (i.e., K occurred), then there are 4 remaining black socks of the 7 in the drawer. So, $P(M|K) = \frac{4}{7}$.

   Step 3. If the sock on Jim's left foot is blue (L occurred), then there are 2 blue socks among the 7 remaining in the drawer. So, $P(M|L) = \frac{2}{7}$.

Your Turn 5.2.2.1

You are about to roll a special 6-sided die that has both a colored letter and a colored number on each face. The faces are labeled with: a red 1 and a blue A, a red 1 and a green A, an orange 1 and a green B, an orange 2 and a red C, a purple 3 and a brown D, an orange 4 and a blue E. Find the given conditional probabilities:

$P(\text{the letter is a vowel}|\text{the number is orange})$

$P(\text{the number is 1}|\text{the number is red})$

$P(\text{the number is red}|\text{the number is 1})$

Checkpoint

In Tree Diagrams, Tables, and Outcomes, we introduced the concept of dependence between stages of a multistage experiment. We stated at the time that two stages were dependent if the result of one stage affects the other stage. We explained that dependence in terms of the sample space, but sometimes that dependence can be a little more subtle; it's more properly understood in terms of conditional probabilities. Two stages of an experiment are dependent if $P(E|F) \neq P(E|F')$ for some outcome of the second stage E and outcome of the first stage F.

Who Knew?: Protecting Bombers in World War II

In his book *How Not to Be Wrong*, Jordan Ellenberg recounts this anecdote: During World War II, the American military wanted to add additional armor plating to bomber aircraft, in order to reduce the chances that they get shot down. So, they collected data on planes after returning from missions. The data showed that the fuselage, wings, and fuel system had many more bullet holes (per unit area) than the engine compartments, so the military brass wanted to add additional armor to the parts of the plane that were hit most often. Luckily, before they added the armor to the planes, they asked for a second opinion. Abraham Wald, a Jewish mathematician who had fled the rising Nazi regime, pointed out that it was far more important that the armor plating be added to areas where there were *fewer* bullet holes. Why? The planes they were studying had already completed their missions, so the military was essentially looking at conditional probabilities: the probability of suffering a bullet strike, *given that* the plane made it back safely. More bullet holes in an area on the plane indicated that was a region that wasn't as important for the plane's survival!

Compound Events Using "And" and the Multiplication Rule

For multistage experiments, the outcomes of the experiment as a whole are often stated in terms of the outcomes of the individual stages. Commonly, those statements are joined with "and." For example, in the sock drawer example just above, one outcome might be "the left sock is black and the right sock is blue." As with "or" compound events, these probabilities can be computed with basic arithmetic.

FORMULA

Multiplication Rule for Probability: If E and F are events associated with the first and second stages of an experiment, then $P(E\text{and } F) = P(E) \times P(F|E)$.

Checkpoint

*In The Addition Rule for Probability, we considered probabilities of events connected with "and" in the statement of the Inclusion/Exclusion Principle. These two scenarios are different; in the statement of the Inclusion/Exclusion Principle, the events connected with "and" are both events associated with the **same** single-stage experiment (or the same stage of a multistage experiment). In the Multiplication Rule, we're looking at events associated with **different** stages of a multistage experiment.*

Example 5.2.2.2: Using the Multiplication Rule for Probability

You are president of a club with 10 members: 4 seniors, 3 juniors, 2 sophomores, and 1 first-year. You need to choose 2 members to represent the club on 2 college committees. The first person selected will be on the Club Awards Committee and the second will be on the New Club Orientation Committee. The same person cannot be selected for both. You decide to select these representatives at random.

1. What is the probability that a senior is chosen for both positions?
2. What is the probability that a junior is chosen first and a sophomore is chosen second?

3. What is the probability that a sophomore is chosen first and a senior is chosen second?

Answer

1. We need the probability that a senior is chosen first and a senior is chosen second. These are two stages of a multistage experiment, so we'll apply the Multiplication Rule for Probability:
$P(\text{senior chosen first and senior chosen second}) = P(\text{senior chosen first}) \times P(\text{senior chosen second}|\text{senior chosen first})$.
Since there are 4 seniors among the 10 members, $P(\text{senior chosen first}) = \frac{4}{10} =$. Next, assuming a senior is chosen first, there are 3 seniors among the 9 remaining members. So, $P(\text{senior chosen second}|\text{senior chosen first}) = \frac{3}{9} = \frac{1}{3}$. Putting this all together, we get $P(\text{senior chosen first and senior chosen second}) = \frac{2}{5} \times \frac{1}{3} = \frac{2}{15}$.
2. There are 3 juniors among the 10 members, so $P(\text{junior chosen first}) = \frac{3}{10}$. Assuming a junior is chosen first, there are 2 sophomores among the remaining 9 members, so $P(\text{sophomore chosen second}|\text{junior chosen first}) = \frac{2}{9}$. Thus, using the Multiplication Rule for Probability, we have $P(\text{junior chosen first and sophomore chosen second}) = \frac{3}{10} \times \frac{2}{9} = \frac{1}{15}$.
3. The probability that a sophomore is chosen first is $\frac{2}{10} = \frac{1}{5}$, and the probability that a senior is chosen second given that a sophomore was chosen first is $\frac{4}{9}$. Thus, using the Multiplication Rule for Probability, we have:
$P(\text{sophomore chosen first and senior chosen second}) = \frac{1}{5} \times \frac{4}{9} =$.

Your Turn 5.2.2.2

You're drawing 2 cards in order from a deck containing only the cards A♡, A♣, K♠, 10♡, 9♡, 9♠, 9♢, and 6♣. Compute the following:

$P(\text{draw a } \heartsuit \text{ first and a } \spadesuit \text{ second})$

$P(\text{draw a 9 first and a 6 second})$

$P(\text{draw two hearts})$

WORK IT OUT: The Birthday Problem

One of the most famous problems in probability theory is the Birthday Problem, which has to do with shared birthdays in a large group. To make the analysis easier, we'll ignore leap days, and assume that the probability of being born on any given date is $\frac{1}{365}$. Now, if you have 366 people in a room, we're guaranteed to have at least one pair of people who share a single birthday. Imagine filling the room by first admitting someone born on January 1, then someone born on January 2, and so on… The 365th person admitted would be born on December 31. If you add one more person to the room, that person's birthday would *have* to match someone else's.

Let's look at the other end of the spectrum. If you choose two people at random, what is the probability that they share a birthday? As with many probability questions, this is best addressed by find out the probability that they do *not* share a birthday. The first person's birthday can be anything (probability 1), and the second person's birthday can be anything other than the first person's birthday (probability $\frac{364}{365}$). The probability that they have *different* birthdays is $1 \times \frac{364}{365} = \frac{364}{365}$. So, the probability that they share a birthday is $1 - \frac{364}{365} = \frac{1}{365}$.

What if we have three people? The probability that they all have different birthdays can be obtained by extending our previous calculation: The probability that two people have different birthdays is $\frac{364}{365}$, so if we add a third to the mix, the probability that they have a different birthday from the other two is $\frac{363}{365}$. So, the probability that all three have different birthdays is $\frac{364}{365} \times \frac{363}{365} \approx 0.9918$, and thus the probability that there's a shared birthday in the group is $1 - 0.9918 \approx 0.0082$

The big question is this: How many people do we need in the room to have the probability of a shared birthday greater than $\frac{1}{2}$? Make a guess, then with a partner keep adding hypothetical people to the group and computing probabilities until you get there!

It is often useful to combine the rules we've seen so far with the techniques we used for finding sample spaces. In particular, trees can be helpful when we want to identify the probabilities of every possible outcome in a multistage experiment. The next example will illustrate this.

Example 5.2.2.3: Using Tree Diagrams to Help Find Probabilities

The board game *Clue* uses a deck of 21 cards: 6 suspects, 6 weapons, and 9 rooms. Suppose you are about to draw 2 cards from this deck. There are 6 possible outcomes for the draw: 2 suspects, 2 weapons, 2 rooms, 1 suspect and 1 weapon, 1 suspect and 1 room, or 1 weapon and 1 room. What are the probabilities for each of these outcomes?

Answer

Step 1: Let's start by building a tree diagram that illustrates both stages of this experiment. Let's use S, W, and R to indicate drawing a suspect, weapon, and room, respectively (Figure 5.2.2.3).

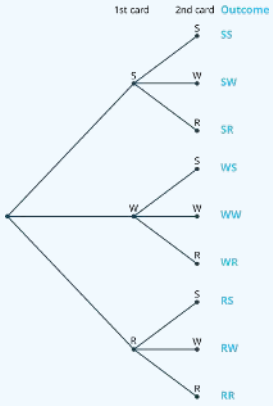


Figure 5.2.2.3

Step 2: We want to start computing probabilities, starting with the first stage. The probability that the first card is a suspect is $\frac{6}{21} = \frac{2}{7}$. The probability that the first card is a weapon is the same: $\frac{2}{7}$. Finally, the probability that the first card is a room is $\frac{9}{21} = \frac{3}{7}$.

Step 3: Let's incorporate those probabilities into our tree: label the edges going into each of the nodes representing the first-stage outcomes with the corresponding probabilities (Figure 5.2.2.4).

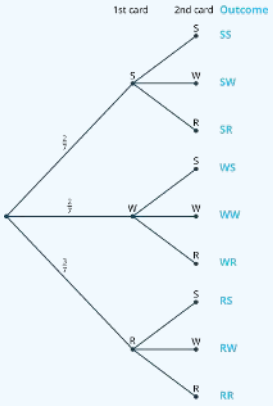


Figure 5.2.2.4

Note that the sum of the probabilities coming out of the initial node is 1; this should always be the case for the probabilities coming out of any node!

Step 4: Let's look at the case where the first card is a suspect. There are 3 edges emanating from that node (leading to the outcomes SS, SW, and SR). We'll label those edges with the appropriate *conditional* probabilities, under the assumption that the first card is a suspect. First, there are 5 remaining suspect cards among the 20 left in the deck, so $P(\text{second is suspect}|\text{first is suspect}) = \frac{5}{20} = \frac{1}{4}$. Using similar reasoning, we can compute $P(\text{second is weapon}|\text{first is suspect}) = \frac{6}{20} = \frac{3}{10}$ and $P(\text{second is room}|\text{first is suspect}) = \frac{9}{20}$.

