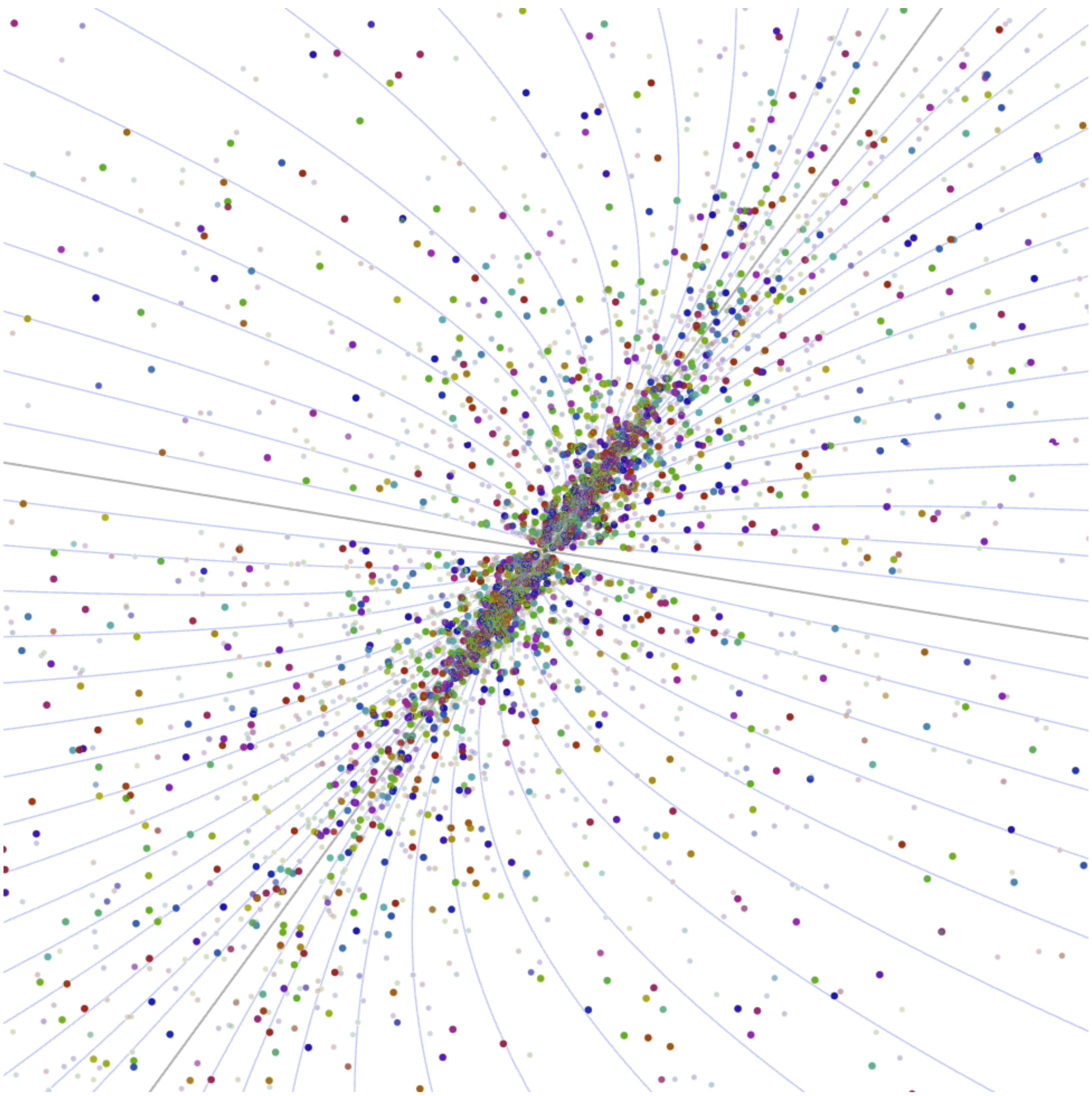
Interactive Linear Algebra



Interactive Linear Algebra

Dan Margalit
Georgia Institute of Technology

Joseph Rabinoff
Georgia Institute of Technology

June 3, 2019

Contributors to this textbook

DAN MARGALIT
School of Mathematics
Georgia Institute of Technology
dmargalit7@math.gatech.edu

LARRY ROLEN
School of Mathematics
Georgia Institute of Technology
larry.rolen@math.gatech.edu

JOSEPH RABINOFF
School of Mathematics
Georgia Institute of Technology
rabinoff@math.gatech.edu

Joseph Rabinoff contributed all of the figures, the demos, and the technical aspects of the project, as detailed below.

- The textbook is written in XML and compiled using a variant of Robert Beezer's MathBook XML, as heavily modified by Rabinoff.
- The mathematical content of the textbook is written in LaTeX, then converted to HTML-friendly SVG format using a collection of scripts called PreTeX: this was coded by Rabinoff and depends heavily on Inkscape for pdf decoding and FontForge for font embedding. The figures are written in PGF/TikZ and processed with PreTeX as well.
- The demonstrations are written in JavaScript+WebGL using Steven Wittens' brilliant framework called MathBox.

All source code can be found on GitHub. It may be freely copied, modified, and redistributed, as detailed in the appendix entitled "GNU Free Documentation License."

Larry Rolen wrote many of the exercises.

Variants of this textbook

There are several variants of this textbook available.

- The master version is the default version of the book.
- The version for math 1553 is fine-tuned to contain only the material covered in Math 1553 at Georgia Tech.

The section numbering is consistent across versions. This explains why Section 6.3 does not exist in the Math 1553 version, for example.

You are currently viewing the *master version*.

Overview

The Subject of This Textbook Before starting with the content of the text, we first ask the basic question: what *is* linear algebra?

- *Linear:* having to do with lines, planes, etc.
- *Algebra:* solving equations involving unknowns.

The name of the textbook highlights an important theme: the synthesis between algebra and geometry. It will be very important to us to understand systems of linear equations both *algebraically* (writing equations for their solutions) and *geometrically* (drawing pictures and visualizing).

Remark. The term “algebra” was coined by the 9th century mathematician Abu Ja’far Muhammad ibn Musa al-Khwarizmi. It comes from the Arabic word *al-jebr*, meaning reunion of broken parts.

At the simplest level, solving a system of linear equations is not very hard. You probably learned in high school how to solve a system like

$$\left\{\begin{aligned} x + 3y - z &= 4 \\ 2x - y + 3z &= 17 \\ y - 4z &= -3. \end{aligned}\right.$$

However, in real life one usually has to be more clever.

- Engineers need to solve many, many equations in many, many variables. Here is a tiny example:

$$\left\{\begin{aligned} 3x_1 + 4x_2 + 10x_3 + 19x_4 - 2x_5 - 3x_6 &= 141 \\ 7x_1 + 2x_2 - 13x_3 - 7x_4 + 21x_5 + 8x_6 &= 2567 \\ -x_1 + 9x_2 + \tfrac{3}{2}x_3 + x_4 + 14x_5 + 27x_6 &= 26 \\ \tfrac{1}{2}x_1 + 4x_2 + 10x_3 + 11x_4 + 2x_5 + x_6 &= -15. \end{aligned}\right.$$

- Often it is enough to know some information about the set of solutions, without having to solve the equations in the first place. For instance, does there exist a solution? What does the solution set look like geometrically? Is there still a solution if we change the 26 to a 27?

- Sometimes the coefficients also contain parameters, like the *eigenvalue equation*

$$\left\{\begin{array}{rrrl}(7-\lambda)x & + \quad y & + \quad 3z & = 0 \\ -3x & + (2-\lambda)y & - \quad 3z & = 0 \\ -3x & - \quad 2y & + (-1-\lambda)z & = 0.\end{array}\right.$$

- In data modeling, a system of equations generally does not actually have a solution. In that case, what is the best approximate solution?

Accordingly, this text is organized into three main sections.

1. Solve the matrix equation $Ax = b$ (chapters 2–4).
   - Solve systems of linear equations using matrices, row reduction, and inverses.
   - Analyze systems of linear equations geometrically using the geometry of solution sets and linear transformations.
2. Solve the matrix equation $Ax = \lambda x$ (chapters 5–6).
   - Solve eigenvalue problems using the characteristic polynomial.
   - Understand the geometry of matrices using similarity, eigenvalues, diagonalization, and complex numbers.
3. Approximately solve the matrix equation $Ax = b$ (chapter 7).
   - Find best-fit solutions to systems of linear equations that have no actual solution using least-squares approximations.
   - Study the geometry of closest vectors and orthogonal projections.

This text is roughly half computational and half conceptual in nature. The main goal is to present a library of linear algebra tools, and more importantly, to teach a conceptual framework for understanding which tools should be applied in a given context.

If Matlab can find the answer faster than you can, then your question is just an algorithm: this is not real problem solving.

The subtle part of the subject lies in understanding *what computation to ask the computer to do for you*—it is far less important to know how to perform computations that a computer can do better than you anyway.

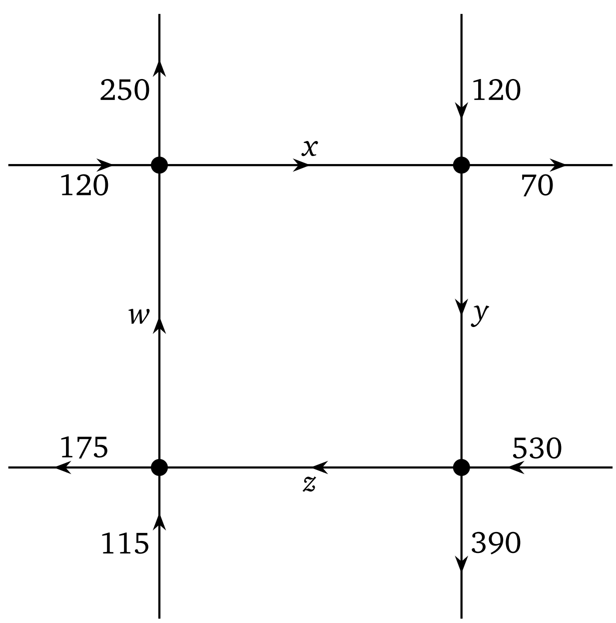
Uses of Linear Algebra in Engineering The vast majority of undergraduates at Georgia Tech have to take a course in linear algebra. There is a reason for this:

> Most engineering problems, no matter how complicated, can be reduced to linear algebra:
>
> $$Ax = b \quad \text{or} \quad Ax = \lambda x \quad \text{or} \quad Ax \approx b.$$

Here we present some sample problems in science and engineering that require linear algebra to solve.

Example (Civil Engineering)**.** The following diagram represents traffic flow around the town square. The streets are all one way, and the numbers and arrows indicate the number of cars per hour flowing along each street, as measured by sensors underneath the roads.

Traffic flow (cars/hr)



There are no sensors underneath some of the streets, so we do not know how much traffic is flowing around the square itself. What are the values of x, y, z, w? Since the number of cars entering each intersection has to equal the number of cars leaving that intersection, we obtain a system of linear equations:

$$\begin{cases} w + 120 = x + 250 \\ x + 120 = y + 70 \\ y + 530 = z + 390 \\ z + 115 = w + 175. \end{cases}$$

Example (Chemical Engineering). A certain chemical reaction (burning) takes ethane and oxygen, and produces carbon dioxide and water:

$$\underline{\ x\ }\ C_2H_6 + \underline{\ y\ }\ O_2 \to \underline{\ z\ }\ CO_2 + \underline{\ w\ }\ H_2O$$

What ratio of the molecules is needed to sustain the reaction? The following three equations come from the fact that the number of atoms of carbon, hydrogen, and oxygen on the left side has to equal the number of atoms on the right, respectively:

$$\begin{aligned} 2x &= z \\ 6x &= 2w \\ 2y &= 2z + w. \end{aligned}$$

Example (Biology). 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 baby rabbits in their first, second, and third years, respectively.

If you know the rabbit population in 2016 (in terms of the number of first, second, and third year rabbits), then what is the population in 2017? The rules for reproduction lead to the following system of equations, where x, y, z represent the number of newborn, first-year, and second-year rabbits, respectively:

$$\begin{cases} \phantom{\tfrac12 x_{2016}} \quad 6y_{2016} + 8z_{2016} = x_{2017} \\ \tfrac{1}{2}x_{2016} \qquad\qquad\qquad\quad = y_{2017} \\ \qquad\qquad \tfrac{1}{2}y_{2016} \qquad\quad\; = z_{2017}. \end{cases}$$

A common question is: what is the *asymptotic* behavior of this system? What will the rabbit population look like in 100 years? This turns out to be an eigenvalue problem.

Use this link to view the online demo

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. What do you notice about the long-term behavior of the ratios? This phenomenon turns out to be due to eigenvectors.

Example (Astronomy)**.** An asteroid has been observed at the following locations:

$$(0,2),\ (2,1),\ (1,-1),\ (-1,-2),\ (-3,1),\ (-1,-1).$$

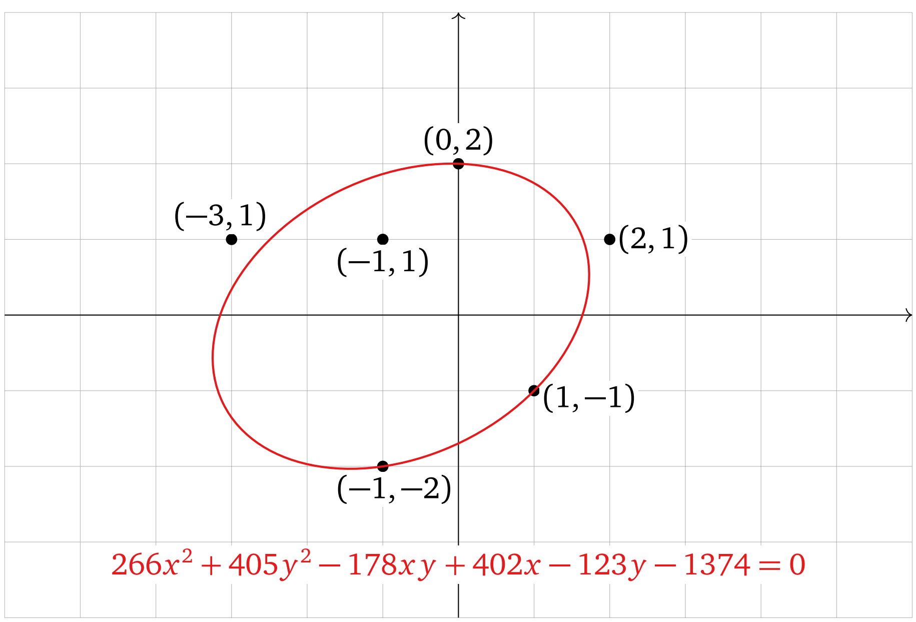
Its orbit around the sun is elliptical; it is described by an equation of the form

$$x^2 + By^2 + Cxy + Dx + Ey + F = 0.$$

What is the most likely orbit of the asteroid, given that there was some significant error in measuring its position? Substituting the data points into the above equation yields the system

$$\begin{aligned}
(0)^2 + B(2)^2 + C(0)(2) + D(0) + E(2) + F &= 0 \\
(2)^2 + B(1)^2 + C(2)(1) + D(2) + E(1) + F &= 0 \\
(1)^2 + B(-1)^2 + C(1)(-1) + D(1) + E(-1) + F &= 0 \\
(-1)^2 + B(-2)^2 + C(-1)(-2) + D(-1) + E(-2) + F &= 0 \\
(-3)^2 + B(1)^2 + C(-3)(1) + D(-3) + E(1) + F &= 0 \\
(-1)^2 + B(-1)^2 + C(-1)(-1) + D(-1) + E(-1) + F &= 0.
\end{aligned}$$

There is no actual solution to this system due to measurement error, but here is the best-fitting ellipse:



Example (Computer Science)**.** Each web page has some measure of importance, which it shares via outgoing links to other pages. This leads to zillions of equations in zillions of variables. Larry Page and Sergei Brin realized that this is a linear algebra problem at its core, and used the insight to found Google. We will discuss this example in detail in Section 5.6.

How to Use This Textbook There are a number of different categories of ideas that are contained in most sections. They are listed at the top of the section, under *Objectives*, for easy review. We classify them as follows.

- *Recipes:* these are algorithms that are generally straightforward (if sometimes tedious), and are usually done by computer in real life. They are nonetheless important to learn and to practice.

- *Vocabulary words:* forming a conceptual understanding of the subject of linear algebra means being able to communicate much more precisely than in ordinary speech. The vocabulary words have precise definitions, which must be learned and used correctly.

- *Essential vocabulary words:* these vocabulary words are essential in that they form the essence of the subject of linear algebra. For instance, if you do not know the definition of an eigenvector, then by definition you cannot claim to understand linear algebra.

- *Theorems:* these describe in a precise way how the objects of interest relate to each other. Knowing which recipe to use in a given situation generally means recognizing which vocabulary words to use to describe the situation, and understanding which theorems apply to that problem.

- *Pictures:* visualizing the geometry underlying the algebra means interpreting and drawing pictures of the objects involved. The pictures are meant to be a core part of the material in the text: they are not just a pretty add-on.

This textbook is exclusively targeted at Math 1553 at Georgia Tech. As such, it contains exactly the material that is taught in that class; no more, and no less: *students in Math 1553 are responsible for understanding all visible content.* In the online version some extra material (most examples and proofs, for instance) is hidden, in that one needs to click on a link to reveal it, like this:

Hidden Content. Hidden content is meant to enrich your understanding of the topic, but is not an official part of Math 1553. That said, the text will be very hard to follow without understanding the examples, and studying the proofs is an excellent way to learn the conceptual part of the material. (Not applicable to the PDF version.)

Finally, we remark that there are over 140 interactive demos contained in the text, which were created to illustrate the geometry of the topic. Click the “view in a new window” link, and play around with them! You will need a modern browser. Internet Explorer is not a modern browser; try Safari, Chrome, or Firefox. Here is a demo from Section 6.5:

Use this link to view the online demo

Click and drag the points on the grid on the right.

Feedback Every page of the online version has a link on the bottom for providing feedback. This will take you to the GitHub Issues page for this book. It requires a Georgia Tech login to access.