Step 5: Checking our work, we see that the sum of these 3 probabilities is again equal to 1. Let's add those to our tree (Figure 5.2.2.5).

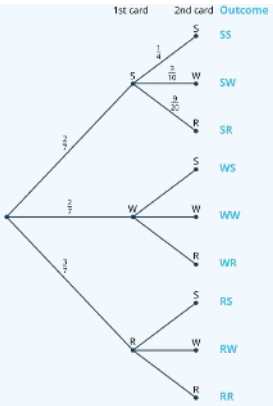


Figure ✔ 5.2.2.5

Step 6: Let's continue filling in the conditional probabilities at the other nodes, always checking to make sure the sum of the probabilities coming out of any node is equal to 1 (Figure ✔ 5.2.2.6).

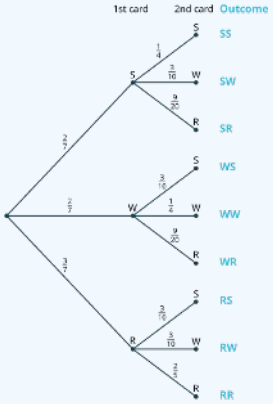


Figure ✔ 5.2.2.6

Step 7: We can compute the probability of landing on any final node by multiplying the probabilities along the path we would take to get there. For example, the probability of drawing a suspect first and a weapon second (i.e., ending up on the node labeled "SW") is $\frac{2}{7} \times \frac{3}{10} = \frac{3}{35}$

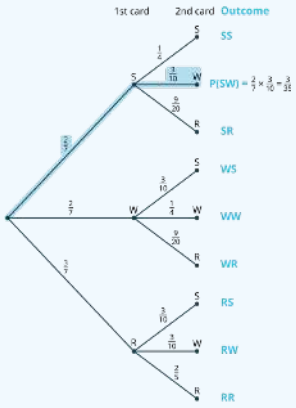


Figure ✔ 5.2.2.7

Step 8: Let's fill in the rest of the probabilities (Figure ✔ 5.2.2.8).

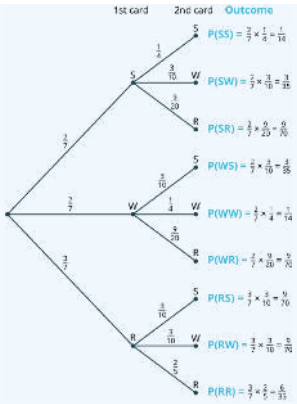


Figure ✔ 5.2.2.8

Step 9: A helpful feature of tree diagrams is that the final outcomes are always mutually exclusive, so the Addition Rule can be directly applied. For example, the probability of drawing one suspect and one room (in any order) would be $P(SR) + P(RS) = \frac{9}{70} + \frac{9}{70} =$ 935P(SR)+P(RS)=970+970=935. We can find the probabilities of the other outcomes in a similar fashion, as shown in the following table:

Outcome	Probability
2 suspects	$\frac{1}{14}$
2 weapons	$\frac{1}{14}$
2 rooms	$\frac{6}{35}$
1 suspect and 1 weapon	$\frac{3}{35} + \frac{3}{35} = \frac{6}{35}$
1 suspect and 1 room	$\frac{9}{70} + \frac{9}{70} = \frac{9}{35}$
1 weapon and 1 room	$\frac{9}{70} + \frac{9}{70} = \frac{9}{35}$

Checking once again, the sum of these 6 probabilities is 1, as expected.

Your Turn 5.2.2.31

You are about to perform the following two-stage experiment. First, you will flip a coin. If the result is heads, roll a standard 6-sided die. If the result of the coin flip is tails, roll a modified 6-sided die with faces labeled 1, 1, 1, 2, 2, 3. Use a tree diagram to find the probability of rolling each of the numbers from 1 to 6.

WORK IT OUT: The Monty Hall Problem

On the original version of the game show *Let's Make a Deal*, originally hosted by Monty Hall and now hosted by Wayne Brady, one contestant was chosen to play a game for the grand prize of the day (often a car). Here's how it worked: On the stage were three areas concealed by numbered curtains. The car was hidden behind one of the curtains; the other two curtains hid worthless prizes (called "Zonks" on the show). The contestant would guess which curtain concealed the car. To build tension, Monty would then reveal what was behind one of the *other* curtains, which was always one of the Zonks (Since Monty knew where the car was hidden, he always had at least one Zonk curtain that hadn't been chosen that he could reveal). Monty then turned to the contestant and asked: "Do you want to stick with your original choice, or do you want to switch your choice to the other curtain?" What should the contestant do? Does it matter?

With a partner or in a small group, simulate this game. You can do that with a small candy (the prize) hidden under one of three cups, or with three playing cards (just decide ahead of time which card represents the "Grand Prize"). One person plays the host, who knows where the prize is hidden. Another person plays the contestant and tries to guess where the prize is hidden. After the guess is made, the host should reveal a losing option that wasn't chosen by the contestant. The contestant then has the option to stick with the original choice or switch to the other, unrevealed option. Play about 20 rounds, taking turns in each role and making sure that both contestant strategies (stick or switch) are used equally often. After each round, make a note of whether the contestant chose "stick" or "switch"

and whether the contestant won or lost. Find the empirical probability of winning under each strategy. Then, see if you can use tree diagrams to verify your findings.

Check Your Understanding

For the following exercises, you are rolling two 6-sided dice, each of which has 3 orange faces, 2 green faces, and 1 blue face.

1. What is the probability of rolling 2 orange faces?
2. What is the probability of rolling 2 green faces?
3. What is the probability of rolling 1 orange and 1 green face (in any order)?
4. For the following exercises, you are about to draw 2 cards at random (without replacement) from a deck containing only these 10 cards: A♡, A♠, A♣, A♢, K♠, K♣, Q♡, Q♠, J♡, J♠.
5. What is the probability of drawing 2 aces?
6. What is the probability of drawing an ace first and a king second?
7. What is the probability of drawing a ♣ and a ♠ (in any order)?

This page titled 5.2.2: Conditional Probability and the Multiplication Rule is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **7.9: Conditional Probability and the Multiplication Rule** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

5.3: Binomial Distributions

5.3: Binomial Distributions is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

5.3.1: The Binomial Distribution



Figure 5.3.1.1: If one baseball team has a 65% chance of beating another in any single game, what's the likelihood that they win a best-of-seven series? (credit: "baseball game" by Britt Reints/Flickr, CC BY 2.0)

Learning Objectives

1. Identify binomial experiments.
2. Use the binomial distribution to analyze binomial experiments.

It's time for the World Series, which determines the champion for this season in Major League Baseball. The scrappy Los Angeles Angels are facing the powerhouse Cincinnati Reds. Computer models put the chances of the Reds winning any single game against the Angels at about 65%. The World Series, as its name implies, isn't just one game, though: it's what's known as a "best-of-seven" contest: the teams play each other repeatedly until one team wins 4 games (which could take up to 7 games total, if each team wins three of the first 6 matchups). If the Reds truly have a 65% chance of winning a single game, then the probability that they win the series should be greater than 65%. Exactly how much bigger?

If you have the patience for it, you could use a tree diagram like we used in Example 7.33 to trace out all of the possible outcomes, find all the related probabilities, and add up the ones that result in the Reds winning the series. Such a tree diagram would have $2^7 = 128$ final nodes, though, so the calculations would be very tedious. Fortunately, we have tools at our disposal that allow us to find these probabilities very quickly. This section will introduce those tools and explain their use.

Binomial Experiments

The tools of this section apply to multistage experiments that satisfy some pretty specific criteria. Before we move on to the analysis, we need to introduce and explain those criteria so that we can recognize experiments that fall into this category. Experiments that satisfy each of these criteria are called binomial experiments. A binomial experiment is an experiment with a fixed number of repeated independent binomial trials, where each trial has the same probability of success.

Repeated Binomial Trials

The first criterion involves the structure of the stages. Each stage of the experiment should be a replication of every other stage; we call these replications trials. An example of this is flipping a coin 10 times; each of the ten flips is a trial, and they all occur under the same conditions as every other. Further, each trial must have only two possible outcomes. These two outcomes are typically labeled "success" and "failure," even if there is not a positive or negative connotation associated with those outcomes. Experiments with more than two outcomes in their sample spaces are sometimes reconsidered in a way that forces just two outcomes; all we need to do is completely divide the sample space into two parts that we can label "success" and "failure." For example, your grade on an exam might be recorded as A, B, C, D, or F, but we could instead think of the grades A, B, C, and D as "success" and a grade of F as "failure." Trials with only two outcomes are called binomial trials (the word *binomial* derives from Latin and Greek roots that mean "two parts").

Independent Trials

The next criterion that we'll be looking for is independence of trials. Back in Tree Diagrams, Tables, and Outcomes, we said that two stages of an experiment are independent if the outcome of one stage doesn't affect the other stage. Independence is necessary for the experiments we want to analyze in this section.

Fixed Number of Trials

Next, we require that the number of trials in the experiment be decided before the experiment begins. For example, we might say "flip a coin 10 times." The number of trials there is fixed at 10. However, if we say "flip a coin until you get 5 heads," then the number of trials could be as low as 5, but theoretically it could be 50 or a 100 (or more)! We can't apply the tools from this section in cases where the number of trials is indeterminate.

Constant Probability

The next criterion needed for binomial experiments is related to the independence of the trials. We must make sure that the probability of success in each trial is the same as the probability of success in every other trial.

✔ Example ✔ 5.3.1.1: Identifying Binomial Experiments

Decide whether each of the following is a binomial experiment. For those that aren't, identify which criterion or criteria are violated.

1. You roll a standard 6-sided die 10 times and write down the number that appears each time.
2. You roll a standard 6-sided die 10 times and write down whether the die shows a 6 or not.
3. You roll a standard 6-sided die until you get a 6.
4. You roll a standard 6-sided die 10 times. On the first roll, we define "success" as rolling a 4 or greater. After the first roll, we define "success" as rolling a number greater than the result of the previous roll.

Answer

1. Since we're noting 1 of 6 possible outcomes, the trials are not binomial. So, this isn't a binomial experiment.
2. We have 2 possible outcomes ("6" and "not 6"), the trials are independent, the probability of success is the same every time, and the number of trials is fixed. This is a binomial experiment.
3. Since the number of trials isn't fixed (we don't know if we'll get our first 6 after 1 roll or 20 rolls or somewhere in between), this isn't a binomial experiment.
4. Here, the probability of success might change with every roll (on the first roll, that probability is $\frac{1}{2}$; if the first roll is a 6, the probability of success on the next roll is zero). So, this is not a binomial experiment.