Contents

Chapter 1

Systems of Linear Equations: Algebra

Primary Goal. Solve a system of linear equations algebraically in parametric form.

This chapter is devoted to the algebraic study of systems of linear equations and their solutions. We will learn a systematic way of solving equations of the form

$$\begin{cases} 3x_1 + 4x_2 + 10x_3 + 19x_4 - 2x_5 - 3x_6 = 141 \\ 7x_1 + 2x_2 - 13x_3 - 7x_4 + 21x_5 + 8x_6 = 2567 \\ -x_1 + 9x_2 + \frac{3}{2}x_3 + x_4 + 14x_5 + 27x_6 = 26 \\ \frac{1}{2}x_1 + 4x_2 + 10x_3 + 11x_4 + 2x_5 + x_6 = -15. \end{cases}$$

In Section 1.1, we will introduce *systems of linear equations*, the class of equations whose study forms the subject of linear algebra. In Section 1.2, will present a procedure, called *row reduction*, for finding all solutions of a system of linear equations. In Section 1.3, you will see hnow to express all solutions of a system of linear equations in a unique way using the *parametric form* of the general solution.

1.1 Systems of Linear Equations

Objectives

1. Understand the definition of $\mathbf{R}^n$, and what it means to use $\mathbf{R}^n$ to label points on a geometric object.

2. *Pictures:* solutions of systems of linear equations, parameterized solution sets.

3. *Vocabulary words:* **consistent**, **inconsistent**, **solution set**.

During the first half of this textbook, we will be primarily concerned with understanding the solutions of systems of linear equations.

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

For instance,

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

are linear equations, but

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

are not.

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

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

$$\left\{ \begin{array}{rrrrrrr} x & + & 2y & + & 3z & = & 6 \\ 2x & - & 3y & + & 2z & = & 14 \\ 3x & + & y & - & z & = & -2. \end{array} \right. \qquad (1.1.1)$$

Definition (Solution sets)**.**

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

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

$$\left\{ \begin{array}{rrrrr} x & + & 2y & = & 3 \\ x & + & 2y & = & -3. \end{array} \right.$$

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

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

A solution of a system of equations in n variables is a list of n numbers. For example, $(x, y, z) = (1, -2, 3)$ is a solution of (1.1.1). As we will be studying solutions of systems of equations throughout this text, now is a good time to fix our notions regarding lists of numbers.

1.1.1 Line, Plane, Space, Etc.

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

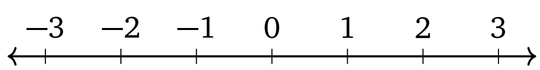
Definition. Let n be a positive whole number. We define

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

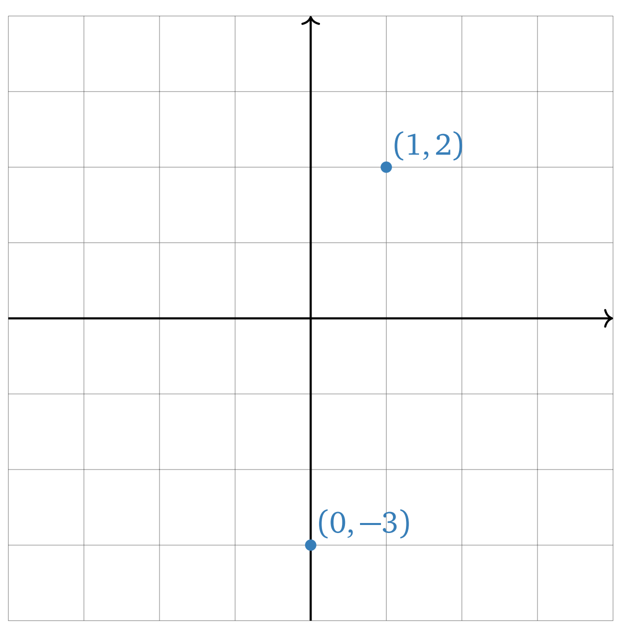
An n-tuple of real numbers is called a **point** of $\mathbf{R}^n$.

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

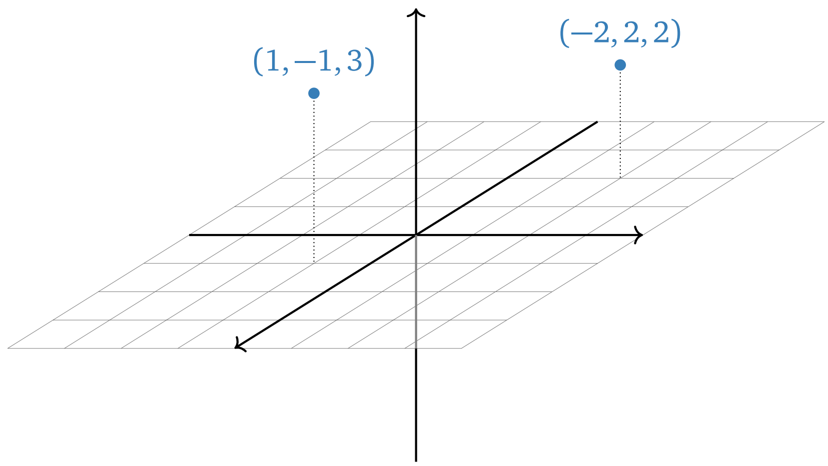
Example (The number line)**.** When $n = 1$, we just get **R** back: $\mathbf{R}^1 = \mathbf{R}$. Geometrically, this is the number line.



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



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



Interactive: Points in 3-Space.

Use this link to view the online demo

A point in 3-space, and its coordinates. Click and drag the point, or move the sliders.

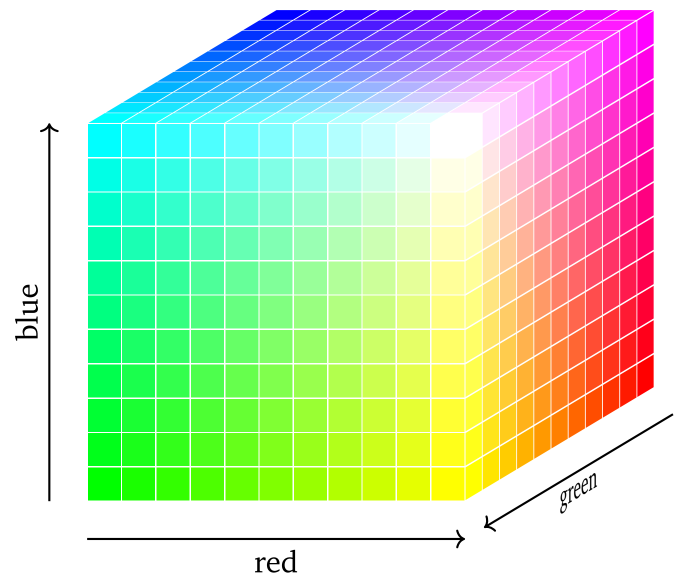
So what is $\mathbf{R}^4$? or $\mathbf{R}^5$? or $\mathbf{R}^n$? These are harder to visualize, so you have to go back to the definition: $\mathbf{R}^n$ is the set of all ordered n-tuples of real numbers $(x_1, x_2, x_3, \ldots, x_n)$.

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

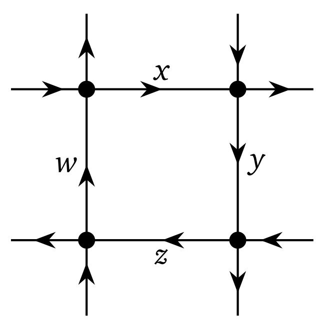
We will make definitions and state theorems that apply to any $\mathbf{R}^n$, but we will only draw pictures for $\mathbf{R}^2$ and $\mathbf{R}^3$.

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

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



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



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

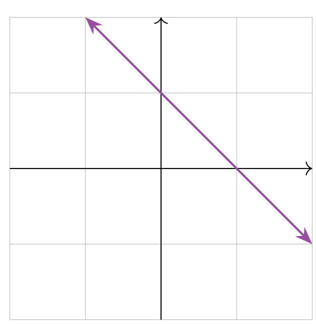
The QR code for this textbook is a 29×29 *array of black/white squares.*

In the above examples, it was useful from a psychological perspective to replace a list of four numbers (representing traffic flow) or of 841 numbers (representing a QR code) by a single piece of data: a point in some $\mathbf{R}^n$. This is a powerful concept; starting in Section 2.2, we will almost exclusively record solutions of systems of linear equations in this way.

1.1.2 Pictures of Solution Sets

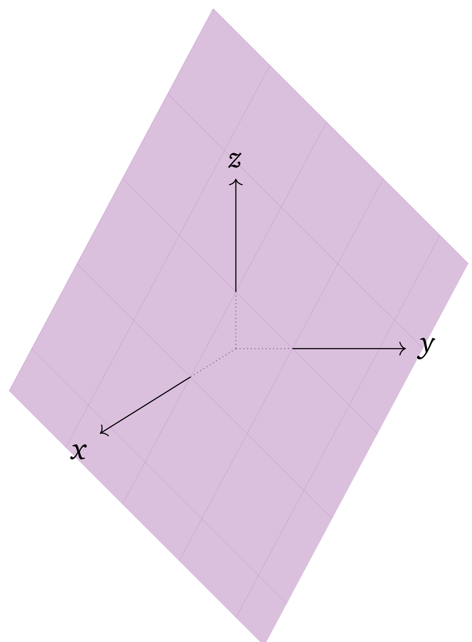
Before discussing how to solve a system of linear equations below, it is helpful to see some pictures of what these solution sets look like geometrically.

One Equation in Two Variables. Consider the linear equation $x + y = 1$. We can rewrite this as $y = 1 - x$, which defines a line in the plane: the slope is -1, and the x-intercept is 1.



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

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



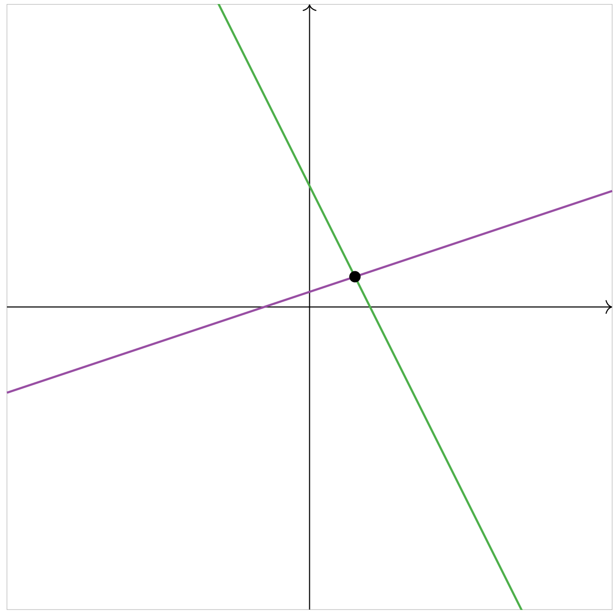
Definition (Planes)**.** A **plane** is a flat sheet that is infinite in all directions.

Remark. The equation $x + y + z + w = 1$ defines a "3-plane" in 4-space, and more generally, a single linear equation in n variables defines an "$(n-1)$-plane" in n-space. We will make these statements precise in Section 2.7.

Two Equations in Two Variables. Now consider the system of two linear equations

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

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

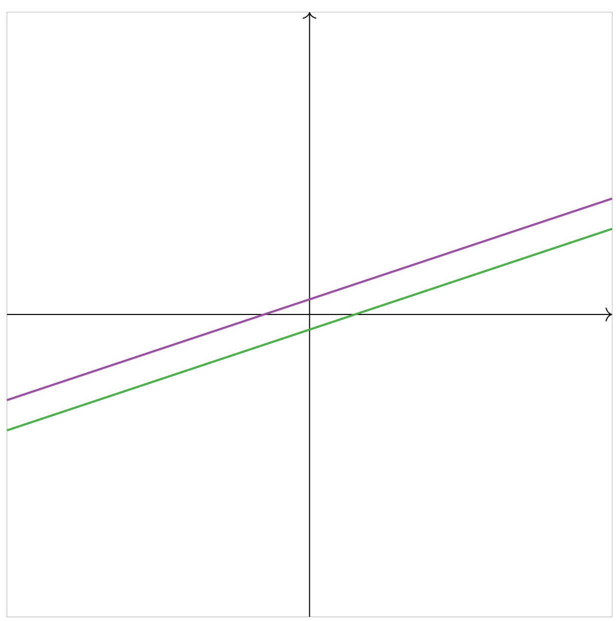


A solution to the *system* of both equations is a pair of numbers (x, y) that makes both equations true at once. In other words, it as a point that lies on both lines simultaneously. We can see in the picture above that there is only one point where the lines intersect: therefore, this system has exactly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

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

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

These define *parallel* lines in the plane.

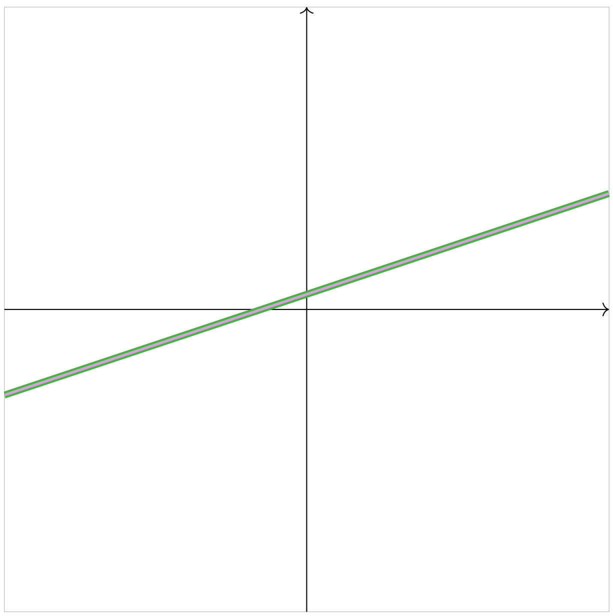


The fact that that the lines do not intersect means that the system of equations has no solution. Of course, this is easy to see algebraically: if $x-3y=-3$, then it is cannot also be the case that $x-3y=3$.

There is one more possibility. Consider the system of equations

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

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



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

Two Equations in Three Variables. Consider the system of two linear equations

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

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

Use this link to view the online demo

The planes defined by the equations $x+y+z=1$ and $x-z=0$ intersect in the red line, which is the solution set of the system of both equations.

Remark. In general, the solutions of a system of equations in n variables is the intersection of "$(n-1)$-planes" in n-space. This is always some kind of linear space, as we will discuss in Section 2.4.

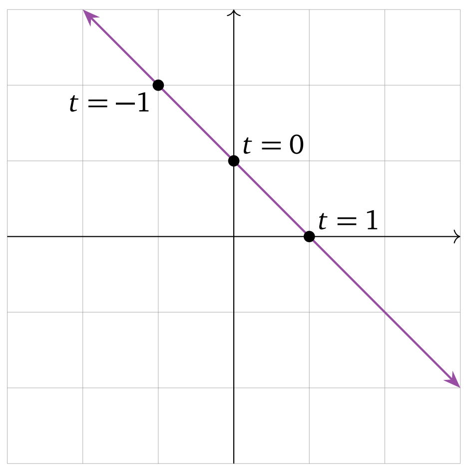
1.1.3 Parametric Description of Solution Sets

According to this definition, solving a system of equations means writing down all solutions in terms of some number of parameters. We will give a systematic way of doing so in Section 1.3; for now we give parametric descriptions in the examples of the previous subsection.

Lines. Consider the linear equation $x + y = 1$ of this example. In this context, we call $x + y = 1$ an **implicit equation** of the line. We can write the same line in **parametric form** as follows:

$$(x, y) = (t, 1 - t) \quad \text{for any} \quad t \in \mathbf{R}.$$

This means that every point on the line has the form $(t, 1-t)$ for some real number t. In this case, we call t a **parameter**, as it *parameterizes* the points on the line.



Now consider the system of two linear equations

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

of this example. These collectively form the **implicit equations** for a line in $\mathbf{R}^3$. (At least two equations are needed to define a line in space.) This line also has a **parametric form** with one **parameter** t:

$$(x, y, z) = (t, 1 - 2t, t).$$

Use this link to view the online demo

The planes defined by the equations $x + y + z = 1$ and $x - z = 0$ intersect in the yellow line, which is parameterized by $(x, y, z) = (t, 1 - 2t, t)$. Move the slider to change the parameterized point.

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

Planes. Consider the linear equation $x+y+z=1$ of this example. This is an **implicit equation** of a plane in space. This plane has an equation in **parametric form**: we can write every point on the plane as

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

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

Use this link to view the online demo

The plane in $\mathbf{R}^3$ defined by the equation $x+y+z=1$. This plane is parameterized by two numbers t,w; move the sliders to change the parameterized point.

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

When there is a unique solution, as in this example, it is not necessary to use parameters to describe the solution set.

1.2 Row Reduction

Objectives

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

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

1.2.1 The Elimination Method

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

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

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

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

$$\left\{\begin{aligned} x + 2y + 3z &= 6 \\ 2x - 3y + 2z &= 14 \\ 3x + y - z &= -2 \end{aligned}\right. \xrightarrow{\text{2nd} \;=\; \text{2nd}-2\times\text{1st}} \left\{\begin{aligned} x + 2y + 3z &= 6 \\ -7y - 4z &= 2 \\ 3x + y - z &= -2 \end{aligned}\right.$$

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

$$\left\{\begin{aligned} x + 2y + 3z &= 6 \\ 2x - 3y + 2z &= 14 \\ 3x + y - z &= -2 \end{aligned}\right. \xrightarrow{\text{3rd} \longleftrightarrow \text{1st}} \left\{\begin{aligned} 3x + y - z &= -2 \\ 2x - 3y + 2z &= 14 \\ x + 2y + 3z &= 6 \end{aligned}\right.$$

Example. Solve (1.1.1) using the elimination method.

Solution.

$$\left\{\begin{aligned} x + 2y + 3z &= 6 \\ 2x - 3y + 2z &= 14 \\ 3x + y - z &= -2 \end{aligned}\right. \xrightarrow{\text{2nd} \;=\; \text{2nd}-2\times\text{1st}} \left\{\begin{aligned} x + 2y + 3z &= 6 \\ -7y - 4z &= 2 \\ 3x + y - z &= -2 \end{aligned}\right.$$

$$\xrightarrow{\text{3rd} \;=\; \text{3rd}-3\times\text{1st}} \left\{\begin{aligned} x + 2y + 3z &= 6 \\ -7y - 4z &= 2 \\ -5y - 10z &= -20 \end{aligned}\right.$$

$$\xrightarrow{\text{2nd} \longleftrightarrow \text{3rd}} \left\{\begin{aligned} x + 2y + 3z &= 6 \\ -5y - 10z &= -20 \\ -7y - 4z &= 2 \end{aligned}\right.$$

$$\xrightarrow{\text{divide 2nd by } -5} \left\{\begin{aligned} x + 2y + 3z &= 6 \\ y + 2z &= 4 \\ -7y - 4z &= 2 \end{aligned}\right.$$

$$\xrightarrow{\text{3rd} \;=\; \text{3rd}+7\times\text{2nd}} \left\{\begin{aligned} x + 2y + 3z &= 6 \\ y + 2z &= 4 \\ 10z &= 30 \end{aligned}\right.$$

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

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

$$\left\{\begin{array}{rrrrl} x & +2y & +3z & = & 6 \\ 2x & -3y & +2z & = & 14 \\ 3x & +\;y & -\;z & = & -2 \end{array}\right. \xrightarrow{\text{substitute}} \left\{\begin{array}{rrrl} 1 & +2\cdot(-2) & +3\cdot 3 & = 6 \\ 2\cdot 1 & -3\cdot(-2) & +2\cdot 3 & = 14 \\ 3\cdot 1 & +\;(-2) & -\;3 & = -2. \end{array}\right.$$

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

$$\left\{\begin{array}{rrrrl} x & +2y & +3z & = & 6 \\ 2x & -3y & +2z & = & 14 \\ 3x & +\;y & -\;z & = & -2 \end{array}\right. \xrightarrow{\text{becomes}} \left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right).$$

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

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

$$\left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right) \xrightarrow{R_1 = R_1 \times -3} \left(\begin{array}{rrr|r} -3 & -6 & -9 & -18 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right)$$

  Here the notation R_1 simply means "the first row", and likewise for R_2, R_3, etc.

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

$$\left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right) \xrightarrow{R_2 = R_2 - 2\times R_1} \left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 3 & 1 & -1 & -2 \end{array}\right)$$

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

$$\left(\begin{array}{rrr|r} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right) \xrightarrow{R_1 \longleftrightarrow R_3} \left(\begin{array}{rrr|r} 3 & 1 & -1 & -2 \\ 2 & -3 & 2 & 14 \\ 1 & 2 & 3 & 6 \end{array}\right)$$

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

Example. Solve (1.1.1) using row operations.

Solution. We start by forming an augmented matrix:

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

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

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

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

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

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & -5 & -10 & -20 \end{array}\right) \xrightarrow{R_3=R_3\div -5} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & 1 & 2 & 4 \end{array}\right)$$

$$\xrightarrow{R_2\longleftrightarrow R_3} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & -7 & -4 & 2 \end{array}\right)$$

$$\xrightarrow{R_3=R_3+7R_2} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 10 & 30 \end{array}\right)$$

$$\xrightarrow{R_3=R_3\div 10} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

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

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

Hence $z = 3$; back-substituting as in this example gives $(x, y, z) = (1, -2, 3)$.

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

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

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

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

Example (An Inconsistent System)**.** Solve the following system of equations using row operations:

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

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

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

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

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

$$\xrightarrow{R_3 = R_3 - 4R_1} \left(\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{array}\right)$$

Now we clear the second entry from the last row.

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

This translates back into the system of equations

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

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

1.2.2 Echelon Forms

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

Definition. A matrix is in **row echelon form** if:

1. All zero rows are at the bottom.

2. The first nonzero entry of a row is to the *right* of the first nonzero entry of the row above.

3. Below the first nonzero entry of a row, all entries are zero.

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

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

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

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

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

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

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

5. Each pivot is equal to 1.

6. Each pivot is the only nonzero entry in its column.

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

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

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

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

Example. The following matrices are in reduced row echelon form:

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 8 & 0 \end{pmatrix} \qquad \left(\begin{array}{cc|c} 1 & 17 & 0 \\ 0 & 0 & 1 \end{array}\right) \qquad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

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

$$\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \qquad \left(\begin{array}{ccc|c} 2 & 7 & 1 & 4 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{array}\right) \qquad \left(\begin{array}{cc|c} 1 & 17 & 0 \\ 0 & 1 & 1 \end{array}\right) \qquad \begin{pmatrix} 2 & 1 & 3 \\ 0 & 0 & 0 \end{pmatrix}.$$

The following matrices are not in echelon form:

$$\left(\begin{array}{ccc|c} 2 & 7 & 1 & 4 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 3 \end{array}\right) \qquad \left(\begin{array}{cc|c} 0 & 17 & 0 \\ 0 & 2 & 1 \end{array}\right) \qquad \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

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

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

Remark (Why the word "pivot"?)**.** Consider the following system of equations:

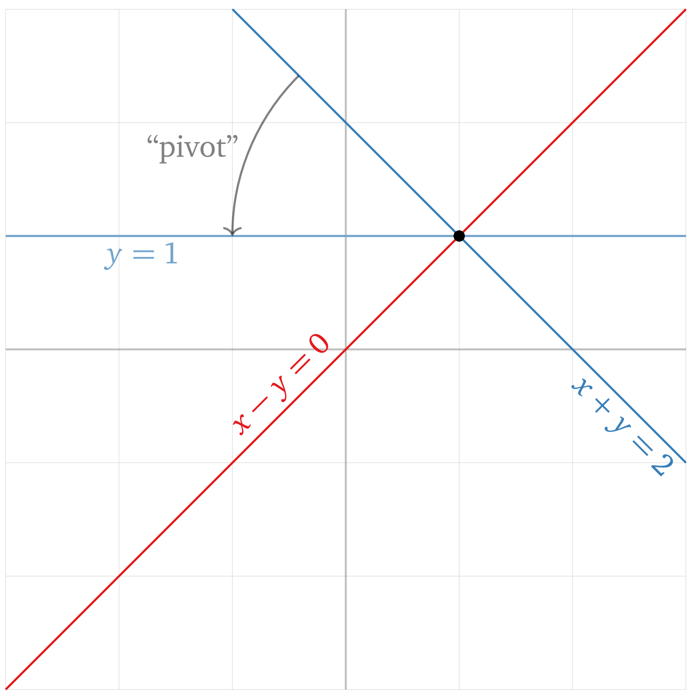
$$\begin{cases} x - y = 0 \\ x + y = 2. \end{cases}$$

We can visualize this system as a pair of lines in $\mathbf{R}^2$ (red and blue, respectively, in the picture below) that intersect at the point $(1,1)$. If we subtract the first equation from the second, we obtain the equation $2y = 2$, or $y = 1$. This results in the system of equations:

$$\begin{cases} x - y = 0 \\ \quad\;\; y = 1. \end{cases}$$

In terms of row operations on matrices, we can write this as:

$$\left(\begin{array}{cc|c} 1 & -1 & 0 \\ 1 & 1 & 2 \end{array}\right) \xrightarrow{R_2 = R_2 - R_1} \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 2 & 2 \end{array}\right)$$
$$\xrightarrow{R_2 = \frac{1}{2}R_2} \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 1 & 1 \end{array}\right)$$



What has happened geometrically is that the original blue line has been replaced with the new blue line $y = 1$. We can think of the blue line as rotating, or pivoting, around the solution $(1,1)$. We used the pivot position in the matrix in order to make the blue line pivot like this. This is one possible explanation for the terminology "pivot".

1.2.3 The Row Reduction Algorithm

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

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

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

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

We will not prove uniqueness, but maybe you can!

Algorithm (Row Reduction)**.**

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

Example. Row reduce this matrix:

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

Solution.

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

Step 1a: Row swap to make this nonzero.

$$\xrightarrow{R_1 \longleftrightarrow R_2} \left(\begin{array}{ccc|c} 2 & 4 & 6 & 12 \\ 0 & -7 & -4 & 2 \\ 3 & 1 & -1 & -2 \end{array}\right)$$

Step 1b: Scale to make this 1.

$$\xrightarrow{R_1 = R_1 \div 2} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 3 & 1 & -1 & -2 \end{array}\right)$$

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

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

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

$$\xrightarrow{R_2 \longleftrightarrow R_3} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -5 & -10 & -20 \\ 0 & -7 & -4 & 2 \end{array}\right)$$

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

$$\xrightarrow{R_2 = R_2 \div -5} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & -7 & -4 & 2 \end{array}\right)$$

Note how Step 2b doesn't create fractions.

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

$$\xrightarrow{R_3 = R_3 + 7R_2} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 10 & 30 \end{array}\right)$$

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

$$\xrightarrow{R_3 = R_3 \div 10} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

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

$$\xrightarrow{R_2 = R_2 - 2R_3} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

$$\xrightarrow{R_1 = R_1 - 3R_3} \left(\begin{array}{ccc|c} 1 & 2 & 0 & -3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

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

$$\xrightarrow{R_1 = R_1 - 2R_2} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

The reduced row echelon form of the matrix is

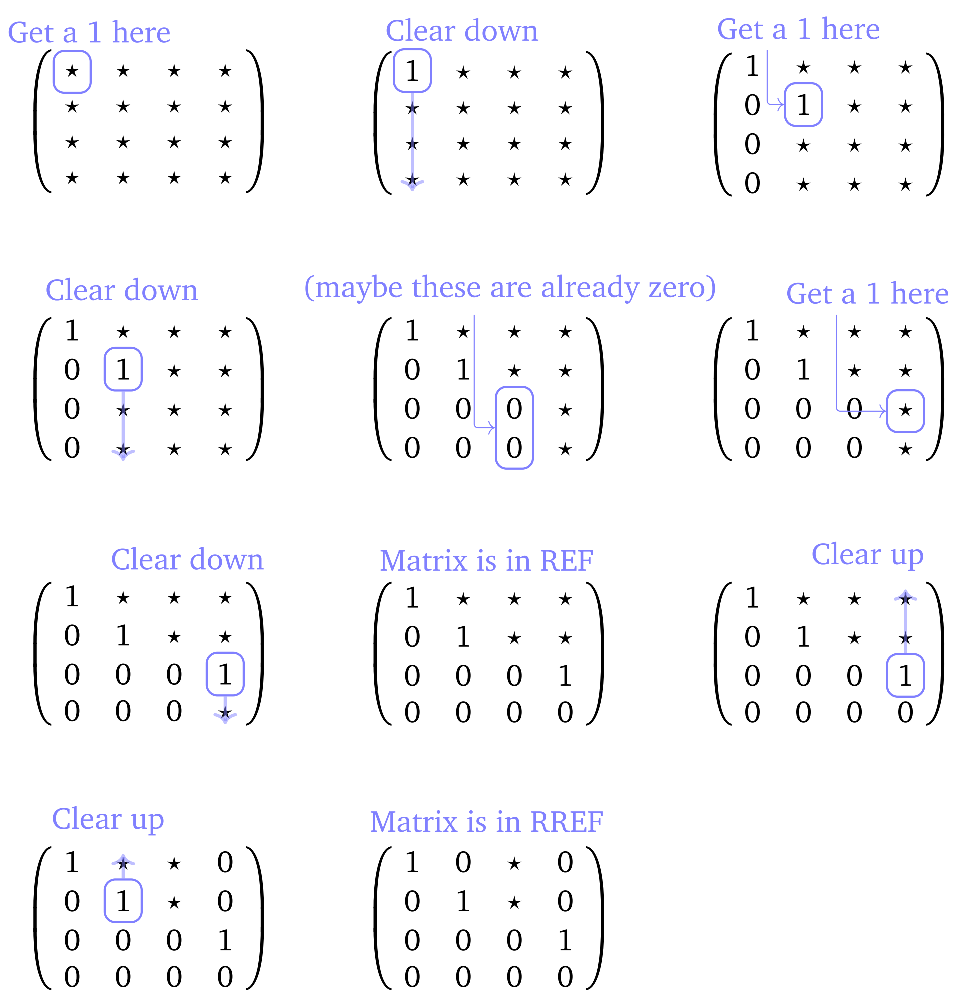
$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right) \xrightarrow{\text{translates to}} \left\{\begin{array}{rrrl} x & & & = 1 \\ & y & & = -2 \\ & & z & = 3. \end{array}\right.$$

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

Use this link to view the online demo

Animated slideshow of the row reduction in this example.

Here is the row reduction algorithm, summarized in pictures.



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

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

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

Example (Pivot Positions)**.** Find the pivot positions and pivot columns of this matrix

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

Solution. We saw in this example that a row echelon form of the matrix is

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

The pivot positions of A are the entries that become pivots in a row echelon form; they are marked in red below:

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

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

Example (An Inconsistent System)**.** Solve the linear system

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

using row reduction.

Solution.

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

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

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

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

This row reduced matrix corresponds to the *inconsistent* system

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

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

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

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

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

We have discussed two classes of matrices so far:

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

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right) \xrightarrow{\text{translates to}} \left\{\begin{array}{rrrl} x & & & = 1 \\ & y & & = -2 \\ & & z & = 3. \end{array}\right.$$

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

$$\left(\begin{array}{cc|c} 1 & 5 & 0 \\ 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{translates to}} \left\{\begin{array}{rl} x + 5y & = 0 \\ 0 & = 1. \end{array}\right.$$

What happens when one of the non-augmented columns lacks a pivot? This is the subject of Section 1.3.

Example (A System with Many Solutions)**.** Solve the linear system

$$\left\{\begin{array}{rl} 2x + y + 12z & = 1 \\ x + 2y + 9z & = -1 \end{array}\right.$$

using row reduction.

Solution.

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

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

$$\xrightarrow{R_2 = R_2 \div -3} \left(\begin{array}{ccc|c} 1 & 2 & 9 & -1 \\ 0 & 1 & 2 & -1 \end{array}\right) \qquad \text{(Step 2b)}$$

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

This row reduced matrix corresponds to the linear system

$$\left\{\begin{array}{rl} x + 5z & = 1 \\ y + 2z & = -1. \end{array}\right.$$

In what sense is the system solved? We will see in Section 1.3.

1.3 Parametric Form

Objectives

1. Learn to express the solution set of a system of linear equations in parametric form.
2. Understand the three possibilities for the number of solutions of a system of linear equations.
3. *Recipe:* parametric form.
4. *Vocabulary word:* **free variable**.

1.3.1 Free Variables

There is one possibility for the row reduced form of a matrix that we did not see in Section 1.2.

Example (A System with a Free Variable)**.** Consider the linear system

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

We solve it using row reduction:

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

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

$$\xrightarrow{R_2 = R_2 \div -3} \left(\begin{array}{ccc|c} 1 & 2 & 9 & -1 \\ 0 & 1 & 2 & -1 \end{array}\right) \qquad \text{(Step 2b)}$$

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

This row reduced matrix corresponds to the linear system

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

In what sense is the system solved? We rewrite as

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

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

We have found all solutions: it is the set of all values x, y, z, where

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

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

Use this link to view the online demo

A picture of the solution set (the yellow line) of the linear system in this example. There is a unique solution for every value of z; move the slider to change z.

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

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

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

In the above example, the variable z was free because the reduced row echelon form matrix was

$$\left(\begin{array}{ccc|c} 1 & 0 & 5 & 1 \\ 0 & 1 & 2 & -1 \end{array}\right).$$

In the matrix

$$\left(\begin{array}{cccc|c} 1 & \star & 0 & \star & \star \\ 0 & 0 & 1 & \star & \star \end{array}\right),$$

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

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

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

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

Implicit Versus Parameterized Equations. The solution set of the system of linear equations

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

is a line in $\mathbf{R}^3$, as we saw in this example. These equations are called the **implicit equations** for the line: the line is defined *implicitly* as the simultaneous solutions to those two equations.

The parametric form

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

can be written as follows:

$$(x, y, z) = (1 - 5z, -1 - 2z, z) \qquad z \text{ any real number.}$$

This called a **parameterized equation** for the same line. It is an expression that produces all points of the line in terms of one parameter, z.

One should think of a system of equations as being an implicit equation for its solution set, and of the parametric form as being the parameterized equation for the same set. The parameteric form is much more explicit: it gives a concrete recipe for producing *all* solutions.

> You can choose *any value* for the free variables in a (consistent) linear system. Free variables come from the *columns without pivots* in a matrix in row echelon form.

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

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 3 & 2 \\ 0 & 0 & 1 & 4 & -1 \end{array}\right).$$

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

This translates into the system of equations

$$\begin{cases} x_1 + 3x_4 = 2 \\ x_3 + 4x_4 = -1 \end{cases} \xrightarrow{\text{parametric form}} \begin{cases} x_1 = 2 - 3x_4 \\ x_3 = -1 - 4x_4. \end{cases}$$

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

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

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

Example (A Parameterized Plane)**.** The system of one linear equation

$$x + y + z = 1$$

comes from the matrix

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \end{array}\right),$$

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

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

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

Use this link to view the online demo

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

1.3.2 Number of Solutions

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

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

$$\left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$$

comes from a linear system with no solutions.

2. ***Every column except the last column is a pivot column.*** In this case, the system has a *unique* solution. For example, the matrix

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

tells us that the unique solution is $(x, y, z) = (a, b, c)$.