📌 Your Turn 📌 5.3.1.1

Decide whether the following experiments are binomial experiments:

1. Draw a card from a well-shuffled deck, note its suit, and replace it. Repeat this process 5 times.
2. Draw 5 cards from a well-shuffled deck and count the number of ♣.
3. Draw a card from a well-shuffled deck, note whether it is a ♣ or not, and replace it. Repeat this process 5 times.
4. Draw cards from a well-shuffled deck until you get a ♣.

The Binomial Formula

If we flip a coin 100 times, you might expect the number of heads to be around 50, but you probably wouldn't be surprised if the actual number of heads was 47 or 52. What is the probability that the number of heads is exactly 50? Or falls between 45 and 55? It seems unlikely that we would get more than 70 heads. Exactly how unlikely is that?

Each of these questions is a question about the number of successes in a binomial experiment (flip a coin 100 times, "success" is flipping heads). We could theoretically use the techniques we've seen in earlier sections to answer each of these, but the number of calculations we'd have to do is astronomical; just building the tree diagram that represents this situation is more than we could complete in a lifetime; it would have $2^{100} \approx 1.3 \times 10^{30}$ final nodes! To put that number in perspective, if we could draw 1,000

dots every second, and we started at the moment of the Big Bang, we'd currently be about 0.00000003% of the way to drawing out those final nodes. Luckily, there's a shortcut called the Binomial Formula that allows us to get around doing all those calculations!

FORMULA

Binomial Formula: Suppose we have a binomial experiment with n trials and the probability of success in each trial is p. Then:

$$P(\text{ number of successes is } a) = {}_nC_a \times p^a \times (1-p)^{n-a}$$

We can use this formula to answer one of our questions about 100 coin flips. What is the probability of flipping exactly 50 heads? In this case, $n = 100, p = \frac{1}{2}$, and $a = 50$, so $P(\text{ flip 50 heads }) = {}_{100}C_{50} \times \left(\frac{1}{2}\right)^{50} \times \left(1 - \frac{1}{2}\right)^{100-50}$. Unfortunately, many calculators will balk at this calculation; that first factor $({}_{100}C_{50})$ is an enormous number, and the other two factors are very close to zero. Even if your calculator can handle numbers that large or small, the arithmetic can create serious errors in rounding off.

Tech Check

Luckily, spreadsheet programs have alternate methods for doing this calculation. In Google Sheets, we can use the BINOMDIST function to do this calculation for us. Open up a new sheet, click in any empty cell, and type "=BINOMDIST(50,100,0.5,FALSE)" followed by the Enter key. The cell will display the probability we seek; it's about 8%. Let's break down the syntax of that function in Google Sheets: enter "=BINOMDIST(a, n, p, FALSE)" to find the probability of a successes in n trials with probability of success p.

Example 5.3.1.2: Using the Binomial Formula

1. Find the probability of rolling a standard 6-sided die 4 times and getting exactly one 6 *without using technology*.
2. Find the probability of rolling a standard 6-sided die 60 times and getting exactly ten 6s using technology.
3. Find the probability of rolling a standard 6-sided die 60 times and getting exactly eight 6s using technology.

Answer

1. We'll apply the Binomial Formula, where $n = 4, a = 1$, and $p = \frac{1}{6}$:

$$\begin{aligned} P(\text{ rolling one 6}) &= {}_4C_1 \times \left(\frac{1}{6}\right)^1 \times \left(\frac{5}{6}\right)^{4-1} \\ &= \frac{4!}{1!(4-1)!} \times \frac{1}{6} \times \left(\frac{5}{6}\right)^3 \\ &= 4 \times \frac{1}{6} \times \frac{5^3}{6^3} \\ &= \frac{4 \times 5^3}{6^4} \\ &= \frac{500}{1,296} \end{aligned}$$

2. Here, $n = 60, a = 10$, and $p = \frac{1}{6}$. In Google Sheets, we'll enter "=BINOMDIST(10, 60, 1/6, FALSE)" to get our result: 0.137 .

3. This experiment is the same as in Exercise 2 of this example; we're simply changing the number of successes from 10 to 8 . Making that change in the formula in Google Sheets, we get the probability 0.116.

Your Turn 5.3.1.2

Compute the probabilities (rounded to 3 decimal places) of the following events related to rolling a standard 4-sided die (with faces labeled 1, 2, 3, and 4):

1. You roll the die 10 times and get exactly four 2s.
2. You roll the die 20 times and get exactly four 2s.
3. You roll the die 30 times and get exactly four 2s.

The Binomial Distribution

If we are interested in the probability of more than just a single outcome in a binomial experiment, it's helpful to think of the Binomial Formula as a function, whose input is the number of successes and whose output is the probability of observing that many successes. Generally, for a small number of trials, we'll give that function in table form, with a complete list of the possible outcomes in one column and the probability in the other.

For example, suppose Kristen is practicing her basketball free throws. Assume Kristen always makes 82% of those shots. If she attempts 5 free throws, then the Binomial Formula gives us these probabilities:

Shots Made	Probability
0	0.000189
1	0.004304
2	0.0392144
3	0.1786432
4	0.4069096
5	0.3707398

A table that lists all possible outcomes of an experiment along with the probabilities of those outcomes is an example of a probability density function (PDF). A PDF may also be a formula that you can use to find the probability of any outcome of an experiment.

Checkpoint

Because they refer to the same thing, some sources will refer to the Binomial Formula as the Binomial PDF.

If we want to know the probability of a range of outcomes, we could add up the corresponding probabilities. Going back to Kristen's free throws, we can find the probability that she makes 3 or fewer of her 5 attempts by adding up the probabilities associated with the corresponding outcomes (in this case: 0, 1, 2, or 3):

$$\begin{aligned} P(\text{makes 3 or fewer}) &= P(a=0)+P(a=1)+P(a=2)+P(a=3) \\ &= 0.000189+0.004304+0.0392144+0.1786432 \\ &= 0.2223506 \end{aligned}$$

The probability that the outcome of an experiment is less than or equal to a given number is called a *cumulative probability*. A table of the cumulative probabilities of all possible outcomes of an experiment is an example of a cumulative distribution function (CDF). A CDF may also be a formula that you can use to find those cumulative probabilities.

Checkpoint

Cumulative probabilities are always associated with events that are defined using $\leq$. If other inequalities are used to define the event, we must restate the definition so that it uses the correct inequality.

Here are the PDF and CDF for Kristen's free throws:

Shots Made	Probability	Cumulative
0	0.000189	0.000189
1	0.004304	0.004493

Shots Made	Probability	Cumulative
2	0.0392144	0.0437073
3	0.1786432	0.2223506
4	0.4069096	0.6292602
5	0.3707398	1

Tech Check

Google Sheets can also compute cumulative probabilities for us; all we need to do is change the “FALSE” in the formulas we used before to "TRUE."

Example 5.3.1.3: Using the Binomial CDF

Suppose we are about to flip a fair coin 50 times. Let H represent the number of heads that result from those flips. Use technology to find the following:

1. $P(H \leq 22)$
2. $P(H < 26)$
3. $P(H > 28)$
4. $P(H \geq 20)$
5. $P(20 < H < 25)$

Answer

1. The event here is defined by $H \leq$22H≤22, which is the inequality we need to have if we want to use the Binomial CDF. In Google Sheets, we’ll enter “=BINOMDIST(22, 50, 0.5, TRUE)” to get our answer: 0.2399.
2. This event uses the wrong inequality, so we need to do some preliminary work. If $H <$26H<26, that means $H \leq$25H≤25 (because H has to be a whole number). So, we’ll enter “=BINOMDIST(25, 50, 0.5, TRUE)” to find $P(H < 26) = P(H \leq 25) = 0.5561$.
3. The inequality associated with this event is pointing in the wrong direction. If E is the event $H >$28H>28, that means that E contains the outcomes {29, 30, 31, 32, 33, …}. Thus, E' must contain the outcomes {…, 25, 26, 27, 28}. In other words, E' is defined by $H \leq$28H≤28. Since it uses $\leq$, we can find $P(E')$ using “=BINOMDIST(28, 50, 0.5, TRUE)”: 0.8389 So, using the formula for probabilities of complements, we have
$$\begin{aligned} P(E) &= 1 - P(E') \\ &= 1 - 0.8389 \\ &= 0.1611. \end{aligned}$$
4. As in part 3, this inequality is pointing in the wrong direction. If F is the event $H \geq$20H≥20, then F contains the outcomes {20, 21, 22, 23, …}. That means F' contains the outcomes {…, 16, 17, 18, 19}, and so F' is defined by $H \leq$19H≤19. So, we can find $P(F')$ using “=BINOMDIST(19, 50, 0.5, TRUE)”: 0.0595. Finally, using the formula for probabilities of complements, we get:
$$\begin{aligned} P(F) &= 1 - P(F') \\ &= 1 - 0.0595 \\ &= 0.9405. \end{aligned}$$
5. If $20 < H < 25$, that means we are interested in the outcomes {21, 22, 23, 24}. This doesn’t look like any of the previous situations, but there is a way to find this probability using the Binomial CDF. We need to put everything in terms of “less than or equal to,” so we’ll first note that all of our outcomes are less than or equal to 24. But we don’t want to include values that are less than or equal to 20. So, we have three events: let I be the event defined by $20 < H < 25$ (note that we’re trying to find $P(I)$). Let J be defined by $H \leq$24H≤24, and let K be defined by $H \leq$20H≤20. Of these three events, J contains the most outcomes. If J occurs, then either K or I must have occurred.

Moreover, K and I are mutually exclusive. Thus, $P(J) = P(K) + P(I)$, by the Addition Rule. Solving for the probability that we want, we get

$$\begin{aligned} P(I) &= P(J) - P(K) \\ &= P(H \leq 24) - P(H \leq 20) \\ &= 0.44386 - 0.10132 \\ &= 0.34254. \end{aligned}$$

Your Turn 5.3.1.3

You are about to roll a standard 6-sided die 20 times. Let S denote a success, which will be rolling a number greater than 4. Find the probabilities of the following events, rounded to 4 decimal places:

1. $S \leq 5$
2. $S < 8$
3. $S > 10$
4. $S \geq 7$
5. $5 < S < 8$

Finally, we can answer the question posed at the beginning of this section. Remember that the Reds are facing the Angels in the World Series, which is won by the team who is first to win 4 games. The Reds have a 65% chance to win any game against the Angels. So, what is the probability that the Reds win the World Series? At first glance, this is not a binomial experiment: The number of games played is not fixed, since the series ends as soon as one team wins 4 games. However, we can extend this situation to a binomial experiment: Let's assume that 7 games are always played in the World Series, and the winner is the team who wins more games. In a way, this is what happens in reality; it's as though the first team to lose 4 games (and thus cannot win more than the other team) forfeits the rest of their games. So, we can treat the actual World Series as a binomial experiment with seven trials. If W is the number of games won by the Reds, the probability that the Reds win the World Series is $P(W \geq 4)$. Using the techniques from the last example, we get $P(\text{Reds win the series}) = 0.8002$.

People in Mathematics: Abraham de Moivre

Abraham de Moivre was born in 1667 in France to a Protestant family. Though he was educated in Catholic schools, he remained true to his faith; in 1687, he fled with his brother to London to escape persecution under the reign of King Louis XIV. Once he arrived in England, he supported himself as a freelance math tutor while he conducted his own research. Among his interests was probability; in 1711, he published the first edition of *The Doctrine of Chances: A Method of Calculating the Probabilities of Events in Play*. This book was the second textbook on probability (after Cardano's *Liber de ludo aleae*). De Moivre discovered an important connection between the binomial distribution and the *normal distribution* (an important concept in statistics; we'll explore that distribution and its connection to the binomial distribution in Chapter 8). De Moivre also discovered some properties of a new probability distribution that later became known as the Poisson distribution.

Check Your Understanding

You are rolling a 6-sided die with 3 orange faces, 2 green faces, and 1 blue face.

1. If you roll the die 5 times and note the color showing on each roll, is this a binomial experiment?
2. If you roll the die 5 times and count the number times you roll a green face, is this a binomial experiment?
3. If you count how many times you roll the die until you get a blue face, is this a binomial experiment?

Suppose you're rolling the same colored 6-sided die 10 times. Let O, G, and B represent the number of times the die lands with an orange, green, and blue side up, respectively. Find these probabilities (round to 4 decimal places):

4. $P(O \leq 7)$
5. $P(G = 5)$
6. $P(1 \leq B \leq 4)$

This page titled 5.3.1: The Binomial Distribution is shared under a CC BY 4.0 license and was authored, remixed, and/or curated by OpenStax via source content that was edited to the style and standards of the LibreTexts platform.

- **7.10: The Binomial Distribution** by OpenStax is licensed CC BY 4.0. Original source: https://openstax.org/details/books/contemporary-mathematics.

5.4: Markov Chains

5.4: Markov Chains is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.

5.4.1: Stochastic Matrices

Objectives

1. Learn examples of stochastic matrices and applications to difference equations.
2. Understand Google's PageRank algorithm.
3. *Recipe:* find the steady state of a positive stochastic matrix.
4. *Picture:* dynamics of a positive stochastic matrix.
5. *Theorem:* the Perron–Frobenius theorem.
6. *Vocabulary words:* ***difference equation, (positive) stochastic matrix, steady state, importance matrix, Google matrix***.

This section is devoted to one common kind of application of eigenvalues: to the study of difference equations, in particular to Markov chains. We will introduce stochastic matrices, which encode this type of difference equation, and will cover in detail the most famous example of a stochastic matrix: the Google Matrix.

Difference Equations

Suppose that we are studying a system whose state at any given time can be described by a list of numbers: for instance, the numbers of rabbits aged $0, 1$, and 2 years, respectively, or the number of copies of Prognosis Negative in each of the Red Box kiosks in Atlanta. In each case, we can represent the state at time t by a vector v_t. We assume that t represents a discrete time quantity: in other words, v_t is the state "on day t" or "at year t". Suppose in addition that the state at time $t+1$ is related to the state at time t in a linear way: $v_{t+1} = Av_t$ for some matrix A. This is the situation we will consider in this subsection.

Definition 5.4.1.1: Difference Equation

A ***difference equation*** is an equation of the form

$$v_{t+1} = Av_t$$

for an $n \times n$ matrix A and vectors $v_0, v_1, v_2, \ldots$ in $\mathbb{R}^n$.

In other words:

- v_t is the "state at time t,"
- v_{t+1} is the "state at time $t+1$," and
- $v_{t+1} = Av_t$ means that A is the "change of state matrix."

Note that

$$v_t = Av_{t-1} = A^2v_{t-2} = \cdots = A^tv_0,$$

which should hint to you that the long-term behavior of a difference equation is an eigenvalue problem.

We will use the following example in this subsection and the next. Understanding this section amounts to understanding this example.

Example 5.4.1.1

Red Box has kiosks all over Atlanta where you can rent movies. You can return them to any other kiosk. For simplicity, pretend that there are three kiosks in Atlanta, and that every customer returns their movie the next day. Let v_t be the vector whose entries x_t, y_t, z_t are the number of copies of Prognosis Negative at kiosks $1, 2$, and 3, respectively. Let A be the matrix whose i, j-entry is the probability that a customer renting Prognosis Negative from kiosk j returns it to kiosk i. For example, the matrix

$$A = \begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix}$$

encodes a 30% probability that a customer renting from kiosk 3 returns the movie to kiosk 2, and a 40% probability that a movie rented from kiosk 1 gets returned to kiosk 3. The second row (for instance) of the matrix A says:

> *The number of movies returned to kiosk 2 will be (on average):*
>
> 30% of the movies from kiosk 1
> 40% of the movies from kiosk 2
> 30% of the movies from kiosk 3

Applying this to all three rows, this means

$$A\begin{pmatrix} x_t \\ y_t \\ z_t \end{pmatrix} = \begin{pmatrix} .3x_t + .4y_t + .5z_t \\ .3x_t + .4y_t + .3z_t \\ .4x_t + .2y_t + .2z_t \end{pmatrix}.$$

Therefore, Av_t represents the number of movies in each kiosk the next day:

$$Av_t = v_{t+1}.$$

This system is modeled by a difference equation.

An important question to ask about a difference equation is: what is its long-term behavior? How many movies will be in each kiosk after 100 days? In the next subsection, we will answer this question for a particular type of difference equation.

Example 5.4.1.2: Rabbit population

In a population of rabbits,

1. half of the newborn rabbits survive their first year;
2. of those, half survive their second year;
3. the maximum life span is three years;
4. rabbits produce 0, 6, 8 rabbits in their first, second, and third years, respectively.

Let v_t be the vector whose entries x_t, y_t, z_t are the number of rabbits aged $0, 1$, and 2, respectively. The rules above can be written as a system of equations:

$$\begin{aligned} x_{t+1} &= 6y_t + 8z_t \\ y_{t+1} &= \tfrac{1}{2}x_t \\ z_{t+1} &= \tfrac{1}{2}y_t. \end{aligned}$$

In matrix form, this says:

$$\begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} v_t = v_{t+1}.$$

This system is modeled by a difference equation.

Define

$$A = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

We compute A has eigenvalues 2 and -1, and that an eigenvector with eigenvalue 2 is

$$v = \begin{pmatrix} 16 \\ 4 \\ 1 \end{pmatrix}.$$

This partially explains why the ratio $x_t : y_t : z_t$ approaches $16 : 4 : 1$ and why all three quantities eventually double each year in this demo:

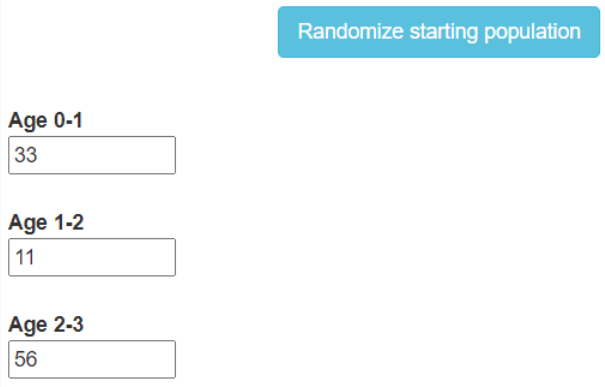


Figure 5.4.1.1: Left: the population of rabbits in a given year. Right: the proportions of rabbits in that year. Choose any values you like for the starting population, and click "Advance 1 year" several times. Notice that the ratio $x_t : y_t : z_t$ approaches $16 : 4 : 1$, and that all three quantities eventually double each year.

Stochastic Matrices and the Steady State

In this subsection, we discuss difference equations representing *probabilities*, like the Red Box example 5.4.1.1. Such systems are called *Markov chains*. The most important result in this section is the Perron–Frobenius theorem, which describes the long-term behavior of a Markov chain.

Definition 5.4.1.2: Stochastic

A square matrix A is ***stochastic*** if all of its entries are nonnegative, and the entries of each column sum to 1.

A matrix is *positive* if all of its entries are positive numbers.

Note 5.4.1.1

A positive stochastic matrix is a stochastic matrix whose entries are all positive numbers. In particular, no entry is equal to zero. For instance, the first matrix below is a positive stochastic matrix, and the second is not:

$$\begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Remark

More generally, a ***regular*** stochastic matrix is a stochastic matrix A such that A^n is positive for some $n \geq 1$. The Perron–Frobenius theorem below also applies to regular stochastic matrices.

Example 5.4.1.3

Continuing with Red Box example 5.4.1.1, the matrix

$$A = \begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix}$$

is a positive stochastic matrix. The fact that the columns sum to 1 says that all of the movies rented from a particular kiosk must be returned to *some* other kiosk (remember that every customer returns their movie the next day). For instance, the first column says:

Of the movies rented from kiosk 1,

$$30\% \text{ will be returned to kiosk } 1$$
$$30\% \text{ will be returned to kiosk } 2$$
$$40\% \text{ will be returned to kiosk } 3.$$

The sum is 100%*, as all of the movies are returned to one of the three kiosks.*

The matrix A represents the change of state from one day to the next:

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \\ z_{t+1} \end{pmatrix} = A \begin{pmatrix} x_t \\ y_t \\ z_t \end{pmatrix} = \begin{pmatrix} .3x_t + .4y_t + .5z_t \\ .3x_t + .4y_t + .3z_t \\ .4x_t + .2y_t + .2z_t \end{pmatrix}.$$

If we sum the entries of v_{t+1}, we obtain

$$\begin{aligned} &(.3x_t + .4y_t + .5z_t) + (.3x_t + .4y_t + .3z_t) + (.4x_t + .2y_t + .2z_t) \\ &\quad = (.3 + .3 + .4)x_t + (.4 + .4 + .2)y_t + (.5 + .3 + .2)z_t \\ &\quad = x_t + y_t + z_t. \end{aligned}$$

This says that the *total* number of copies of Prognosis Negative in the three kiosks does not change from day to day, as we expect.