3. ***The last column is not a pivot column, and some other column is not a pivot column either.*** In this case, the system has *infinitely many* solutions, corresponding to the infinitely many possible values of the free variable(s). For example, in the system corresponding to the matrix

$$\left(\begin{array}{cccc|c} 1 & -2 & 0 & 3 & 1 \\ 0 & 0 & 1 & 4 & -1 \end{array}\right),$$

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

Chapter 2

Systems of Linear Equations: Geometry

Primary Goals. We have already discussed systems of linear equations and how this is related to matrices. In this chapter we will learn how to write a system of linear equations succinctly as a matrix equation, which looks like $Ax = b$, where A is an $m \times n$ matrix, b is a vector in $\mathbf{R}^m$ and x is a variable vector in $\mathbf{R}^n$. As we will see, this is a powerful perspective. We will study two related questions:

1. What is the set of solutions to $Ax = b$?
2. What is the set of b so that $Ax = b$ is consistent?

The first question is the kind you are used to from your first algebra class: what is the set of solutions to $x^2 - 1 = 0$. The second is also something you could have studied in your previous algebra classes: for which b does $x^2 = b$ have a solution? This question is more subtle at first glance, but you can solve it in the same way as the first question, with the quadratic formula.

In order to answer the two questions listed above, we will use geometry. This will be analogous to how you used parabolas in order to understand the solutions to a quadratic equation in one variable. Specifically, this chapter is devoted to the geometric study of two objects:

1. the solution set of a matrix equation $Ax = b$, and
2. the set of all b that makes a particular system consistent.

The second object will be called the column space of A. The two objects are related in a beautiful way by the rank theorem in Section 2.9.

Instead of parabolas and hyperbolas, our geometric objects are subspaces, such as lines and planes. Our geometric objects will be something like 13-dimensional planes in $\mathbf{R}^{27}$, etc. It is amazing that we can say anything substantive about objects that we cannot directly visualize.

We will develop a large amount of vocabulary that we will use to describe the above objects: vectors (Section 2.1), spans (Section 2.2), linear independence

(Section 2.5), subspaces (Section 2.6), dimension (Section 2.7), coordinate systems (Section 2.8), etc. We will use these concepts to give a precise geometric description of the solution set of any system of equations (Section 2.4). We will also learn how to express systems of equations more simply using matrix equations (Section 2.3).

2.1 Vectors

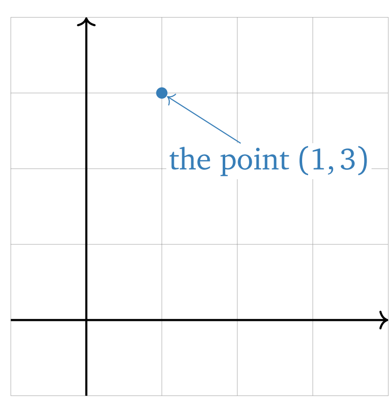
Objectives

1. Learn how to add and scale vectors in $\mathbf{R}^n$, both algebraically and geometrically.
2. Understand linear combinations geometrically.
3. *Pictures:* vector addition, vector subtraction, linear combinations.
4. *Vocabulary words:* **vector**, **linear combination**.

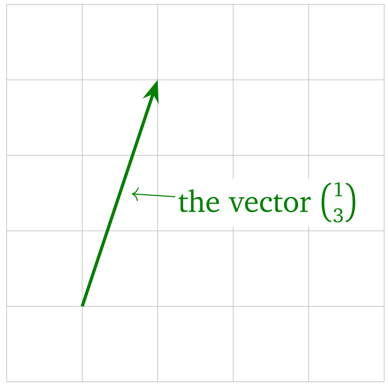
2.1.1 Vectors in $\mathbf{R}^n$

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

Points and Vectors. Again, a **point** in $\mathbf{R}^n$ is drawn as a dot.



A **vector** is a point in $\mathbf{R}^n$, drawn as an arrow.



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

Interactive: A vector in $\mathbf{R}^3$, by coordinates.

Use this link to view the online demo

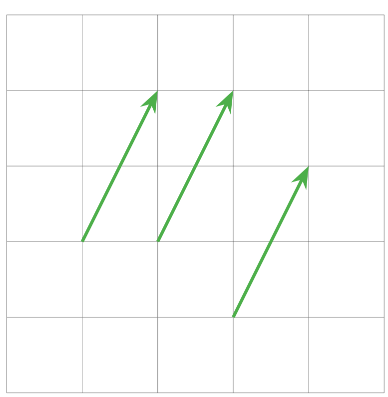
A vector in $\mathbf{R}^3$, and its coordinates. Drag the arrow head and tail.

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

$$v = \begin{pmatrix} 1 \\ 3 \end{pmatrix}.$$

We will also write 0 for the zero vector.

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

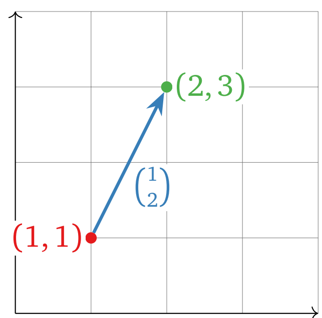


Unless otherwise specified, we will assume that all vectors start at the origin.

Vectors makes sense in the real world: many physical quantities, such as velocity, are represented as vectors. But it makes more sense to think of the velocity of a car as being located at the car.

Remark. Some authors use boldface letters to represent vectors, as in "**v**", or use arrows, as in "$\vec{v}$". As it is usually clear from context if a letter represents a vector, we do not decorate vectors in this way.

Note. Another way to think about a vector is as a *difference* between two points, or the arrow from one point to another. For instance, $\binom{1}{2}$ is the arrow from $(1,1)$ to $(2,3)$.



2.1.2 Vector Algebra and Geometry

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

Vector addition and scalar multiplication.

- We can add two vectors together:

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} + \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a+x \\ b+y \\ c+z \end{pmatrix}.$$

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

$$c\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} c \cdot x \\ c \cdot y \\ c \cdot z \end{pmatrix}.$$

  We call c a **scalar** to distinguish it from a vector. If v is a vector and c is a scalar, then cv is called a **scalar multiple** of v.

Addition and scalar multiplication work in the same way for vectors of length n.

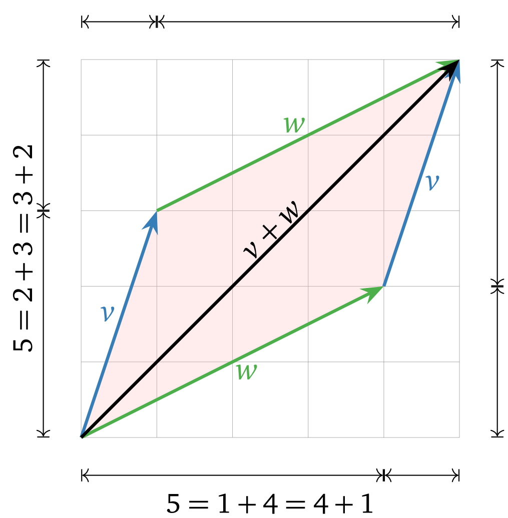
Example.

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

The Parallelogram Law for Vector Addition Geometrically, the sum of two vectors v, w is obtained as follows: place the tail of w at the head of v. Then $v+w$ is the vector whose tail is the tail of v and whose head is the head of w. Doing this both ways creates a parallelogram. For example,

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

Why? The width of $v+w$ is the sum of the widths, and likewise with the heights.



Interactive: The parallelogram law for vector addition.

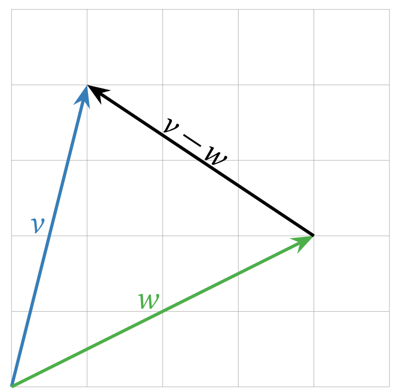
Use this link to view the online demo

The parallelogram law for vector addition. Click and drag the heads of and w.

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

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

Why? If you add $v - w$ to w, you get v.



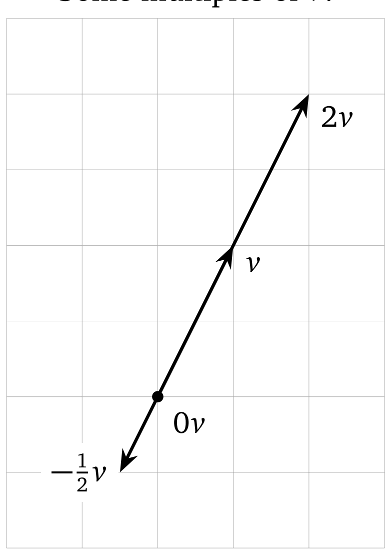
Interactive: Vector subtraction.

Use this link to view the online demo

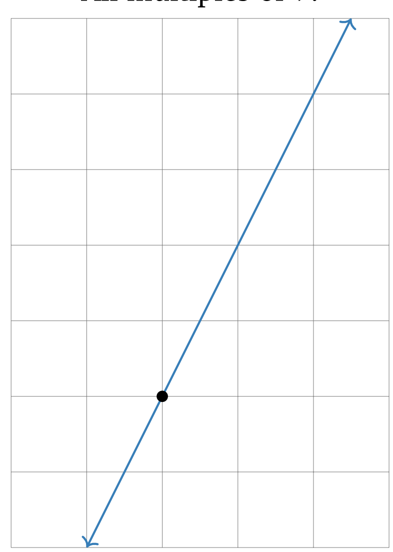
Vector subtraction. Click and drag the heads of v and w.

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

Some multiples of v.



All multiples of v.



Interactive: Scalar multiplication.

Use this link to view the online demo

Scalar multiplication. Drag the slider to change the scalar.

2.1.3 Linear Combinations

We can add and scale vectors in the same equation.

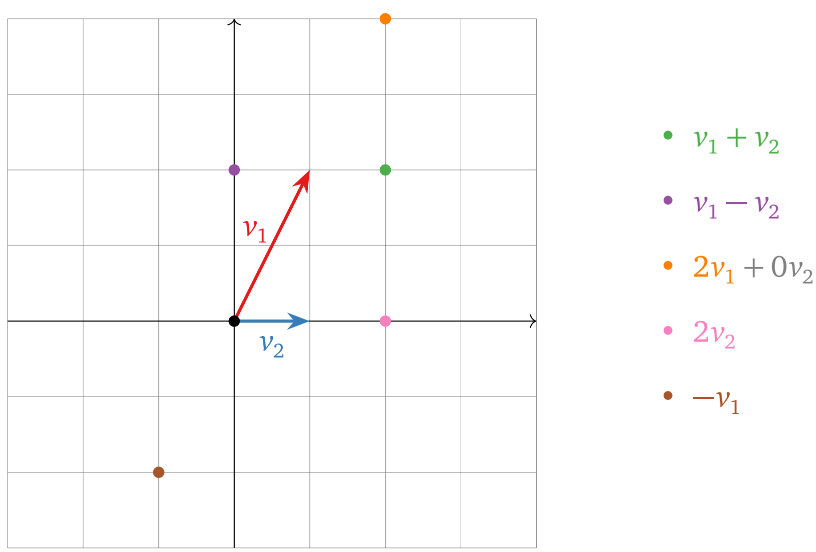
Definition. Let $c_1, c_2, \ldots, c_k$ be scalars, and let $v_1, v_2, \ldots, v_k$ be vectors in $\mathbf{R}^n$. The vector in $\mathbf{R}^n$

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k$$

is called a **linear combination** of the vectors $v_1, v_2, \ldots, v_k$, with **weights** or **coefficients** $c_1, c_2, \ldots, c_k$.

Geometrically, a linear combination is obtained by stretching / shrinking the vectors $v_1, v_2, \ldots, v_k$ according to the coefficients, then adding them together using the parallelogram law.

Example. Let $v_1 = \binom{1}{2}$ and $v_2 = \binom{1}{0}$. Here are some linear combinations of v_1 and v_2, drawn as points.



The locations of these points are found using the parallelogram law for vector addition. Any vector on the plane is a linear combination of v_1 and v_2, with suitable coefficients.

Use this link to view the online demo

Linear combinations of two vectors in $\mathbf{R}^2$*: move the sliders to change the coefficients of* v_1 *and* v_2*. Note that any vector on the plane can be obtained as a linear combination of* v_1, v_2 *with suitable coefficients.*

Interactive: Linear combinations of three vectors.

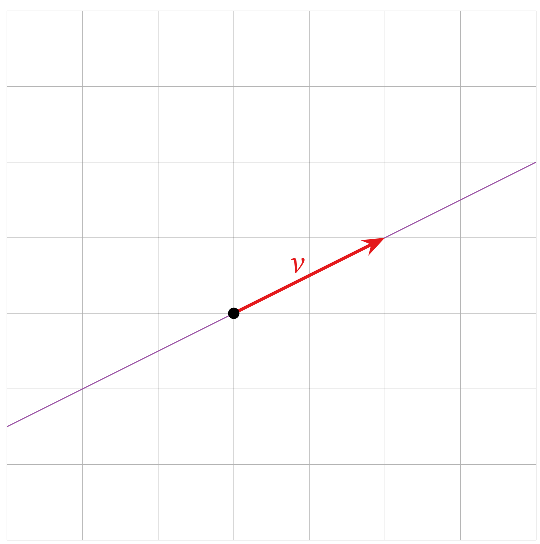
Use this link to view the online demo

Linear combinations of three vectors: move the sliders to change the coefficients of v_1, v_2, v_3*. Note how the parallelogram law for addition of three vectors is more of a "parallepiped law".*

Example (Linear Combinations of a Single Vector)**.** A linear combination of a single vector $v = \binom{1}{2}$ is just a scalar multiple of v. So some examples include

$$v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad \frac{3}{2}v = \begin{pmatrix} 3/2 \\ 3 \end{pmatrix}, \quad -\frac{1}{2}v = \begin{pmatrix} -1/2 \\ -1 \end{pmatrix}, \quad \ldots$$

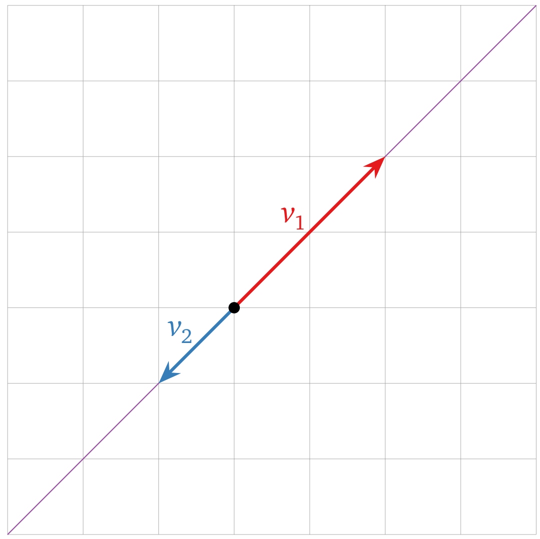
The set of all linear combinations is the *line through* v. (Unless $v = 0$, in which case any scalar multiple of v is again 0.)



Example (Linear Combinations of Collinear Vectors)**.** The set of all linear combinations of the vectors

$$v_1 = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

is the *line* containing both vectors.



The difference between this and a previous example is that both vectors lie on the same line. Hence any scalar multiples of v_1, v_2 lie on that line, as does their sum.

Interactive: Linear combinations of two collinear vectors.

Use this link to view the online demo

Linear combinations of two collinear vectors in $\mathbf{R}^2$*. Move the sliders to change the coefficients of* v_1, v_2*. Note that there is no way to "escape" the line.*

2.2 Vector Equations and Spans

Objectives

1. Understand the equivalence between a system of linear equations and a vector equation.
2. Learn the definition of Span$\{x_1, x_2, \ldots, x_k\}$, and how to draw pictures of spans.
3. *Recipe:* solve a vector equation using augmented matrices / decide if a vector is in a span.
4. *Pictures:* an inconsistent system of equations, a consistent system of equations, spans in $\mathbf{R}^2$ and $\mathbf{R}^3$.
5. *Vocabulary word:* **vector equation**.
6. *Essential vocabulary word:* **span**.

2.2.1 Vector Equations

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

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

simplifies to

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

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

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

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

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

For example the vector equation above is asking if the vector $(8,16,3)$ is a linear combination of the vectors $(1,2,6)$ and $(-1,2,-1)$.

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

A Picture of a Consistent System. Below we will show that the above system of equations is consistent. Equivalently, this means that the above vector equation has a solution. In other words, there is a linear combination of $(1,2,6)$ and $(-1,2,-1)$ that equals $(8,16,3)$. We can visualize the last statement geometrically. Therefore, the following figure gives a *picture of a consistent system of equations*. Compare with figure below, which shows a picture of an inconsistent system.

Use this link to view the online demo

A picture of the above vector equation. Try to solve the equation geometrically by moving the sliders.

In order to actually solve the vector equation

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

one has to solve the system of linear equations

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

This means forming the augmented matrix

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

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

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

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

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

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

$$-\begin{pmatrix}1\\2\\6\end{pmatrix}-9\begin{pmatrix}-1\\-2\\-1\end{pmatrix}=\begin{pmatrix}8\\16\\3\end{pmatrix}.$$

Recipe: Solving a vector equation. In general, the vector equation

$$x_1v_1 + x_2v_2 + \cdots + x_kv_k = b$$

where $v_1, v_2, \ldots, v_k$, b are vectors in $\mathbf{R}^n$ and $x_1, x_2, \ldots, x_k$ are unknown scalars, has the same solution set as the linear system with augmented matrix

$$\left(\begin{array}{cccc|c} | & | & & | & | \\ v_1 & v_2 & \cdots & v_k & b \\ | & | & & | & | \end{array}\right)$$

whose columns are the v_i's and the b's.