The fact that the entries of the vectors v_t and v_{t+1} sum to the same number is a consequence of the fact that the columns of a stochastic matrix sum to 1.

Note 5.4.1.2

Let A be a stochastic matrix, let v_t be a vector, and let $v_{t+1} = Av_t$. Then the sum of the entries of v_t equals the sum of the entries of v_{t+1}.

Computing the long-term behavior of a difference equation turns out to be an eigenvalue problem. The eigenvalues of stochastic matrices have very special properties.

Fact 5.4.1.1

Let A be a stochastic matrix. Then:

1. 1 is an eigenvalue of A.
2. If λ is a (real or complex) eigenvalue of A, then $|\lambda| \leq 1$.

Proof

If A is stochastic, then the *rows* of A^T sum to 1. But multiplying a matrix by the vector $(1, 1, \ldots, 1)$ sums the rows:

$$\begin{pmatrix} .3 & .3 & .4 \\ .4 & .4 & .2 \\ .5 & .3 & .2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} .3 + .3 + .4 \\ .4 + .4 + .2 \\ .5 + .3 + .2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Therefore, 1 is an eigenvalue of A^T. But A and A^T have the same characteristic polynomial:

$$\det(A - \lambda I_n) = \det\big((A - \lambda I_n)^T\big) = \det(A^T - \lambda I_n).$$

Therefore, 1 is an eigenvalue of A.

Now let λ be any eigenvalue of A, so it is also an eigenvalue of A^T. Let $x = (x_1, x_2, \ldots, x_n)$ be an eigenvector of A^T with eigenvalue λ, so $\lambda x = A^T x$. The jth entry of this vector equation is

$$\lambda x_j = \sum_{i=1}^{n} a_{ij} x_i.$$

Choose x_j with the largest absolute value, so $|x_i| \leq |x_j|$ for all i. Then

$$|\lambda| \cdot |x_j| = \left| \sum_{i=1}^{n} a_{ij} x_i \right| \leq \sum_{i=1}^{n} a_{ij} \cdot |x_i| \leq \sum_{i=1}^{n} a_{ij} \cdot |x_j| = 1 \cdot |x_j|,$$

where the last equality holds because $\sum_{i=1}^{n} a_{ij} = 1$. This implies $|\lambda| \leq 1$.

In fact, for a *positive* stochastic matrix A, one can show that if $\lambda \neq 1$ is a (real or complex) eigenvalue of A, then $|\lambda| < 1$. The 1-eigenspace of a stochastic matrix is very important.

Definition 5.4.1.3: Steady State

A *steady state* of a stochastic matrix A is an eigenvector w with eigenvalue 1, such that the entries are *positive* and sum to 1.

The Perron–Frobenius theorem describes the long-term behavior of a difference equation represented by a stochastic matrix. Its proof is beyond the scope of this text.

Theorem 5.4.1.1: Perron-Frobenius Theorem

Let A be a positive stochastic matrix. Then A admits a unique steady state vector w, which spans the 1-eigenspace.

Moreover, for any vector v_0 with entries summing to some number c, the iterates

$$v_1 = Av_0, \ v_2 = Av_1, \ \ldots, \ v_t = Av_{t-1}, \ \ldots$$

approach cw as t gets large.

Translation: The Perron–Frobenius theorem makes the following assertions:

- The 1-eigenspace of a positive stochastic matrix is a line.
- The 1-eigenspace contains a vector with positive entries.
- All vectors approach the 1-eigenspace upon repeated multiplication by A.

One should think of a steady state vector w as a vector of *percentages*. For example, if the movies are distributed according to these percentages today, then they will be have the same distribution tomorrow, since $Aw = w$. And no matter the starting distribution of movies, the long-term distribution will always be the steady state vector.

The sum c of the entries of v_0 is the *total number* of things in the system being modeled. The total number does not change, so the long-term state of the system must approach cw: it is a multiple of w because it is contained in the 1-eigenspace, and the entries of cw sum to c.

Recipe 1: Compute the Steady State Vector

Let A be a positive stochastic matrix. Here is how to compute the steady-state vector of A.

1. Find any eigenvector v of A with eigenvalue 1 by solving $(A - I_n)v = 0$.
2. Divide v by the sum of the entries of v to obtain a vector w whose entries sum to 1.
3. This vector automatically has positive entries. It is the unique steady-state vector.

The above recipe is suitable for calculations by hand, but it does not take advantage of the fact that A is a stochastic matrix. In practice, it is generally faster to compute a steady state vector by computer as follows:

Recipe 2: Approximate the Steady State Vector by Computer

Let A be a positive stochastic matrix. Here is how to approximate the steady-state vector of A with a computer.

1. Choose any vector v_0 whose entries sum to 1 (e.g., a standard coordinate vector).
2. Compute $v_1 = Av_0$, $v_2 = Av_1$, $v_3 = Av_2$, etc.
3. These converge to the steady state vector w.

Example 5.4.1.4: A 2×2 stochastic matrix

Consider the positive stochastic matrix

$$A = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}.$$

This matrix has characteristic polynomial

$$f(\lambda) = \lambda^2 - \operatorname{Tr}(A)\lambda + \det(\lambda) = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = (\lambda - 1)(\lambda - 1/2).$$

Notice that 1 is strictly greater than the other eigenvalue, and that it has algebraic (hence, geometric) multiplicity 1. We compute eigenvectors for the eigenvalues $1, 1/2$ to be, respectively,

$$u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad u_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

The eigenvector u_1 necessarily has positive entries; the steady-state vector is

$$w = \frac{1}{1+1}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 50\% \\ 50\% \end{pmatrix}.$$

The Perron–Frobenius theorem asserts that, for any vector v_0, the vectors $v_1 = Av_0$, $v_2 = Av_1$, $\ldots$ approach a vector whose entries are the same: 50% of the sum will be in the first entry, and 50% will be in the second.

We can see this explicitly, as follows. The eigenvectors u_1, u_2 form a basis $\mathcal{B}$ for $\mathbb{R}^2$; for any vector $x = a_1u_1 + a_2u_2$ in $\mathbb{R}^2$, we have

$$Ax = A(a_1u_1 + a_2u_2) = a_1Au_1 + a_2Au_2 = a_1u_1 + \frac{a_2}{2}u_2.$$

Iterating multiplication by A in this way, we have

$$A^t x = a_1u_1 + \frac{a_2}{2^t}u_2 \;\to\; a_1u_1$$

as $t \to \infty$. This shows that $A^t x$ approaches

$$a_1u_1 = \begin{pmatrix} a_1 \\ a_1 \end{pmatrix}.$$

Note that the sum of the entries of a_1u_1 is equal to the sum of the entries of $a_1u_1 + a_2u_2$, since the entries of u_2 sum to 0.

To illustrate the theorem with numbers, let us choose a particular value for u_0, say $u_0 = \binom{1}{0}$. We compute the values for $u_t = \binom{x_t}{y_t}$ in this table:

t	x_t	y_t
0	1.000	0.000
1	0.750	0.250
2	0.625	0.375
3	0.563	0.438
4	0.531	0.469
5	0.516	0.484
6	0.508	0.492
7	0.504	0.496
8	0.502	0.498
9	0.501	0.499
10	0.500	0.500

We see that u_t does indeed approach $\binom{0.5}{0.5}$.

Now we turn to visualizing the dynamics of (i.e., repeated multiplication by) the matrix A. This matrix is diagonalizable; we have $A = CDC^{-1}$ for

$$C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

The matrix D leaves the x-coordinate unchanged and scales the y-coordinate by $1/2$. Repeated multiplication by D makes the y-coordinate very small, so it "sucks all vectors into the x-axis."

The matrix A does the same thing as D, but with respect to the coordinate system defined by the columns u_1, u_2 of C. This means that A "sucks all vectors into the 1-eigenspace", without changing the sum of the entries of the vectors.

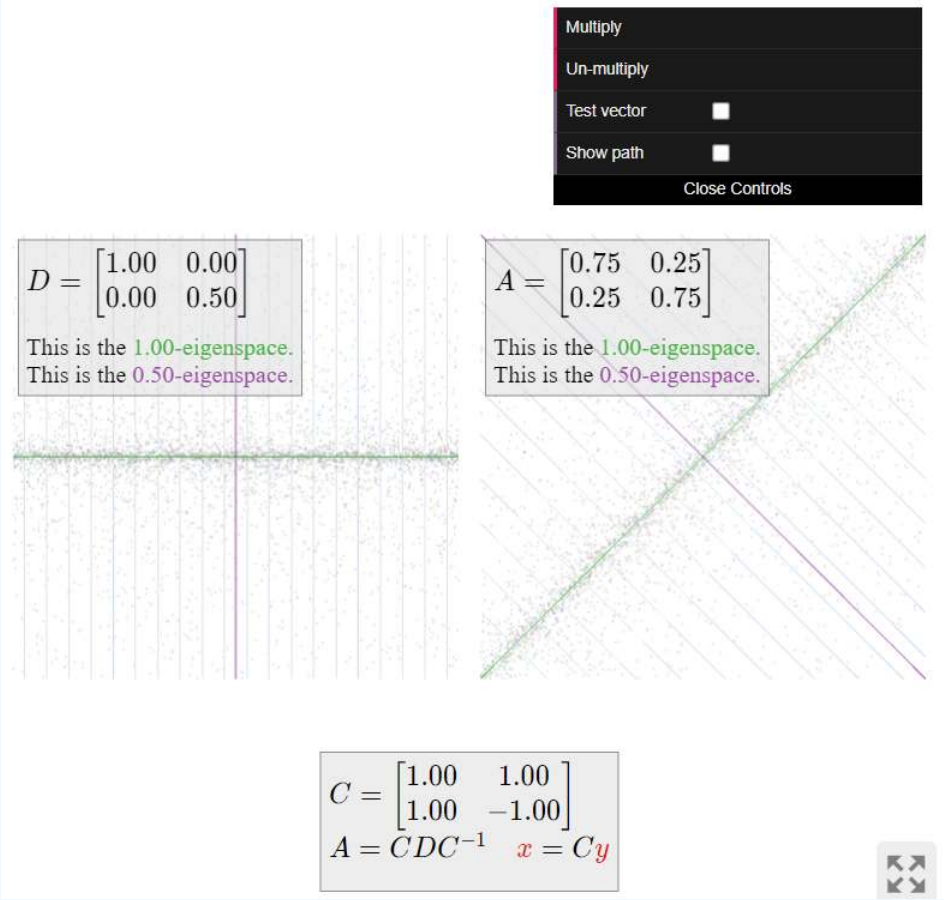


Figure 5.4.1.2: Dynamics of the stochastic matrix A. Click "multiply" to multiply the colored points by D on the left and A on the right. Note that on both sides, all vectors are "sucked into the 1-eigenspace" (the red line).

Example 5.4.1.5

Continuing with Red Box example 5.4.1.1, we can illustrate the Perron–Frobenius theorem explicitly. The matrix

$$A = \begin{pmatrix} .3 & .4 & .5 \\ .3 & .4 & .3 \\ .4 & .2 & .2 \end{pmatrix}$$

has characteristic polynomial

$$f(\lambda) = -\lambda^3 + 0.12\lambda - 0.02 = -(\lambda - 1)(\lambda + 0.2)(\lambda - 0.1).$$

Notice that 1 is strictly greater in absolute value than the other eigenvalues, and that it has algebraic (hence, geometric) multiplicity 1. We compute eigenvectors for the eigenvalues $1, -0.2, 0.1$to be, respectively,

$$u_1 = \begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix} \qquad u_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \qquad u_3 = \begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix}.$$

The eigenvector u_1 necessarily has positive entries; the steady-state vector is

$$w = \frac{1}{7+6+5}u_1 = \frac{1}{18}\begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix}.$$

The eigenvectors u_1, u_2, u_3 form a basis $\mathcal{B}$ for $\mathbb{R}^3$; for any vector $x = a_1u_1 + a_2u_2 + a_3u_3$ in $\mathbb{R}^3$, we have

$$\begin{aligned} Ax &= A(a_1u_1 + a_2u_2 + a_3u_3) \\ &= a_1Au_1 + a_2Au_2 + a_3Au_3 \\ &= a_1u_1 - 0.2a_2u_2 + 0.1a_3u_3. \end{aligned}$$

Iterating multiplication by A in this way, we have

$$A^tx = a_1u_1 - (0.2)^ta_2u_2 + (0.1)^ta_3u_3 \;\to\; a_1u_1$$

as $t \to \infty$. This shows that A^tx approaches a_1u_1, which is an *eigenvector with eigenvalue* 1, as guaranteed by the Perron–Frobenius theorem.

What do the above calculations say about the number of copies of Prognosis Negative in the Atlanta Red Box kiosks? Suppose that the kiosks start with 100 copies of the movie, with 30 copies at kiosk 1, 50 copies at kiosk 2, and 20 copies at kiosk 3. Let $v_0 = (30, 50, 20)$ be the vector describing this state. Then there will be $v_1 = Av_0$ movies in the kiosks the next day, $v_2 = Av_1$ the day after that, and so on. We let $v_t = (x_t, y_t, z_t)$.

t	x_t	y_t	z_t
0	30.000000	50.000000	20.000000
1	39.000000	35.000000	26.000000
2	38.700000	33.500000	27.800000
3	38.910000	33.350000	27.740000
4	38.883000	33.335000	27.782000
5	38.889900	33.333500	27.776600
6	38.888670	33.333350	27.777980
7	38.888931	33.333335	27.777734
8	38.888880	33.333333	27.777786
9	38.888891	33.333333	27.777776
10	38.888889	33.333333	27.777778

(Of course it does not make sense to have a fractional number of movies; the decimals are included here to illustrate the convergence.) The steady-state vector says that eventually, the movies will be distributed in the kiosks according to the percentages

$$w = \frac{1}{18}\begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix} = \begin{pmatrix} 38.888888\% \\ 33.333333\% \\ 27.777778\% \end{pmatrix},$$

which agrees with the above table. Moreover, this distribution is *independent* of the beginning distribution of movies in the kiosks.

Now we turn to visualizing the dynamics of (i.e., repeated multiplication by) the matrix A. This matrix is diagonalizable; we have $A = CDC^{-1}$ for

$$C = \begin{pmatrix} 7 & -1 & 1 \\ 6 & 0 & -3 \\ 5 & 1 & 2 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -.2 & 0 \\ 0 & 0 & .1 \end{pmatrix}.$$

The matrix D leaves the x-coordinate unchanged, scales the y-coordinate by $-1/5$, and scales the z-coordinate by $1/10$. Repeated multiplication by D makes the y- and z-coordinates very small, so it "sucks all vectors into the x-axis."

The matrix A does the same thing as D, but with respect to the coordinate system defined by the columns u_1, u_2, u_3 of C. This means that A "sucks all vectors into the 1-eigenspace", without changing the sum of the entries of the vectors.

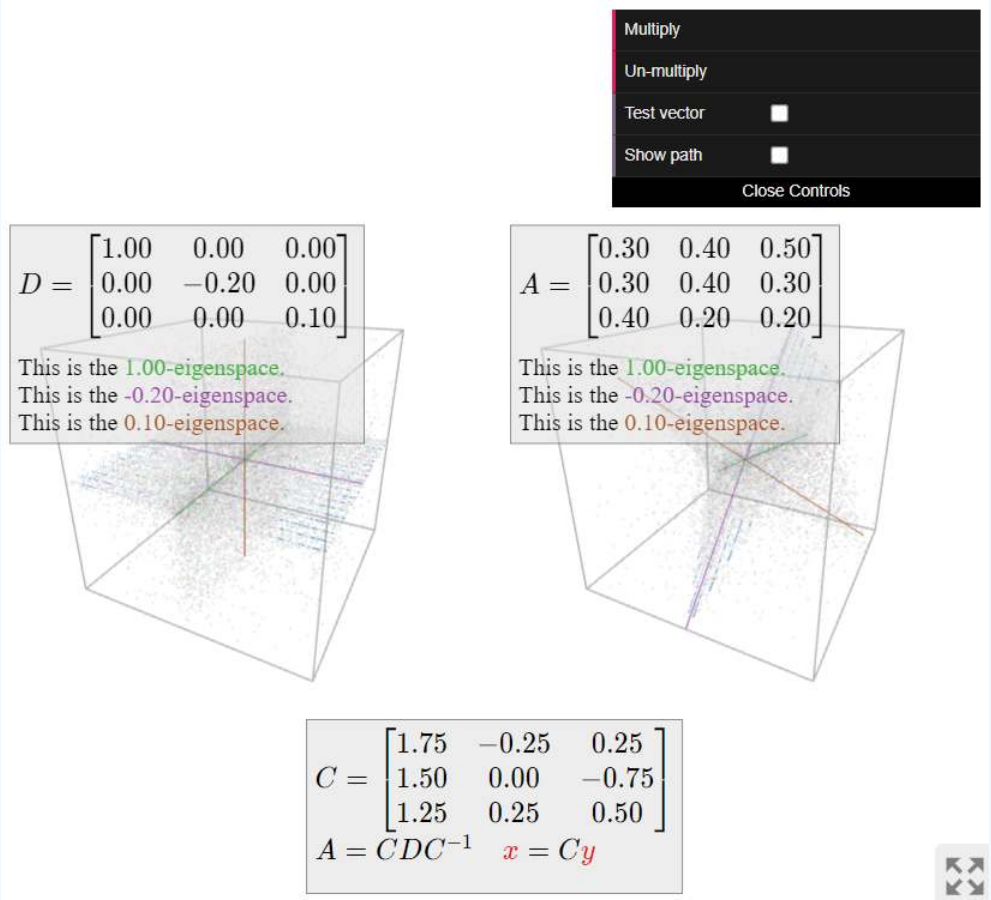


Figure 5.4.1.3: Dynamics of the stochastic matrix A. Click "multiply" to multiply the colored points by D on the left and A on the right. Note that on both sides, all vectors are "sucked into the 1-eigenspace" (the green line). (We have scaled C by $1/4$ so that vectors have roughly the same size on the right and the left. The "jump" that happens when you press "multiply" is a negation of the $-.2$-eigenspace, which is not animated.)

The picture of a positive stochastic matrix is always the same, whether or not it is diagonalizable: *all vectors are "sucked into the 1-eigenspace,"* which is a line, without changing the sum of the entries of the vectors. This is the geometric content of the Perron–Frobenius theorem.

Google's PageRank Algorithm

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 5.4.1.4: 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 5.4.1.5: 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.

Observe that the importance matrix is a stochastic matrix, assuming every page contains a link: if page i has m links, then the ith column contains the number $1/m$, m times, and the number zero in the other entries.

Example 5.4.1.6

Consider the following internet with only four pages. Links are indicated by arrows.

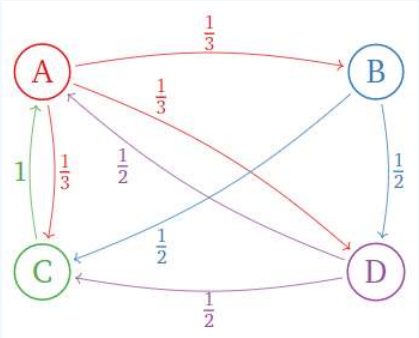


Figure 5.4.1.4

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

$$\begin{pmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} c + \frac{1}{2}d \\ \frac{1}{3}a \\ \frac{1}{3}a + \frac{1}{2}b + \frac{1}{2}d \\ \frac{1}{3}a + \frac{1}{2}b \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

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 5.4.1.1

The rank vector is an eigenvector of the importance matrix with eigenvalue 1.

In light of the key observation, we would like to use the Perron–Frobenius theorem to find the rank vector. Unfortunately, the importance matrix is not always a *positive* stochastic matrix.