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

1. As a system of equations:
$$\begin{cases} 2x_1 + 3x_2 - 2x_3 = 7 \\ x_1 - x_2 - 3x_3 = 5 \end{cases}$$

2. As an augmented matrix:
$$\left(\begin{array}{ccc|c} 2 & 3 & -2 & 7 \\ 1 & -1 & -3 & 5 \end{array}\right)$$

3. As a vector equation ($x_1v_1 + x_2v_2 + \cdots + x_nv_n = b$):
$$x_1\begin{pmatrix}2\\1\end{pmatrix} + x_2\begin{pmatrix}3\\-1\end{pmatrix} + x_3\begin{pmatrix}-2\\-3\end{pmatrix} = \begin{pmatrix}7\\5\end{pmatrix}$$

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

2.2.2 Spans

It will be important to know what are *all* linear combinations of a set of vectors $v_1, v_2, \ldots, v_k$ in $\mathbf{R}^n$. In other words, we would like to understand the set of all vectors b in $\mathbf{R}^n$ such that the vector equation (in the unknowns $x_1, x_2, \ldots, x_k$)

$$x_1v_1 + x_2v_2 + \cdots + x_kv_k = b$$

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

Essential Definition. Let $v_1, v_2, \ldots, v_k$ be vectors in $\mathbf{R}^n$. The **span** of $v_1, v_2, \ldots, v_k$ is the collection of all linear combinations of $v_1, v_2, \ldots, v_k$, and is denoted $\mathrm{Span}\{v_1, v_2, \ldots, v_k\}$. In symbols:

$$\mathrm{Span}\{v_1, v_2, \ldots, v_k\} = \big\{x_1v_1 + x_2v_2 + \cdots + x_kv_k \mid x_1, x_2, \ldots, x_k \text{ in } \mathbf{R}\big\}$$

We also say that $\mathrm{Span}\{v_1, v_2, \ldots, v_k\}$ is the subset **spanned by** or **generated by** the vectors $v_1, v_2, \ldots, v_k$.

The above definition is the first of several *essential definitions* that we will see in this textbook. They are essential in that they form the essence of the subject of linear algebra: learning linear algebra means (in part) learning these definitions. All of the definitions are important, but it is essential that you learn and understand the definitions marked as such.

Set Builder Notation. The notation

$$\{x_1v_1 + x_2v_2 + \cdots + x_kv_k \mid x_1, x_2, \ldots, x_k \text{ in } \mathbf{R}\}$$

reads as: "the set of all things of the form $x_1v_1 + x_2v_2 + \cdots + x_kv_k$ such that $x_1, x_2, \ldots, x_k$ are in $\mathbf{R}$." The vertical line is "such that"; everything to the left of it is "the set of all things of this form", and everything to the right is the condition that those things must satisfy to be in the set. Specifying a set in this way is called **set builder notation**.

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

Three characterizations of consistency. Now we have three equivalent ways of making the same statement:

1. A vector b is in the span of $v_1, v_2, \ldots, v_k$.

2. The vector equation
$$x_1v_1 + x_2v_2 + \cdots + x_kv_k = b$$
has a solution.

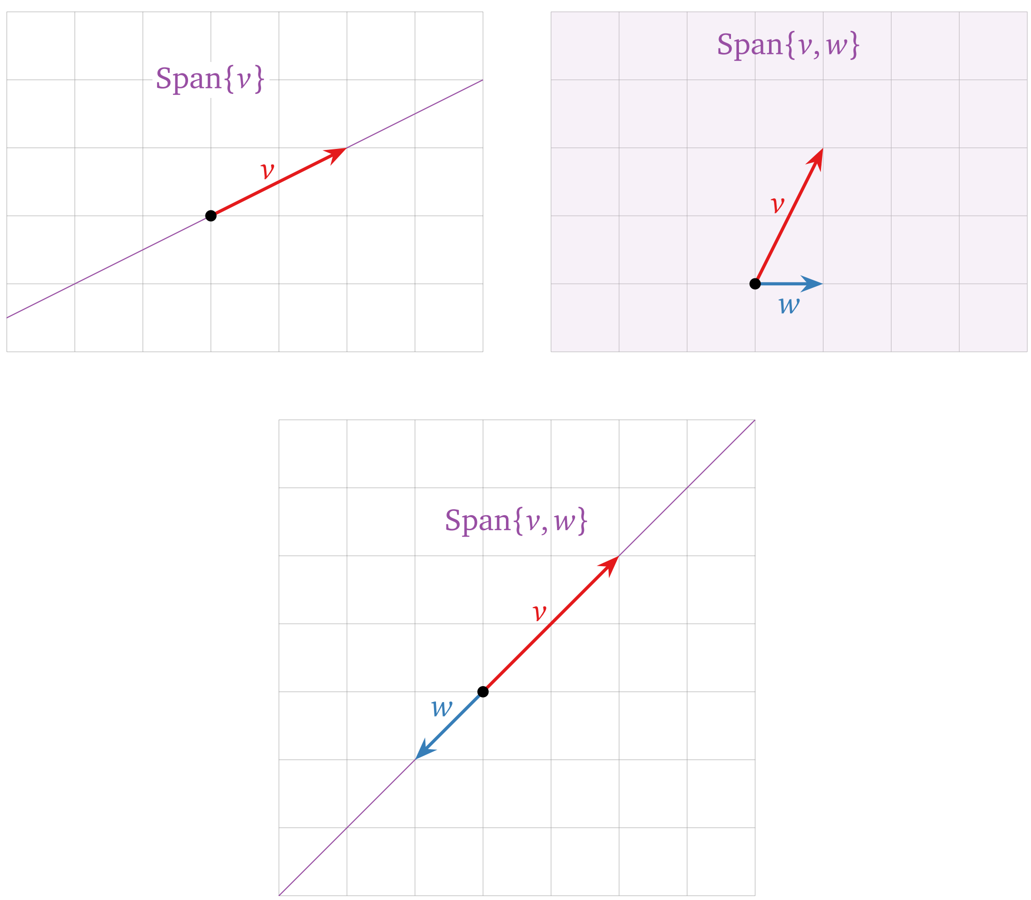
3. The linear system with augmented matrix
$$\left(\begin{array}{cccc|c} | & | & & | & | \\ v_1 & v_2 & \cdots & v_k & b \\ | & | & & | & | \end{array}\right)$$
is consistent.

Equivalent means that, for any given list of vectors $v_1, v_2, \ldots, v_k, b$, either all three statements are true, or all three statements are false.

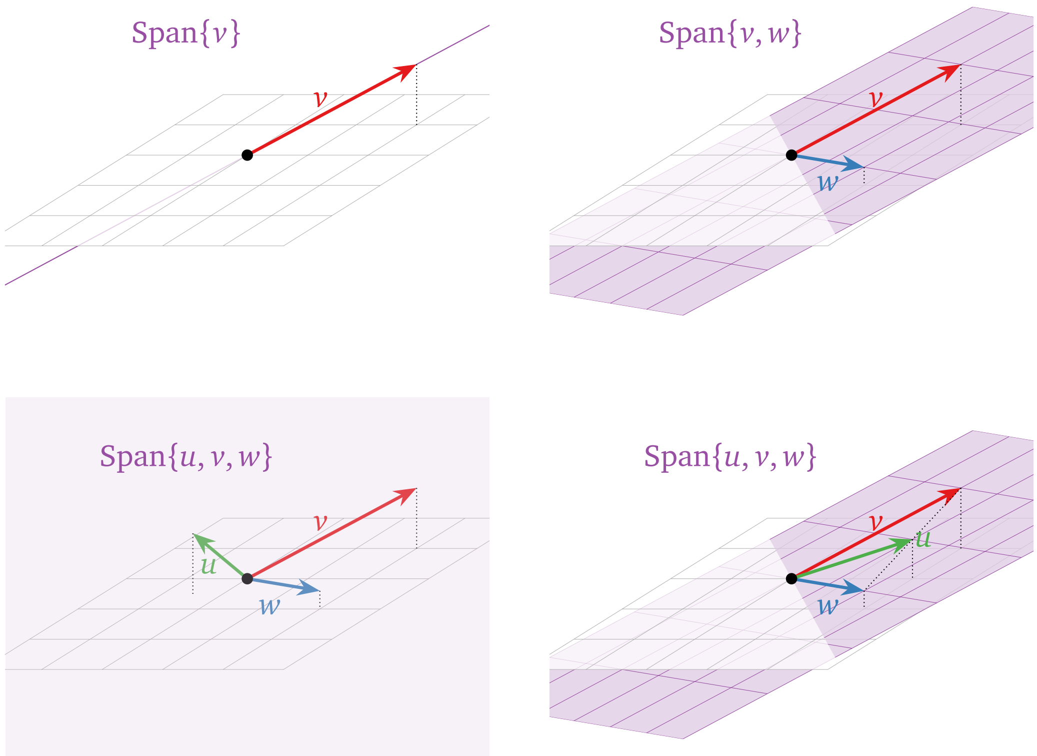
Use this link to view the online demo

This is a picture of an inconsistent *linear system: the vector w on the right-hand side of the equation $x_1v_1 + x_2v_2 = w$ is not in the span of v_1, v_2. Convince yourself of this by trying to solve the equation $x_1v_1 + x_2v_2 = w$ by moving the sliders, and by row reduction. Compare this figure.*

Pictures of Spans Drawing a picture of Span$\{v_1, v_2, \ldots, v_k\}$ is the same as drawing a picture of all linear combinations of $v_1, v_2, \ldots, v_k$.



Pictures of spans in $\mathbf{R}^2$.



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

Interactive: Span of two vectors in $\mathbf{R}^2$**.**

Use this link to view the online demo

Interactive picture of a span of two vectors in $\mathbf{R}^2$*. Check "Show* $x.v + y.w$*" and move the sliders to see how every point in the violet region is in fact a linear combination of the two vectors.*

Interactive: Span of two vectors in $\mathbf{R}^3$**.**

Use this link to view the online demo

Interactive picture of a span of two vectors in $\mathbf{R}^3$*. Check "Show* $x.v + y.w$*" and move the sliders to see how every point in the violet region is in fact a linear combination of the two vectors.*

Interactive: Span of three vectors in $\mathbf{R}^3$**.**

Use this link to view the online demo

Interactive picture of a span of three vectors in $\mathbf{R}^3$*. Check "Show x.v + y.w + z.u" and move the sliders to see how every point in the violet region is in fact a linear combination of the three vectors.*

2.3 Matrix Equations

Objectives

1. Understand the equivalence between a system of linear equations, an augmented matrix, a vector equation, and a matrix equation.
2. Characterize the vectors b such that $Ax = b$ is consistent, in terms of the span of the columns of A.
3. Characterize matrices A such that $Ax = b$ is consistent for all vectors b.
4. *Recipe:* multiply a vector by a matrix (two ways).
5. *Picture:* the set of all vectors b such that $Ax = b$ is consistent.
6. *Vocabulary word:* **matrix equation**.

2.3.1 The Matrix Equation $Ax = b$.

In this section we introduce a very concise way of writing a system of linear equations: $Ax = b$. Here A is a matrix and x, b are vectors (generally of different sizes), so first we must explain how to multiply a matrix by a vector.

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

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

Definition. Let A be an $m \times n$ matrix with columns $v_1, v_2, \ldots, v_n$:

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}$$

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

$$Ax = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n.$$

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

Example.

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

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

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

Properties of the Matrix-Vector Product. *Let A be an $m \times n$ matrix, let u, v be vectors in $\mathbf{R}^n$, and let c be a scalar. Then:*

- $A(u+v) = Au + Av$
- $A(cu) = cAu$

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

In this book we will study two complementary questions about a matrix equation $Ax = b$:

1. Given a specific choice of b, what are all of the solutions to $Ax = b$?
2. What are all of the choices of b so that $Ax = b$ is consistent?

The first question is more like the questions you might be used to from your earlier courses in algebra; you have a lot of practice solving equations like $x^2 - 1 = 0$ for x. The second question is perhaps a new concept for you. The rank theorem in Section 2.9, which is the culmination of this chapter, tells us that the two questions are intimately related.

Matrix Equations and Vector Equations. Let $v_1, v_2, \ldots, v_n$ and b be vectors in $\mathbf{R}^m$. Consider the vector equation

$$x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = b.$$

This is equivalent to the matrix equation $Ax = b$, where

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \quad \text{and} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

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

$$x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = b,$$

where $v_1, v_2, \ldots, v_n$ are the columns of A, and $x_1, x_2, \ldots, x_n$ are the entries of x.

Example. Write the vector equation

$$2v_1 + 3v_2 - 4v_3 = \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}$$

as a matrix equation, where v_1, v_2, v_3 are vectors in $\mathbf{R}^3$.

Solution. Let A be the matrix with columns v_1, v_2, v_3, and let x be the vector with entries $2, 3, -4$. Then

$$Ax = \begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix} = 2v_1 + 3v_2 - 4v_3,$$

so the vector equation is equivalent to the matrix equation $Ax = \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}$.

Four Ways of Writing a Linear System. We now have *four* equivalent ways of writing (and thinking about) a system of linear equations:

1. As a system of equations:

$$\left\{ \begin{aligned} 2x_1 + 3x_2 - 2x_3 &= 7 \\ x_1 - x_2 - 3x_3 &= 5 \end{aligned} \right.$$

2. As an augmented matrix:

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

3. As a vector equation ($x_1v_1 + x_2v_2 + \cdots + x_nv_n = b$):

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

4. As a matrix equation ($Ax = b$):

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

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

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

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

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

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

This is a scalar.

Recipe: The row-column rule for matrix-vector multiplication. If A is an $m \times n$ matrix with rows $r_1, r_2, \ldots, r_m$, and x is a vector in $\mathbf{R}^n$, then

$$Ax = \begin{pmatrix}—r_1—\\—r_2—\\\vdots\\—r_m—\end{pmatrix}x = \begin{pmatrix}r_1x\\r_2x\\\vdots\\r_mx\end{pmatrix}.$$

Example.

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

This is the same answer as before:

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

2.3.2 Spans and Consistency

Let A be a matrix with columns $v_1, v_2, \ldots, v_n$:

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}.$$

Then

$Ax = b$ has a solution

$$\begin{aligned} &\iff \text{there exist } x_1, x_2, \ldots, x_n \text{ such that } A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = b \\ &\iff \text{there exist } x_1, x_2, \ldots, x_n \text{ such that } x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = b \\ &\iff b \text{ is a linear combination of } v_1, v_2, \ldots, v_n \\ &\iff b \text{ is in the span of the columns of } A. \end{aligned}$$

Spans and Consistency. The matrix equation $Ax = b$ has a solution if and only if b is in the span of the columns of A.

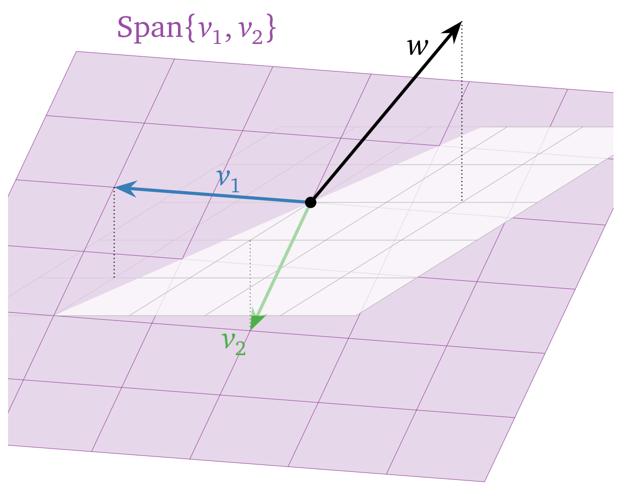
This gives an equivalence between an *algebraic* statement ($Ax = b$ is consistent), and a *geometric* statement (b is in the span of the columns of A).

Example (An Inconsistent System)**.** Let $A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 1 & -1 \end{pmatrix}$. Does the equation $Ax = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$ have a solution?

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

$$v_1 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

and the target vector (on the right-hand side of the equation) is $w = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$. The equation $Ax = w$ is consistent if and only if w is contained in the span of the columns of A. So we draw a picture:



It does not appear that w lies in $\text{Span}\{v_1, v_2\}$, so the equation is inconsistent.

Use this link to view the online demo

The vector w is not contained in $\text{Span}\{v_1, v_2\}$*, so the equation $Ax = b$ is inconsistent. (Try moving the sliders to solve the equation.)*

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

$$\left(\begin{array}{cc|c} 2 & 1 & 0 \\ -1 & 0 & 2 \\ 1 & -1 & 2 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right).$$

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

$$\begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 1 & -1 \end{pmatrix} x = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$$

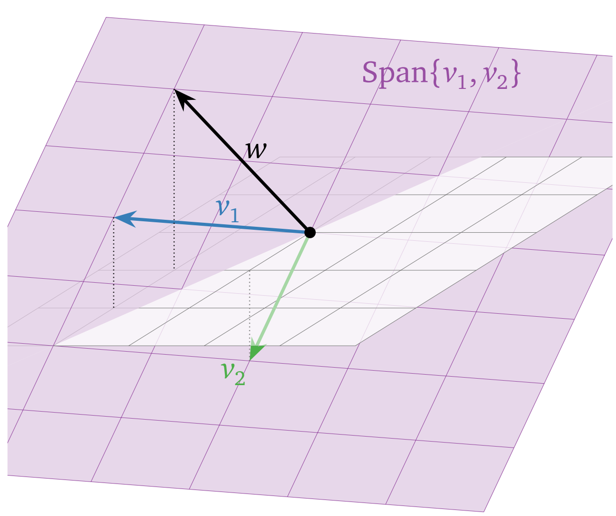
has no solution.

Example (A Consistent System). Let $A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \\ 1 & -1 \end{pmatrix}$. Does the equation $Ax = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ have a solution?

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

$$v_1 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

and the target vector (on the right-hand side of the equation) is $w = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$. The equation $Ax = w$ is consistent if and only if w is contained in the span of the columns of A. So we draw a picture:



It appears that w is indeed contained in the span of the columns of A; in fact, we can see

$$w = v_1 - v_2 \implies x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Use this link to view the online demo

The vector w is contained in $\mathrm{Span}\{v_1, v_2\}$, *so the equation $Ax = b$ is consistent. (Move the sliders to solve the equation.)*

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

$$\left(\begin{array}{cc|c} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 2 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right).$$

This gives us $x = 1$ and $y = -1$, which is consistent with the picture:

$$1\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} - 1\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \quad \text{or} \quad A\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

When Solutions Always Exist Building on this note, we have the following criterion for when $Ax = b$ is consistent for *every* choice of b.

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

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

Proof. The equivalence of 1 and 2 is established by this note as applied to every b in $\mathbf{R}^m$.

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

$$\begin{pmatrix} 1 & 0 & \star & 0 & \star \\ 0 & 1 & \star & 0 & \star \\ 0 & 0 & 0 & 1 & \star \end{pmatrix},$$

and therefore $\left(A \mid b\right)$ reduces to this:

$$\left(\begin{array}{ccccc|c} 1 & 0 & \star & 0 & \star & \star \\ 0 & 1 & \star & 0 & \star & \star \\ 0 & 0 & 0 & 1 & \star & \star \end{array}\right).$$

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

$$\begin{pmatrix} 1 & 0 & \star & 0 & \star \\ 0 & 1 & \star & 0 & \star \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

which can give rise to an inconsistent system after augmenting with b:

$$\left(\begin{array}{ccccc|c} 1 & 0 & \star & 0 & \star & 0 \\ 0 & 1 & \star & 0 & \star & 0 \\ 0 & 0 & 0 & 0 & 0 & 16 \end{array}\right). \qquad \square$$

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

Be careful when reading the statement of the above theorem. The first two conditions look very much like this note, but they are logically quite different because of the quantifier "*for all b*".

Interactive: The criteria of the theorem are satisfied.

Use this link to view the online demo

An example where the criteria of the above theorem are satisfied. The violet region is the span of the columns v_1, v_2, v_3 of A, which is the same as the set of all b such that $Ax = b$ has a solution. If you drag b, the demo will solve $Ax = b$ for you and move x.

Interactive: The critera of the theorem are not satisfied.

Use this link to view the online demo

An example where the criteria of the above theorem are not *satisfied. The violet line is the span of the columns v_1, v_2, v_3 of A, which is the same as the set of all b such that $Ax = b$ has a solution. Try dragging b in and out of the column span.*

Remark. We will see in this corollary in Section 2.7 that the dimension of the span of the columns is equal to the number of pivots of A. That is, the columns of A span a line if A has one pivot, they span a plane if A has two pivots, etc. The whole space $\mathbf{R}^m$ has dimension m, so this generalizes the fact that the columns of A span $\mathbf{R}^m$ when A has m pivots.

2.4 Solution Sets

Objectives

1. Understand the relationship between the solution set of $Ax = 0$ and the solution set of $Ax = b$.
2. Understand the difference between the solution set and the column span.
3. *Recipes:* parametric vector form, write the solution set of a homogeneous system as a span.
4. *Pictures:* solution set of a homogeneous system, solution set of an inhomogeneous system, the relationship between the two.
5. *Vocabulary words:* **homogeneous/inhomogeneous**, **trivial solution**.

In this section we will study the geometry of the solution set of any matrix equation $Ax = b$.

2.4.1 Homogeneous Systems

The equation $Ax = b$ is easier to solve when $b = 0$, so we start with this case.

Definition. A system of linear equations of the form $Ax = 0$ is called **homogeneous**.

A system of linear equations of the form $Ax = b$ for $b \neq 0$ is called **inhomogeneous**.

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

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

Observation. The equation $Ax = 0$ has a nontrivial solution $\iff$ there is a free variable $\iff$ A has a column without a pivot position.

Example (No nontrivial solutions)**.** What is the solution set of $Ax = 0$, where

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

Solution. We form an augmented matrix and row reduce:

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

The only solution is the trivial solution $x = 0$.

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

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

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

When the homogeneous equation $Ax = 0$ does have nontrivial solutions, it turns out that the solution set can be conveniently expressed as a span.

Parametric Vector Form (homogeneous case). Consider the following matrix in reduced row echelon form:

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

The matrix equation $Ax = 0$ corresponds to the system of equations

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

We can write the parametric form as follows:

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

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

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

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

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

Here is the general procedure.

Recipe: Parametric vector form (homogeneous case). Let A be an $m \times n$ matrix. Suppose that the free variables in the homogeneous equation $Ax = 0$ are, for example, x_3, x_6, and x_8.

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

The solutions to $Ax = 0$ will then be expressed in the form

$$x = x_3 v_3 + x_6 v_6 + x_8 v_8$$

for some vectors v_3, v_6, v_8 in $\mathbf{R}^n$, and any scalars x_3, x_6, x_8. This is called the **parametric vector form** of the solution.
In this case, the solution set can be written as Span$\{v_3, v_6, v_8\}$.

We emphasize the following fact in particular.

The set of solutions to a homogeneous equation $Ax = 0$ is a span.

Example (The solution set is a line)**.** Compute the parametric vector form of the solution set of $Ax = 0$, where

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

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

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

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

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

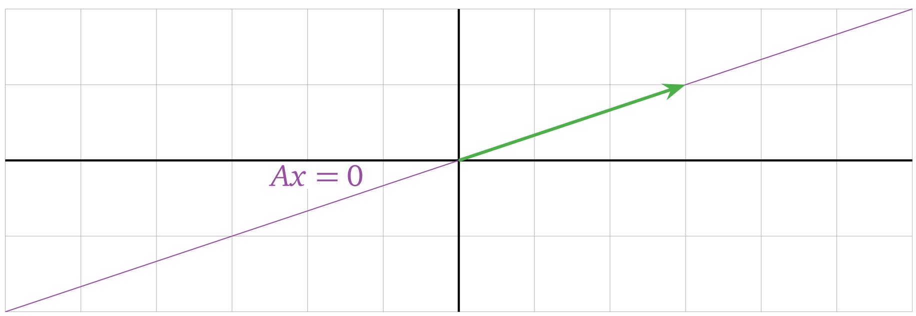
We turn these into a single vector equation:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

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

$$\mathrm{Span}\left\{ \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\}.$$

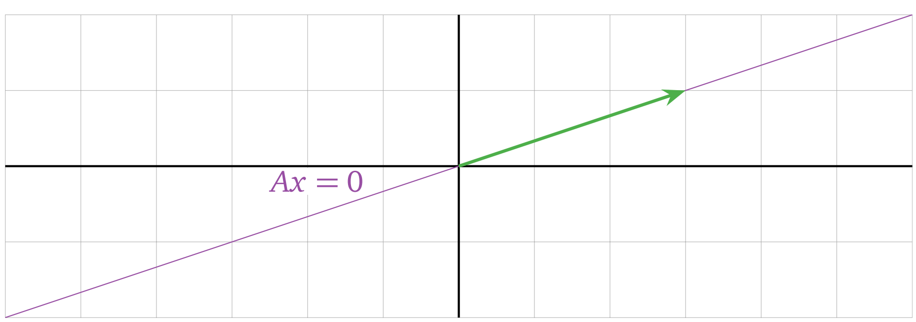
We know how to draw the picture of a span of a vector: it is a line. Therefore, this is a picture of the the solution set:



Use this link to view the online demo

Interactive picture of the solution set of $Ax = 0$. If you drag x along the line spanned by $\binom{3}{1}$, the product Ax is always equal to zero. This is what it means for $\mathrm{Span}\{\binom{3}{1}\}$ *to be the solution set of $Ax = 0$.*

Since there were *two* variables in the above example, the solution set is a subset of $\mathbf{R}^2$. Since *one* of the variables was free, the solution set is a *line*:



In order to actually *find* a nontrivial solution to $Ax = 0$ in the above example, it suffices to substitute any nonzero value for the free variable x_2. For instance, taking $x_2 = 1$ gives the nontrivial solution $x = 1 \cdot \binom{3}{1} = \binom{3}{1}$. Compare to this important note in Section 1.3.

Example (The solution set is a plane)**.** Compute the parametric vector form of the solution set of $Ax = 0$, where

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

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

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

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

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

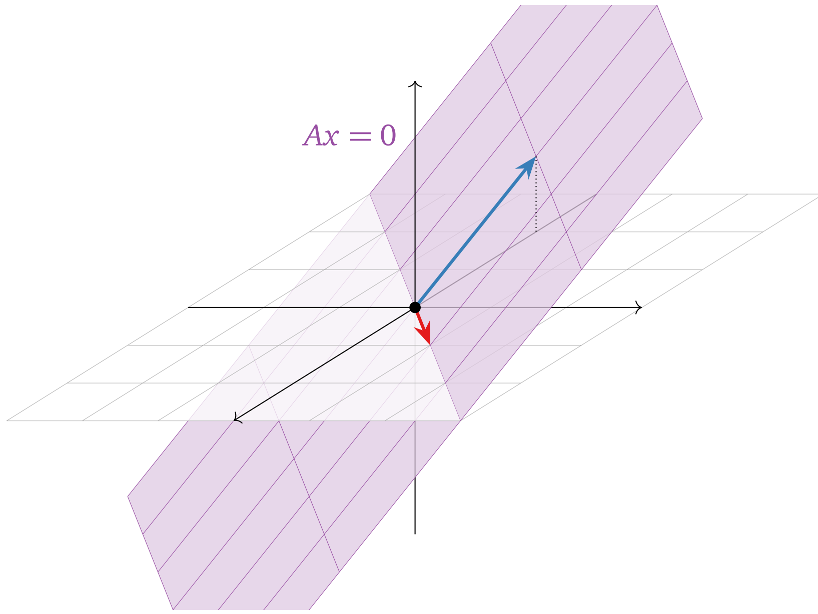
We turn these into a single vector equation:

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

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

$$\text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

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



Use this link to view the online demo

Interactive picture of the solution set of $Ax = 0$*. If you drag* x *along the violet plane, the product* Ax *is always equal to zero. This is what it means for the plane to be the solution set of* $Ax = 0$*.*

Since there were *three* variables in the above example, the solution set is a subset of $\mathbf{R}^3$. Since *two* of the variables were free, the solution set is a *plane*.

There is a natural question to ask here: is it possible to write the solution to a homogeneous matrix equation using fewer vectors than the one given in the above recipe? We will see in example in Section 2.5 that the answer is *no*: the vectors from the recipe are always linearly independent, which means that there is no way to write the solution with fewer vectors.

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

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

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

We will develop a rigorous definition of dimension in Section 2.7, but for now the dimension will simply mean the number of free variables. Compare with this important note in Section 2.5.

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

2.4.2 Inhomogeneous Systems

Recall that a matrix equation $Ax = b$ is called **inhomogeneous** when $b \neq 0$.

Example (The solution set is a line)**.** What is the solution set of $Ax = b$, where

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

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

Solution. We row reduce the associated augmented matrix:

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

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

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

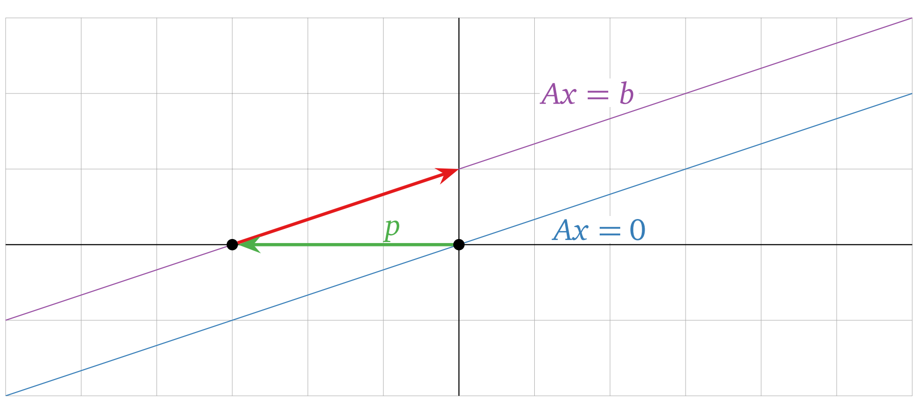
We turn the above system into a *vector equation*:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \end{pmatrix}.$$

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

$$\text{Span}\left\{\begin{pmatrix} 3 \\ 1 \end{pmatrix}\right\} + \begin{pmatrix} -3 \\ 0 \end{pmatrix}.$$

Here is a picture of the the solution set:



Use this link to view the online demo

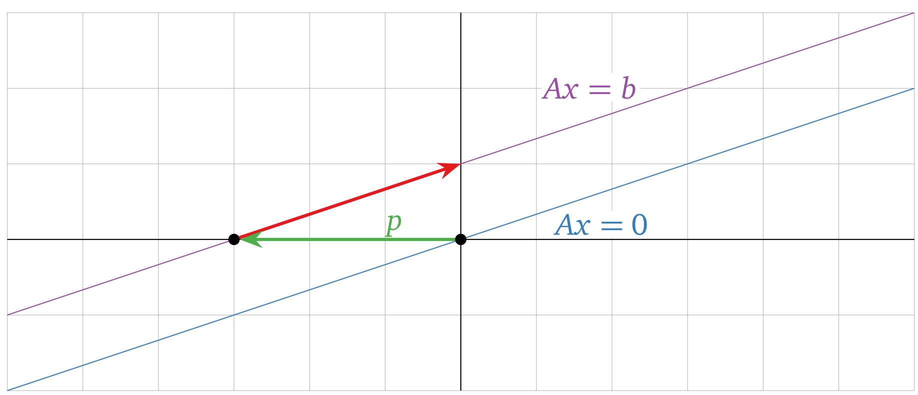
Interactive picture of the solution set of $Ax = b$. If you drag x along the violet line, the product Ax is always equal to b. This is what it means for the line to be the solution set of $Ax = b$.

In the above example, the solution set was all vectors of the form

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

where x_2 is any scalar. The vector $p = \binom{-3}{0}$ is also a solution of $Ax = b$: take $x_2 = 0$. We call p a **particular solution**.

In the solution set, x_2 is allowed to be anything, and so the solution set is obtained as follows: we take all scalar multiples of $\binom{3}{1}$ and then add the particular solution $p = \binom{-3}{0}$ to each of these scalar multiples. Geometrically, this is accomplished by first drawing the span of $\binom{3}{1}$, which is a line through the origin (and, not coincidentally, the solution to $Ax = 0$), and we *translate*, or push, this line along $p = \binom{-3}{0}$. The translated line contains p and is parallel to $\text{Span}\{\binom{3}{1}\}$: it is a *translate of a line*.



Example (The solution set is a plane). What is the solution set of $Ax = b$, where

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

(Compare this example.)

Solution. We row reduce the associated augmented matrix:

$$\left(\begin{array}{ccc|c} 1 & -1 & 2 & 1 \\ -2 & 2 & -4 & -2 \end{array}\right) \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

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

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

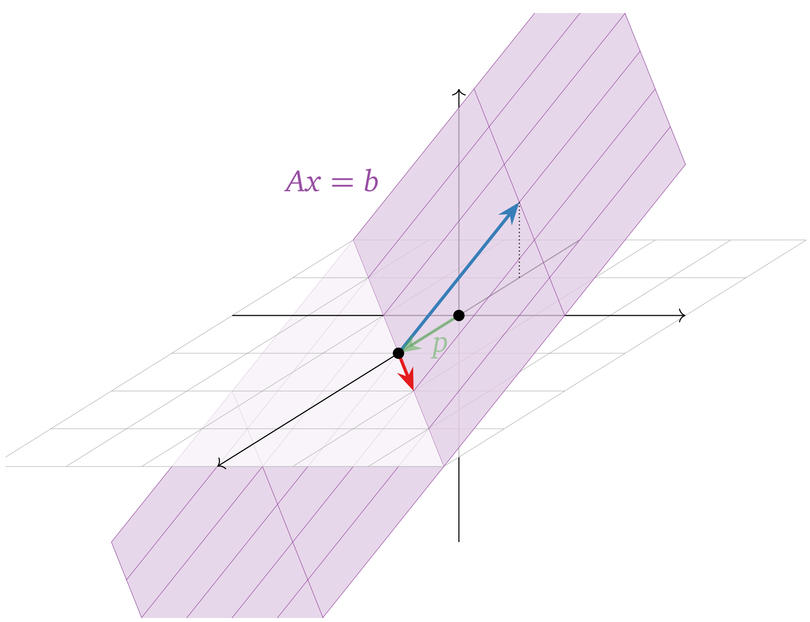
We turn the above system into a *vector equation*:

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

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

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

Here is a picture of the solution set:



Use this link to view the online demo

Interactive picture of the solution set of $Ax = b$. If you drag x along the violet plane, the product Ax is always equal to b. This is what it means for the plane to be the solution set of $Ax = b$.

In the above example, the solution set was all vectors of the form

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

where x_2 and x_3 are any scalars. In this case, a particular solution is $p = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

In the previous example and the example before it, the parametric vector form of the solution set of $Ax = b$ was exactly the same as the parametric vector form of the solution set of $Ax = 0$ (from this example and this example, respectively), plus a particular solution.

Key Observation. If $Ax = b$ is consistent, the set of solutions to is obtained by taking one **particular solution** p of $Ax = b$, and adding all solutions of $Ax = 0$.
In particular, if $Ax = b$ is consistent, the solution set is a *translate of a span*.

The **parametric vector form** of the solutions of $Ax = b$ is just the parametric vector form of the solutions of $Ax = 0$, plus a particular solution p.

It is not hard to see why the key observation is true. If p is a particular solution, then $Ap = b$, and if x is a solution to the homogeneous equation $Ax = 0$, then

$$A(x + p) = Ax + Ap = 0 + b = b,$$

so $x + p$ is another solution of $Ax = b$. On the other hand, if we start with any solution x to $Ax = b$ then $x - p$ is a solution to $Ax = 0$ since

$$A(x - p) = Ax - Ap = b - b = 0.$$

Remark. Row reducing to find the parametric vector form will give you one particular solution p of $Ax = b$. But the key observation is true for any solution p. In other words, if we row reduce in a different way and find a different solution p' to $Ax = b$ then the solutions to $Ax = b$ can be obtained from the solutions to $Ax = 0$ by either adding p or by adding p'.

Example (The solution set is a point)**.** What is the solution set of $Ax = b$, where

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

Solution. We form an augmented matrix and row reduce:

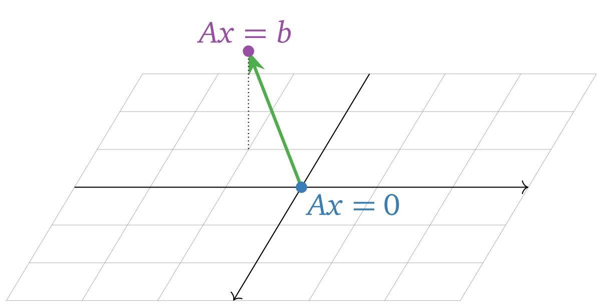
$$\left(\begin{array}{ccc|c} 1 & 3 & 4 & 0 \\ 2 & -1 & 2 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

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

According to the key observation, this is supposed to be a translate of a span by p. Indeed, we saw in the first example that the only solution of $Ax = 0$ is the trivial solution, i.e., that the solution set is the one-point set $\{0\}$. The solution set of the inhomogeneous equation $Ax = b$ is

$$\{0\} + \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

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



See the interactive figures in the next subsection for visualizations of the key observation.