Example 5.4.1.7: A page with no links

Consider the following internet with three pages:

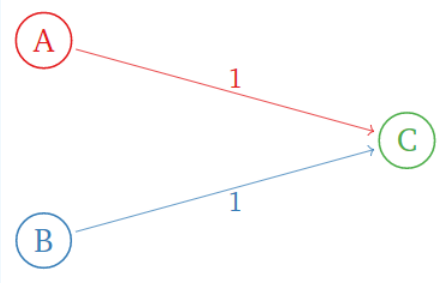


Figure 5.4.1.5

The importance matrix is

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

This has characteristic polynomial

$$f(\lambda) = \det \begin{pmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 1 & 1 & -\lambda \end{pmatrix} = -\lambda^3.$$

So 1 is not an eigenvalue at all: there is no rank vector! The importance matrix is not stochastic because the page C has no links.

Example 5.4.1.8: Disconnected Internet

Consider the following internet:

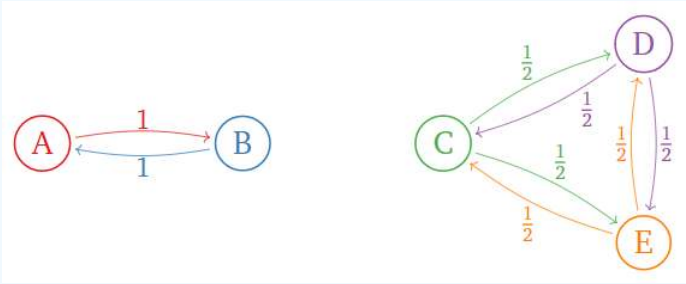


Figure 5.4.1.6

The importance matrix is

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}.$$

This has linearly independent eigenvectors

$$\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

both with eigenvalue 1. So there is more than one rank vector in this case. Here the importance matrix is stochastic, but not positive.

Here is Page and Brin's solution. First we fix the importance matrix by replacing each zero column with a column of $1/n$s, where n is the number of pages:

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{becomes} \quad A' = \begin{pmatrix} 0 & 0 & 1/3 \\ 0 & 0 & 1/3 \\ 1 & 1 & 1/3 \end{pmatrix}.$$

The ***modified importance matrix*** A' is always stochastic.

Now we choose a number p in $(0, 1)$, called the ***damping factor***. (A typical value is $p = 0.15$.)

Definition 5.4.1.6: 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

$$M = (1-p) \cdot A' + p \cdot B \quad \text{where} \quad B = \frac{1}{n} \begin{pmatrix} 1 & 1 & \cdot & 1 \\ 1 & 1 & \cdot & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

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. The reader can verify the following important fact.

Fact 5.4.1.2

The Google Matrix is a positive stochastic matrix.

If we declare that the ranks of all of the pages must sum to 1, then we find:

Note 5.4.1.3: The 25 Billion Dollar Eigenvector

The PageRank vector is the steady state of the Google Matrix.

This exists and has positive entries by the Perron–Frobenius theorem. The hard part is calculating it: in real life, the Google Matrix has zillions of rows.

Example 5.4.1.9: PageRank Vector

What is the PageRank vector for the following internet? (Use the damping factor $p = 0.15$.)

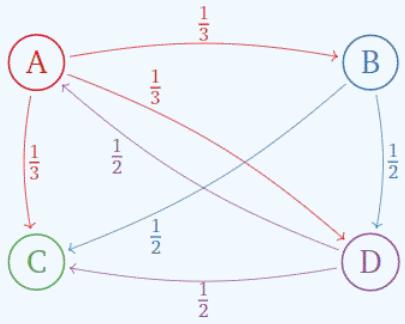


Figure 5.4.1.7

Which page is the most important? Which is the least important?

Solution

First we compute the modified importance matrix:

$$A = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix} \xrightarrow{\text{modify}} A' = \begin{pmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & 0 \end{pmatrix}$$

Choosing the damping factor $p = 0.15$, the Google Matrix is

$$M = 0.85 \cdot \begin{pmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & 0 \end{pmatrix} + 0.15 \cdot \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

$$\approx \begin{pmatrix} 0.0375 & 0.0375 & 0.2500 & 0.4625 \\ 0.3208 & 0.0375 & 0.2500 & 0.0375 \\ 0.3208 & 0.4625 & 0.2500 & 0.4625 \\ 0.3208 & 0.4625 & 0.2500 & 0.0375 \end{pmatrix}.$$

The PageRank vector is the steady state:

$$w \approx \begin{pmatrix} 0.2192 \\ 0.1752 \\ 0.3558 \\ 0.2498 \end{pmatrix}$$

This is the PageRank:

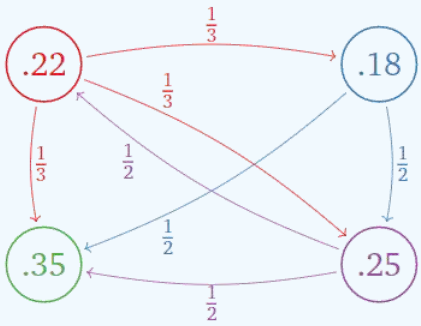


Figure 5.4.1.8

Page C is the most important, with a rank of 0.558, and page B is the least important, with a rank of 0.1752.

This page titled 5.4.1: Stochastic Matrices is shared under a GNU Free Documentation License 1.3 license and was authored, remixed, and/or curated by Dan Margalit & Joseph Rabinoff via source content that was edited to the style and standards of the LibreTexts platform.

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

Index

A

augmented matrix
- 3.2.1: Gaussian Elimination
- 3.3.1: Gaussian Elimination

B

blended cost
- 2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know)

C

cardinality
- 4.2.1: Cardinality

Cartesian Coordinate Plane
- 2.1.4: Graphing Equations by Hand

complement
- 4.1.2: Union, Intersection, and Complement

constant matrix
- 3.2.1: Gaussian Elimination
- 3.3.1: Gaussian Elimination

constraints
- 3.7.1: Maximization Applications

contribution margin
- 2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know)

D

damping factor
- 5.4.1: Stochastic Matrices

difference equations
- 5.4.1: Stochastic Matrices

E

elementary row operations
- 3.2.1: Gaussian Elimination
- 3.3.1: Gaussian Elimination

elements
- 4.1.1: Basics of Sets

elimination method
- 3.1.1: Solving a System of Equations Algebraically

F

fixed cost
- 2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know)

G

gaussian elimination
- 3.2.1: Gaussian Elimination
- 3.3.1: Gaussian Elimination

I

identity matrix
- 3.5.1: The Identity and Inverses

importance matrix
- 5.4.1: Stochastic Matrices

importance rule
- 5.4.1: Stochastic Matrices

intersection
- 4.1.2: Union, Intersection, and Complement

K

Kronecker delta
- 3.5.1: The Identity and Inverses

L

linear programming
- 3.7.1: Maximization Applications

M

matrix
- 3.4.1: Matrix Arithmetic

matrix inverse
- 3.5.2: Finding the Inverse of a Matrix

mixed constraints
- 3.7.1: Maximization Applications

O

objective function
- 3.7.1: Maximization Applications

ordered pairs
- 2.1.4: Graphing Equations by Hand

P

PageRank
- 5.4.1: Stochastic Matrices

Perron–Frobenius theorem
- 5.4.1: Stochastic Matrices

proper subset
- 4.1.1: Basics of Sets

R

random surfer interpretation
- 5.4.1: Stochastic Matrices

rank
- 5.4.1: Stochastic Matrices

S

set
- 4.1.1: Basics of Sets

slope
- 2.2.1: Slope of a Line

square matrix
- 3.4.1: Matrix Arithmetic

standard minimization linear program
- 3.7.2: Minimization Applications

stochastic matrices
- 5.4.1: Stochastic Matrices

subset
- 4.1.1: Basics of Sets

substitution method
- 3.1.1: Solving a System of Equations Algebraically

U

union
- 4.1.2: Union, Intersection, and Complement

unit variable cost
- 2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know)

universal set
- 4.1.2: Union, Intersection, and Complement

V

variable cost
- 2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know)

Venn diagrams
- 4.1.3: Venn Diagrams

Z

zero matrix
- 3.4.1: Matrix Arithmetic

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: MAT 149: Topics in Finite Mathematics (Holz)

Webpages: 85

Applicable Restrictions: Noncommercial, No Derivatives

All licenses found:

- Undeclared: 49.4% (42 pages)
- CC BY 4.0: 30.6% (26 pages)
- CC BY-SA 3.0: 8.2% (7 pages)
- CC BY-NC-SA 4.0: 4.7% (4 pages)
- CC BY-NC-ND 3.0: 1.2% (1 page)
- CC BY-NC-SA 2.5: 1.2% (1 page)
- CC BY-NC-ND 4.0: 1.2% (1 page)
- CC BY-NC 3.0: 1.2% (1 page)
- CC BY-NC-SA 3.0: 1.2% (1 page)
- GNU Free Documentation License: 1.2% (1 page)

By Page

- MAT 149: Topics in Finite Mathematics (Holz) - *Undeclared*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Logic - *Undeclared*
    - 1.1: Statements and Truth Values - *Undeclared*
      - 1.1.1: Introduction - *CC BY 4.0*
      - 1.1.2: Statements and Quantifiers - *CC BY 4.0*
      - 1.1.3: Compound Statements - *CC BY 4.0*
    - 1.2: Truth Tables - *Undeclared*
      - 1.2.1: Constructing Truth Tables - *CC BY 4.0*
      - 1.2.2: Truth Tables for the Conditional and Biconditional - *CC BY 4.0*
    - 1.3: Conditional Statements - *Undeclared*
      - 1.3.1: Equivalent Statements - *CC BY 4.0*
    - 1.4: Deductive Arguments - *Undeclared*
      - 1.4.1: Evaluating Deductive Arguments with Truth Tables - *CC BY-SA 3.0*
      - 1.4.2: Forms of Valid Arguments - *CC BY-SA 3.0*
  - 2: Functions - *Undeclared*
    - 2.1: Functions - *Undeclared*
      - 2.1.1: Function Notation - *CC BY-NC-SA 4.0*
      - 2.1.2: Domain and Range of a Function - *CC BY-NC-SA 4.0*
      - 2.1.3: Piecewise-defined Functions - *CC BY-NC-SA 4.0*
      - 2.1.4: Graphing Equations by Hand - *CC BY-NC-ND 3.0*
    - 2.2: Linear Functions - *Undeclared*
      - 2.2.1: Slope of a Line - *CC BY 4.0*
        - 2.2.1E: Exercises - *CC BY 4.0*
      - 2.2.2: Find the Equation of a Line - *CC BY 4.0*
    - 2.3: Linear Models - *Undeclared*
      - 2.3.1: Linear Models - *CC BY-NC-SA 2.5*
      - 2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know) - *CC BY-NC-SA 4.0*
    - 2.4: Graphing Linear Functions - *Undeclared*
      - 2.4.1: Graphing Linear Equations - *Undeclared*
    - 2.5: Linear Regression - *Undeclared*
  - 3: Systems and Matrices - *Undeclared*
    - 3.1: Systems of Linear Equations - *Undeclared*
      - 3.1.1: Solving a System of Equations Algebraically - *CC BY-NC-ND 4.0*
    - 3.2: Matrices - *Undeclared*
      - 3.2.1: Gaussian Elimination - *CC BY 4.0*
    - 3.3: Solving Using Matrices - *Undeclared*
      - 3.3.1: Gaussian Elimination - *CC BY 4.0*
    - 3.4: Matrix Algebra - *Undeclared*
      - 3.4.1: Matrix Arithmetic - *CC BY 4.0*
      - 3.4.2: Multiplication of Matrices - *CC BY 4.0*
      - 3.4.3: The ijth Entry of a Product - *CC BY 4.0*
    - 3.5: Matrix Inverses - *Undeclared*
      - 3.5.1: The Identity and Inverses - *CC BY 4.0*