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

Again compare with this important note in Section 2.5.

2.4.3 Solution Sets and Column Spans

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

- The **solution set**: for fixed b, this is the set of all x such that $Ax = b$.
  - This is a span if $b = 0$, and it is a translate of a span if $b \neq 0$ (and $Ax = b$ is consistent).
  - It is a subset of $\mathbf{R}^n$.
  - It is computed by solving a system of equations: usually by row reducing and finding the parametric vector form.
- The **span of the columns of** A: this is the set of all b such that $Ax = b$ is consistent.
  - This is always a span.
  - It is a subset of $\mathbf{R}^m$.
  - It is not computed by solving a system of equations: row reduction plays no role.

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

Interactive: Solution set and span of the columns (1).

Use this link to view the online demo

Left: the solution set of $Ax = b$ is in violet. Right: the span of the columns of A is in violet. As you move x, you change b, so the solution set changes—but all solution sets are parallel planes. If you move b within the span of the columns, the solution set also changes, and the demo solves the equation to find a particular solution x. If you move b outside of the span of the columns, the system becomes inconsistent, and the solution set disappears.

Interactive: Solution set and span of the columns (2).

Use this link to view the online demo

Left: the solution set of $Ax = b$ is in violet. Right: the span of the columns of A is in violet. As you move x, you change b, so the solution set changes—but all solution sets are parallel planes. If you move b within the span of the columns, the solution set also changes, and the demo solves the equation to find a particular solution x. If you move b outside of the span of the columns, the system becomes inconsistent, and the solution set disappears.

Interactive: Solution set and span of the columns (3).

Use this link to view the online demo

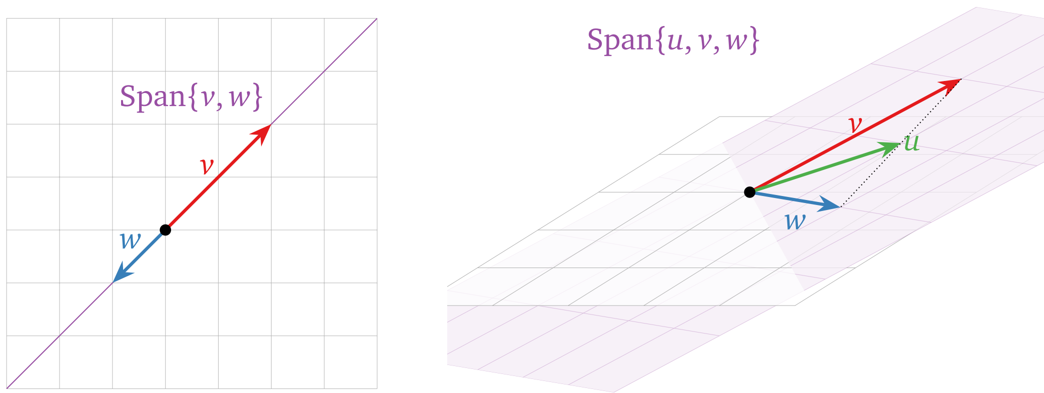
Left: the solution set of $Ax = b$ is in violet. Right: the span of the columns of A is in violet. As you move x, you change b, so the solution set changes—but all solution sets are parallel planes. If you move b within the span of the columns, the solution set also changes, and the demo solves the equation to find a particular solution x. If you move b outside of the span of the columns, the system becomes inconsistent, and the solution set disappears.

2.5 Linear Independence

Objectives

1. Understand the concept of linear independence.
2. Learn two criteria for linear independence.
3. Understand the relationship between linear independence and pivot columns / free variables.
4. *Recipe:* test if a set of vectors is linearly independent / find an equation of linear dependence.
5. *Picture:* whether a set of vectors in $\mathbf{R}^2$ or $\mathbf{R}^3$ is linearly independent or not.
6. *Vocabulary words:* **linear dependence relation** / **equation of linear dependence**.
7. *Essential vocabulary words:* **linearly independent**, **linearly dependent**.

Sometimes the span of a set of vectors is "smaller" than you expect from the number of vectors, as in the picture below. This means that (at least) one of the vectors is redundant: it can be removed without affecting the span. In the present section, we formalize this idea in the notion of *linear independence*.



Pictures of sets of vectors that are linearly dependent. Note that in each case, one vector is in the span of the others—so it doesn't make the span bigger.

2.5.1 The Definition of Linear Independence

Essential Definition. A set of vectors $\{v_1, v_2, \ldots, v_k\}$ is **linearly independent** if the vector equation

$$x_1v_1 + x_2v_2 + \cdots + x_kv_k = 0$$

has only the trivial solution $x_1 = x_2 = \cdots = x_k = 0$. The set $\{v_1, v_2, \ldots, v_k\}$ is **linearly dependent** otherwise.

In other words, $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent if there exist numbers $x_1, x_2, \ldots, x_k$, not all equal to zero, such that

$$x_1v_1 + x_2v_2 + \cdots + x_kv_k = 0.$$

This is called a **linear dependence relation** or **equation of linear dependence**.

Note that linear dependence and linear independence are notions that apply to a *collection of vectors*. It does not make sense to say things like "this vector is linearly dependent on these other vectors," or "this matrix is linearly independent."

Example (Checking linear dependence)**.** Is the set

$$\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\2\end{pmatrix}, \begin{pmatrix}3\\1\\4\end{pmatrix}\right\}$$

linearly independent?

Solution. Equivalently, we are asking if the homogeneous vector equation

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

has a nontrivial solution. We solve this by forming a matrix and row reducing (we do not augment because of this observation in Section 2.4):

$$\begin{pmatrix}1&1&3\\1&-1&1\\1&2&4\end{pmatrix}\xrightarrow{\text{row reduce}}\begin{pmatrix}1&0&2\\0&1&1\\0&0&0\end{pmatrix}$$

This says $x=-2z$ and $y=-z$. So there exist nontrivial solutions: for instance, taking $z=1$ gives this equation of linear dependence:

$$-2\begin{pmatrix}1\\1\\1\end{pmatrix}-\begin{pmatrix}1\\-1\\2\end{pmatrix}+\begin{pmatrix}3\\1\\4\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Use this link to view the online demo

Move the sliders to solve the homogeneous vector equation in this example. Do you see why the vectors need to be coplanar in order for there to exist a nontrivial solution?

Example (Checking linear independence)**.** Is the set

$$\left\{\begin{pmatrix}1\\1\\-2\end{pmatrix},\begin{pmatrix}1\\-1\\2\end{pmatrix},\begin{pmatrix}3\\1\\4\end{pmatrix}\right\}$$

linearly independent?

Solution. Equivalently, we are asking if the homogeneous vector equation

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

has a nontrivial solution. We solve this by forming a matrix and row reducing (we do not augment because of this observation in Section 2.4):

$$\begin{pmatrix}1&1&3\\1&-1&1\\-2&2&4\end{pmatrix}\xrightarrow{\text{row reduce}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

This says $x = y = z = 0$, i.e., the only solution is the trivial solution. We conclude that the set is linearly independent.

Use this link to view the online demo

Move the sliders to solve the homogeneous vector equation in this example. Do you see why the vectors would need to be coplanar in order for there to exist a nontrivial solution?

Example (Vector parametric form)**.** An important observation is that the vectors coming from the parametric vector form of the solution of a matrix equation $Ax = 0$ are linearly independent. In this example in Section 2.4 we saw that the solution set of $Ax = 0$ for

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

is

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

Let's explain why the vectors $(1,1,0)$ and $(-2,0,1)$ are linearly independent. Suppose that

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_2 - 2x_3 \\ x_2 \\ x_3 \end{pmatrix}.$$

Comparing the second and third coordinates, we see that $x_2 = x_3 = 0$. This reasoning will work in any example, since the entries corresponding to the free variables are all equal to 1 or 0, and are only equal to 1 in one of the vectors. This observation forms part of this theorem in Section 2.7.

The above examples lead to the following recipe.

Recipe: Checking linear independence. A set of vectors $\{v_1, v_2, \ldots, v_k\}$ is linearly independent if and only if the vector equation

$$x_1 v_1 + x_2 v_2 + \cdots + x_k v_k = 0$$

has only the trivial solution, if and only if the matrix equation $Ax = 0$ has only the trivial solution, where A is the matrix with columns $v_1, v_2, \ldots, v_k$:

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_k \\ | & | & & | \end{pmatrix}.$$

This is true if and only if A has a pivot position in every column.
Solving the matrix equatiion $Ax = 0$ will either verify that the columns $v_1, v_2, \ldots, v_k$ are linearly independent, or will produce a linear dependence relation by substituting any nonzero values for the free variables.

(Recall that $Ax = 0$ has a nontrivial solution if and only if A has a column without a pivot: see this observation in Section 2.4.)

Suppose that A has more columns than rows. Then A cannot have a pivot in every column (it has at most one pivot per row), so its columns are automatically linearly dependent.

A wide matrix (a matrix with more columns than rows) has linearly dependent columns.

For example, four vectors in $\mathbf{R}^3$ are automatically linearly dependent. Note that a tall matrix may or may not have linearly independent columns.

Facts about linear independence.

1. *Two vectors are linearly dependent if and only if they are collinear, i.e., one is a scalar multiple of the other.*

2. *Any set containing the zero vector is linearly dependent.*

3. *If a subset of $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent, then $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent as well.*

Proof.

1. If $v_1 = cv_2$ then $v_1 - cv_2 = 0$, so $\{v_1, v_2\}$ is linearly dependent. In the other direction, if $x_1 v_1 + x_2 v_2 = 0$ with $x_1 \neq 0$ (say), then $v_1 = -\frac{x_2}{x_1} v_2$.

2. It is easy to produce a linear dependence relation if one vector is the zero vector: for instance, if $v_1 = 0$ then
$$1 \cdot v_1 + 0 \cdot v_2 + \cdots + 0 \cdot v_k = 0.$$

3. After reordering, we may suppose that $\{v_1, v_2, \ldots, v_r\}$ is linearly dependent, with $r < p$. This means that there is an equation of linear dependence
$$x_1 v_1 + x_2 v_2 + \cdots + x_r v_r = 0,$$
with at least one of $x_1, x_2, \ldots, x_r$ nonzero. This is also an equation of linear dependence among $\{v_1, v_2, \ldots, v_k\}$, since we can take the coefficients of $v_{r+1}, \ldots, v_k$ to all be zero.

□

With regard to the first fact, note that the zero vector is a multiple of any vector, so it is collinear with any other vector. Hence facts 1 and 2 are consistent with each other.

2.5.2 Criteria for Linear Independence

In this subsection we give two criteria for a set of vectors to be linearly independent. Keep in mind, however, that the actual definition is above.

Theorem. *A set of vectors $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent if and only if one of the vectors is in the span of the other ones.*

Any such vector may be removed without affecting the span.

Proof. Suppose, for instance, that v_3 is in $\mathrm{Span}\{v_1, v_2, v_4\}$, so we have an equation like
$$v_3 = 2v_1 - \frac{1}{2}v_2 + 6v_4.$$
We can subract v_3 from both sides of the equation to get
$$0 = 2v_1 - \frac{1}{2}v_2 - v_3 + 6v_4.$$
This is a linear dependence relation.

In this case, any linear combination of v_1, v_2, v_3, v_4 is already a linear combination of v_1, v_2, v_4:
$$\begin{aligned} x_1v_1 + x_2v_2 + x_3v_3 + x_4v_4 &= x_1v_1 + x_2v_2 + x_3\left(2v_1 - \frac{1}{2}v_2 + 6v_4\right) + x_4v_4 \\ &= (x_1 + 2x_3)v_1 + \left(x_2 - \frac{1}{2}x_3\right)v_2 + (x_4 + 6)v_4. \end{aligned}$$
Therefore, $\mathrm{Span}\{v_1, v_2, v_3, v_4\}$ is contained in $\mathrm{Span}\{v_1, v_2, v_4\}$. Any linear combination of v_1, v_2, v_4 is also a linear combination of v_1, v_2, v_3, v_4 (with the v_3-coefficient

equal to zero), so $\mathrm{Span}\{v_1, v_2, v_4\}$ is also contained in $\mathrm{Span}\{v_1, v_2, v_3, v_4\}$, and thus they are equal.

In the other direction, if we have a linear dependence relation like

$$0 = 2v_1 - \frac{1}{2}v_2 + v_3 - 6v_4,$$

then we can move any nonzero term to the left side of the equation and divide by its coefficient:

$$v_1 = \frac{1}{2}\left(\frac{1}{2}v_2 - v_3 + 6v_4\right).$$

This shows that v_1 is in $\mathrm{Span}\{v_2, v_3, v_4\}$.

We leave it to the reader to generalize this proof for any set of vectors. □

> **Warning.** In a linearly dependent set $\{v_1, v_2, \ldots, v_k\}$, it is not generally true that *any* vector v_j is in the span of the others, only that *at least one* of them is.

For example, the set $\left\{\binom{1}{0}, \binom{2}{0}, \binom{0}{1}\right\}$ is linearly dependent, but $\binom{0}{1}$ is not in the span of the other two vectors. Also see this figure below.

The previous theorem makes precise in what sense a set of linearly dependent vectors is redundant.

Theorem (Increasing Span Criterion)**.** *A set of vectors* $\{v_1, v_2, \ldots, v_k\}$ *is linearly independent if and only if, for every* j*, the vector* v_j *is not in* $\mathrm{Span}\{v_1, v_2, \ldots, v_{j-1}\}$*.*

Proof. It is equivalent to show that $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent if and only if v_j is in $\mathrm{Span}\{v_1, v_2, \ldots, v_{j-1}\}$ for some j. The "if" implication is an immediate consequence of the previous theorem. Suppose then that $\{v_1, v_2, \ldots, v_k\}$ is linearly dependent. This means that some v_j is in the span of the others. Choose the largest such j. We claim that this v_j is in $\mathrm{Span}\{v_1, v_2, \ldots, v_{j-1}\}$. If not, then

$$v_j = x_1v_1 + x_2v_2 + \cdots + x_{j-1}v_{j-1} + x_{j+1}v_{j+1} + \cdots + x_kv_k$$

with not all of $x_{j+1}, \ldots, x_k$ equal to zero. Suppose for simplicity that $x_k \neq 0$. Then we can rearrange:

$$v_k = -\frac{1}{x_k}\big(x_1v_1 + x_2v_2 + \cdots + x_{j-1}v_{j-1} - v_j + x_{j+1}v_{j+1} + \cdots + x_{p-1}v_{p-1}\big).$$

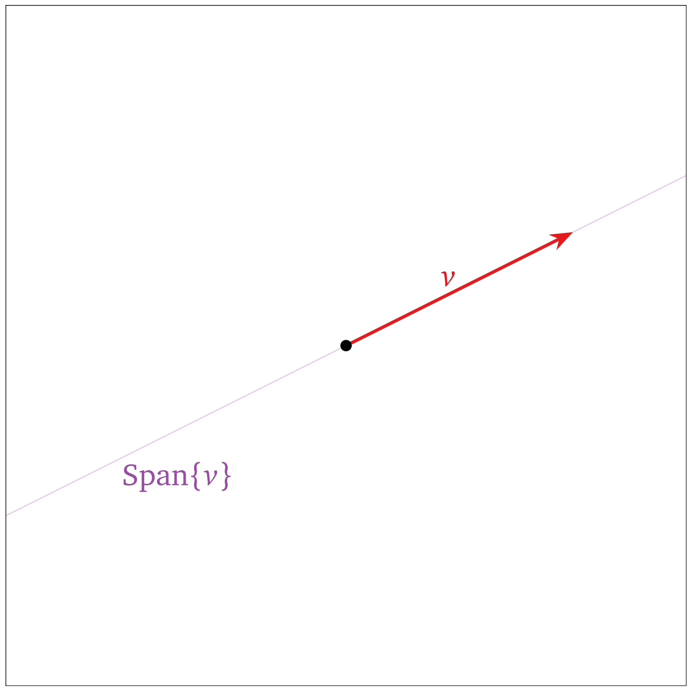
This says that v_k is in the span of $\{v_1, v_2, \ldots, v_{p-1}\}$, which contradicts our assumption that v_j is the last vector in the span of the others. □

We can rephrase this as follows:

> If you make a set of vectors by adding one vector at a time, and if the span got bigger every time you added a vector, then your set is linearly independent.

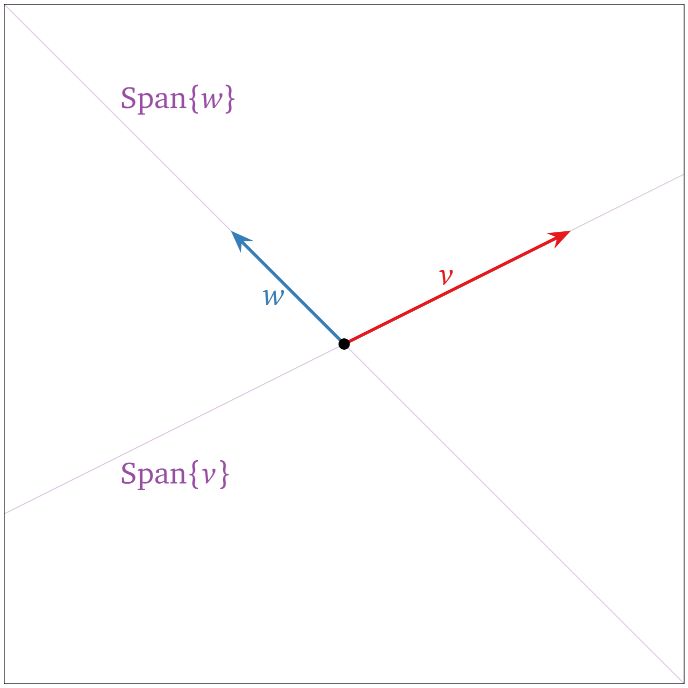
2.5.3 Pictures of Linear Independence

A set containg one vector $\{v\}$ is linearly independent when $v \neq 0$, since $xv = 0$ implies $x = 0$.