- 3.5.2: Finding the Inverse of a Matrix - *CC BY 4.0*
- 3.5.3: The Matrix Inverse - *CC BY-NC 3.0*
  - 3.5.3.1: Exercises 2.6 - *Undeclared*
- 3.6: Linear Inequalities - *Undeclared*
  - 3.6.1: Solving Systems of Linear Inequalities (Two Variables) - *CC BY-NC-SA 3.0*
- 3.7: Linear Programming - *Undeclared*
  - 3.7.1: Maximization Applications - *CC BY 4.0*
    - 3.7.1.1: Maximization Applications (Exercises) - *CC BY 4.0*
  - 3.7.2: Minimization Applications - *CC BY 4.0*

4: Sets and Counting - *Undeclared*

- 4.1: Sets and Operations - *Undeclared*
  - 4.1.1: Basics of Sets - *CC BY-SA 3.0*
  - 4.1.2: Union, Intersection, and Complement - *CC BY-SA 3.0*
  - 4.1.3: Venn Diagrams - *CC BY-SA 3.0*
- 4.2: Number of Elements in a Set - *Undeclared*
  - 4.2.1: Cardinality - *CC BY-SA 3.0*
- 4.3: Multiplication Principle - *Undeclared*
  - 4.3.1: Tree Diagrams, Tables, and Outcomes - *CC BY 4.0*
  - 4.3.2: The Multiplication Rule for Counting - *CC BY 4.0*
- 4.4: Permutations and Combinations - *Undeclared*
  - 4.4.1: Permutations - *CC BY 4.0*
  - 4.4.2: Combinations - *CC BY 4.0*

5: Probability - *Undeclared*

- 5.1: Sample Spaces and Probability - *Undeclared*
  - 5.1.1: Basic Concepts - *CC BY-SA 3.0*
- 5.2: Probability Rules - *Undeclared*
  - 5.2.1: The Addition Rule for Probability - *CC BY 4.0*
  - 5.2.2: Conditional Probability and the Multiplication Rule - *CC BY 4.0*
- 5.3: Binomial Distributions - *Undeclared*
  - 5.3.1: The Binomial Distribution - *CC BY 4.0*
- 5.4: Markov Chains - *Undeclared*
  - 5.4.1: Stochastic Matrices - *GNU Free Documentation License*

Back Matter - *Undeclared*

- Index - *Undeclared*
- Glossary - *Undeclared*
- Detailed Licensing - *Undeclared*
- Detailed Licensing - *Undeclared*

Detailed Licensing

Overview

Title: MAT 149: Topics in Finite Mathematics (Holz)

Webpages: 85

Applicable Restrictions: Noncommercial, No Derivatives

All licenses found:

- Undeclared: 49.4% (42 pages)
- CC BY 4.0: 30.6% (26 pages)
- CC BY-SA 3.0: 8.2% (7 pages)
- CC BY-NC-SA 4.0: 4.7% (4 pages)
- CC BY-NC-ND 3.0: 1.2% (1 page)
- CC BY-NC-SA 2.5: 1.2% (1 page)
- CC BY-NC-ND 4.0: 1.2% (1 page)
- CC BY-NC 3.0: 1.2% (1 page)
- CC BY-NC-SA 3.0: 1.2% (1 page)
- GNU Free Documentation License: 1.2% (1 page)

By Page

- MAT 149: Topics in Finite Mathematics (Holz) - *Undeclared*
  - Front Matter - *Undeclared*
    - TitlePage - *Undeclared*
    - InfoPage - *Undeclared*
    - Table of Contents - *Undeclared*
    - Licensing - *Undeclared*
  - 1: Logic - *Undeclared*
    - 1.1: Statements and Truth Values - *Undeclared*
      - 1.1.1: Introduction - *CC BY 4.0*
      - 1.1.2: Statements and Quantifiers - *CC BY 4.0*
      - 1.1.3: Compound Statements - *CC BY 4.0*
    - 1.2: Truth Tables - *Undeclared*
      - 1.2.1: Constructing Truth Tables - *CC BY 4.0*
      - 1.2.2: Truth Tables for the Conditional and Biconditional - *CC BY 4.0*
    - 1.3: Conditional Statements - *Undeclared*
      - 1.3.1: Equivalent Statements - *CC BY 4.0*
    - 1.4: Deductive Arguments - *Undeclared*
      - 1.4.1: Evaluating Deductive Arguments with Truth Tables - *CC BY-SA 3.0*
      - 1.4.2: Forms of Valid Arguments - *CC BY-SA 3.0*
  - 2: Functions - *Undeclared*
    - 2.1: Functions - *Undeclared*
      - 2.1.1: Function Notation - *CC BY-NC-SA 4.0*
      - 2.1.2: Domain and Range of a Function - *CC BY-NC-SA 4.0*
      - 2.1.3: Piecewise-defined Functions - *CC BY-NC-SA 4.0*
      - 2.1.4: Graphing Equations by Hand - *CC BY-NC-ND 3.0*
    - 2.2: Linear Functions - *Undeclared*
      - 2.2.1: Slope of a Line - *CC BY 4.0*
        - 2.2.1E: Exercises - *CC BY 4.0*
      - 2.2.2: Find the Equation of a Line - *CC BY 4.0*
    - 2.3: Linear Models - *Undeclared*
      - 2.3.1: Linear Models - *CC BY-NC-SA 2.5*
      - 2.3.2: Cost-Revenue-Net Income Analysis (Need to Be in the Know) - *CC BY-NC-SA 4.0*
    - 2.4: Graphing Linear Functions - *Undeclared*
      - 2.4.1: Graphing Linear Equations - *Undeclared*
    - 2.5: Linear Regression - *Undeclared*
  - 3: Systems and Matrices - *Undeclared*
    - 3.1: Systems of Linear Equations - *Undeclared*
      - 3.1.1: Solving a System of Equations Algebraically - *CC BY-NC-ND 4.0*
    - 3.2: Matrices - *Undeclared*
      - 3.2.1: Gaussian Elimination - *CC BY 4.0*
    - 3.3: Solving Using Matrices - *Undeclared*
      - 3.3.1: Gaussian Elimination - *CC BY 4.0*
    - 3.4: Matrix Algebra - *Undeclared*
      - 3.4.1: Matrix Arithmetic - *CC BY 4.0*
      - 3.4.2: Multiplication of Matrices - *CC BY 4.0*
      - 3.4.3: The ijth Entry of a Product - *CC BY 4.0*
    - 3.5: Matrix Inverses - *Undeclared*
      - 3.5.1: The Identity and Inverses - *CC BY 4.0*

- 3.5.2: Finding the Inverse of a Matrix - *CC BY 4.0*
- 3.5.3: The Matrix Inverse - *CC BY-NC 3.0*
  - 3.5.3.1: Exercises 2.6 - *Undeclared*
- 3.6: Linear Inequalities - *Undeclared*
  - 3.6.1: Solving Systems of Linear Inequalities (Two Variables) - *CC BY-NC-SA 3.0*
- 3.7: Linear Programming - *Undeclared*
  - 3.7.1: Maximization Applications - *CC BY 4.0*
    - 3.7.1.1: Maximization Applications (Exercises) - *CC BY 4.0*
  - 3.7.2: Minimization Applications - *CC BY 4.0*

4: Sets and Counting - *Undeclared*

- 4.1: Sets and Operations - *Undeclared*
  - 4.1.1: Basics of Sets - *CC BY-SA 3.0*
  - 4.1.2: Union, Intersection, and Complement - *CC BY-SA 3.0*
  - 4.1.3: Venn Diagrams - *CC BY-SA 3.0*
- 4.2: Number of Elements in a Set - *Undeclared*
  - 4.2.1: Cardinality - *CC BY-SA 3.0*
- 4.3: Multiplication Principle - *Undeclared*
  - 4.3.1: Tree Diagrams, Tables, and Outcomes - *CC BY 4.0*
  - 4.3.2: The Multiplication Rule for Counting - *CC BY 4.0*
- 4.4: Permutations and Combinations - *Undeclared*
  - 4.4.1: Permutations - *CC BY 4.0*
  - 4.4.2: Combinations - *CC BY 4.0*

5: Probability - *Undeclared*

- 5.1: Sample Spaces and Probability - *Undeclared*
  - 5.1.1: Basic Concepts - *CC BY-SA 3.0*
- 5.2: Probability Rules - *Undeclared*
  - 5.2.1: The Addition Rule for Probability - *CC BY 4.0*
  - 5.2.2: Conditional Probability and the Multiplication Rule - *CC BY 4.0*
- 5.3: Binomial Distributions - *Undeclared*
  - 5.3.1: The Binomial Distribution - *CC BY 4.0*
- 5.4: Markov Chains - *Undeclared*
  - 5.4.1: Stochastic Matrices - *GNU Free Documentation License*

Back Matter - *Undeclared*

- Index - *Undeclared*
- Glossary - *Undeclared*
- Detailed Licensing - *Undeclared*
- Detailed Licensing - *Undeclared*