A set of two noncollinear vectors $\{v, w\}$ is linearly independent:

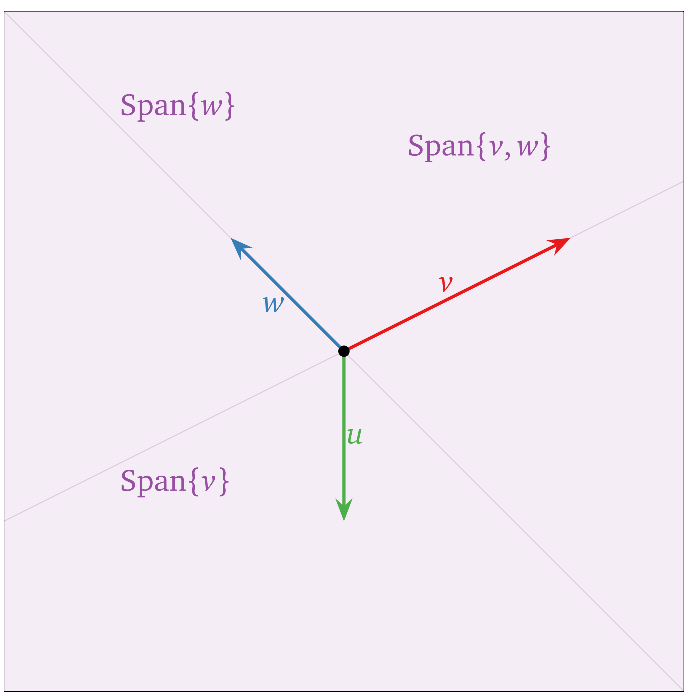
- Neither is in the span of the other, so we can apply the first criterion.
- The span got bigger when we added w, so we can apply the increasing span criterion.



The set of three vectors $\{v, w, u\}$ below is linearly dependent:

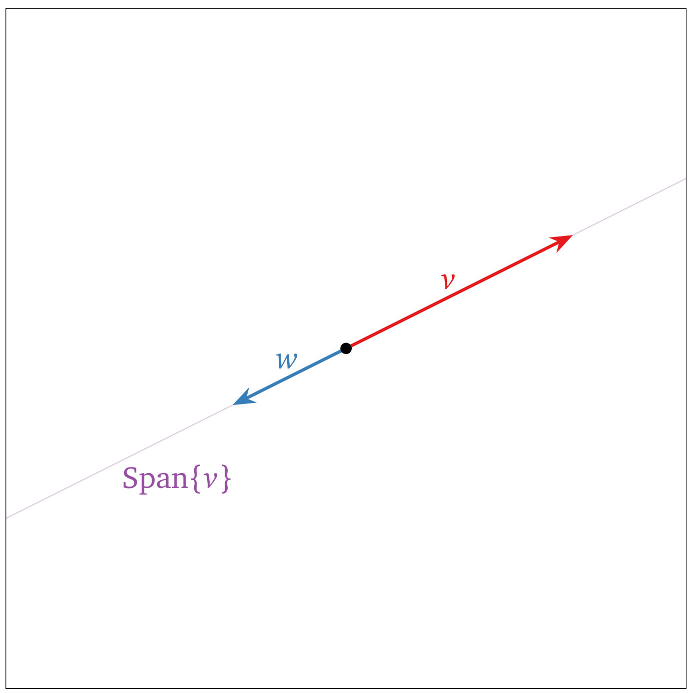
- u is in $\mathrm{Span}\{v, w\}$, so we can apply the first criterion.
- The span did not increase when we added u, so we can apply the increasing span criterion.

In the picture below, note that v is in $\mathrm{Span}\{u, w\}$, and w is in $\mathrm{Span}\{u, v\}$, so we can remove any of the three vectors without shrinking the span.

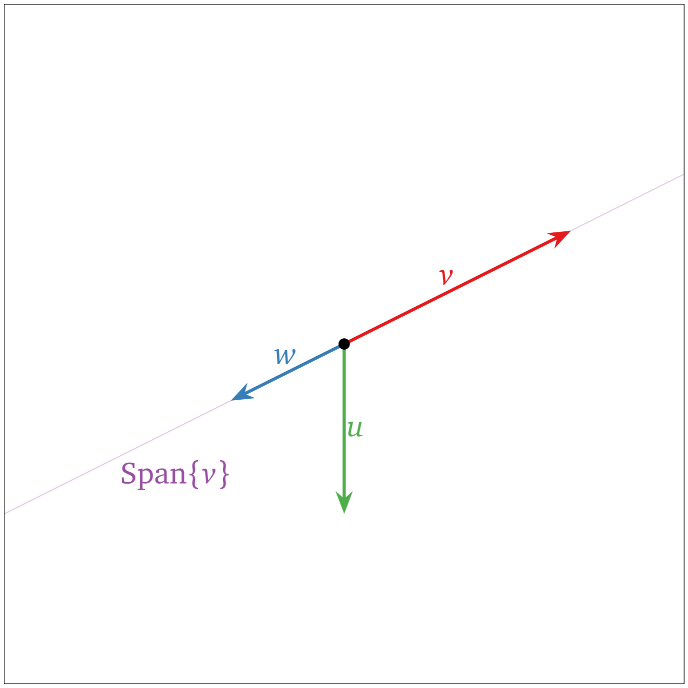


Two collinear vectors are always linearly dependent:

- w is in Span$\{v\}$, so we can apply the first criterion.
- The span did not increase when we added w, so we can apply the increasing span criterion.



These three vectors $\{v, w, u\}$ are linearly dependent: indeed, $\{v, w\}$ is already linearly dependent, so we can use the third fact.



Interactive: Linear independence of two vectors in $\mathbf{R}^2$.

Use this link to view the online demo

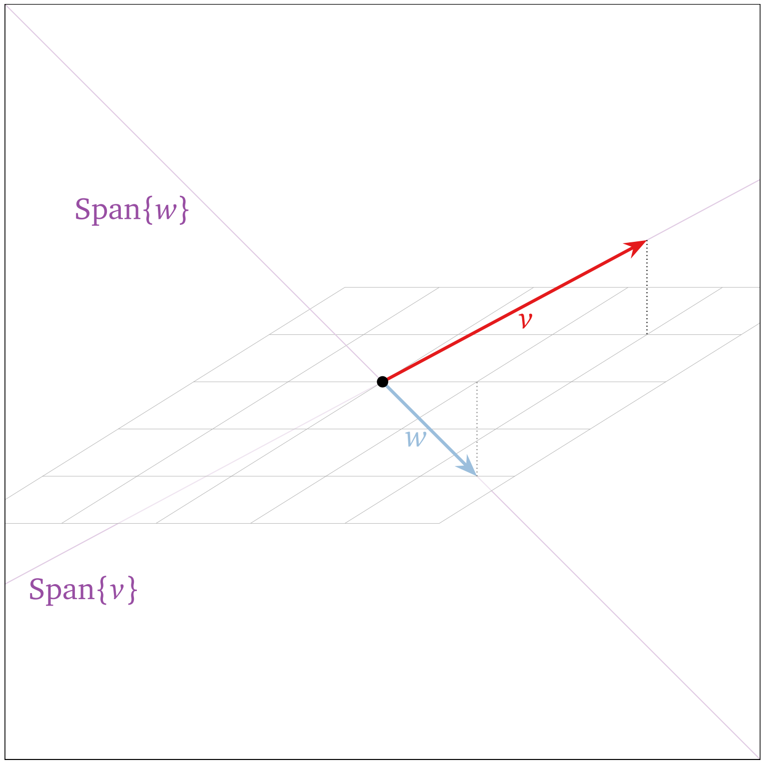
Move the vector heads and the demo will tell you if they are linearly independent and show you their span.

Interactive: Linear dependence of three vectors in $\mathbf{R}^2$.

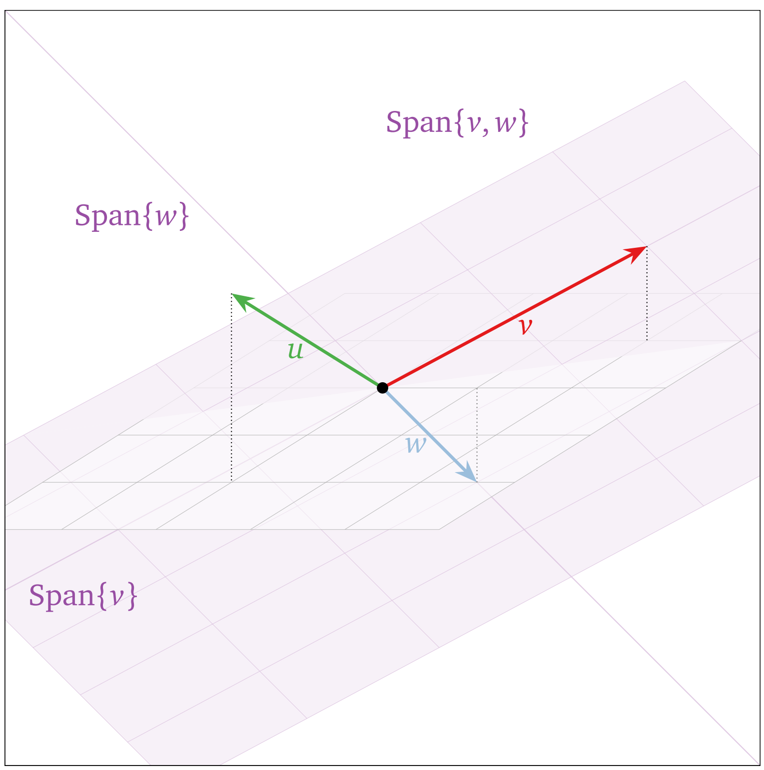
Use this link to view the online demo

Move the vector heads and the demo will tell you that they are linearly dependent and show you their span.

The two vectors $\{v, w\}$ below are linearly independent because they are not collinear.

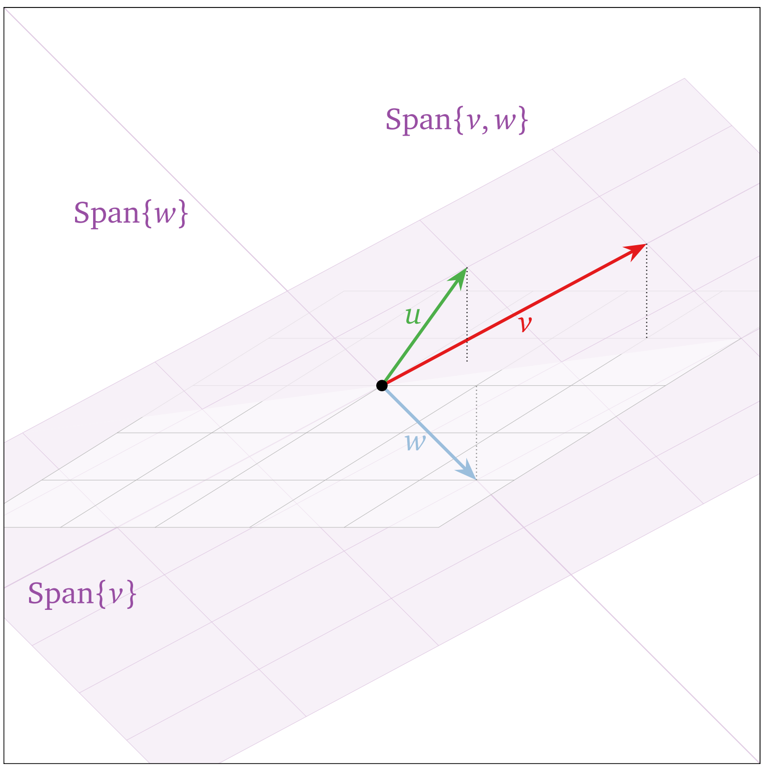


The three vectors $\{v, w, u\}$ below are linearly independent: the span got bigger when we added w, then again when we added u, so we can apply the increasing span criterion.



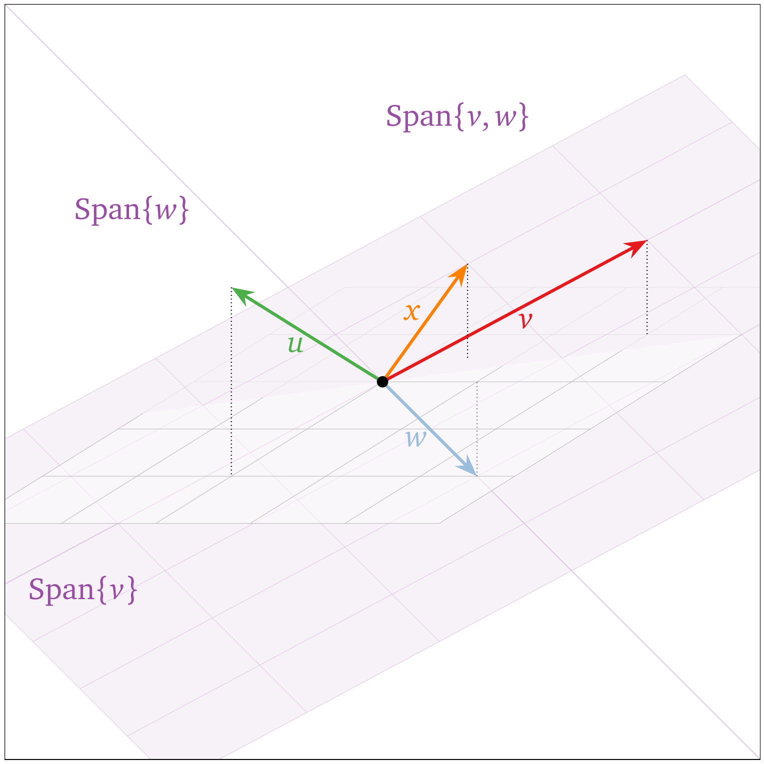
The three coplanar vectors $\{v, w, u\}$ below are linearly dependent:

- u is in $\mathrm{Span}\{v, w\}$, so we can apply the first criterion.
- The span did not increase when we added u, so we can apply the increasing span criterion.



Note that three vectors are linearly dependent if and only if they are *coplanar*. Indeed, $\{v, w, u\}$ is linearly dependent if and only if one vector is in the span of the other two, which is a plane (or a line) (or $\{0\}$).

The four vectors $\{v, w, u, x\}$ below are linearly dependent: they are the columns of a wide matrix. Note however that u is not contained in $\mathrm{Span}\{v, w, x\}$. See this warning.



The vectors $\{v, w, u, x\}$ *are linearly dependent, but u is not contained in* $\text{Span}\{v, w, x\}$.

Interactive: Linear independence of two vectors in $\mathbf{R}^3$.

Use this link to view the online demo

Move the vector heads and the demo will tell you if they are linearly independent and show you their span.

Interactive: Linear independence of three vectors in $\mathbf{R}^3$.

Use this link to view the online demo

Move the vector heads and the demo will tell you if they are linearly independent and show you their span.

2.5.4 Linear Dependence and Free Variables

In light of this important note and this criterion, it is natural to ask which columns of a matrix are redundant, i.e., which we can remove without affecting the column span.

Theorem. *Let $v_1, v_2, \ldots, v_k$ be vectors in $\mathbf{R}^n$, and consider the matrix*

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_k \\ | & | & & | \end{pmatrix}.$$

Then we can delete the columns of A without *pivots (the columns corresponding to the free variables), without changing* $\operatorname{Span}\{v_1, v_2, \ldots, v_k\}$.

The pivot columns are linearly independent, so we cannot delete any more columns without changing the span.

Proof. If the matrix is in reduced row echelon form:

$$A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

then the column without a pivot is visibly in the span of the pivot columns:

$$\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

and the pivot columns are linearly independent:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = x_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_4\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix} \implies x_1 = x_2 = x_4 = 0.$$

If the matrix is not in reduced row echelon form, then we row reduce:

$$A = \begin{pmatrix} 1 & 7 & 23 & 3 \\ 2 & 4 & 16 & 0 \\ -1 & -2 & -8 & 4 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The following two vector equations have the same solution set, as they come from row-equivalent matrices:

$$x_1\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + x_2\begin{pmatrix} 7 \\ 4 \\ -2 \end{pmatrix} + x_3\begin{pmatrix} 23 \\ 16 \\ -8 \end{pmatrix} + x_4\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} = 0$$

$$x_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} + x_4\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0.$$

We conclude that

$$\begin{pmatrix} 23 \\ 16 \\ -8 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + 3\begin{pmatrix} 7 \\ 4 \\ -2 \end{pmatrix} + 0\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}$$

and that

$$x_1\begin{pmatrix}1\\2\\-1\end{pmatrix}+x_2\begin{pmatrix}7\\4\\-2\end{pmatrix}+x_4\begin{pmatrix}3\\0\\4\end{pmatrix}=0$$

has only the trivial solution. □

Note that it is necessary to row reduce A to find which are its pivot columns. However, the span of the columns of the row reduced matrix is generally *not* equal to the span of the columns of A: one must use the pivot columns of the *original* matrix. See theorem in Section 2.7 for a restatement of the above theorem.

Example. The matrix

$$A=\begin{pmatrix}1&2&0&-1\\-2&-3&4&5\\2&4&0&-2\end{pmatrix}$$

has reduced row echelon form

$$\begin{pmatrix}1&0&-8&-7\\0&1&4&3\\0&0&0&0\end{pmatrix}.$$

Therefore, the first two columns of A are the pivot columns, so we can delete the others without changing the span:

$$\text{Span}\left\{\begin{pmatrix}1\\-2\\2\end{pmatrix},\begin{pmatrix}2\\-3\\4\end{pmatrix}\right\}=\text{Span}\left\{\begin{pmatrix}1\\-2\\2\end{pmatrix},\begin{pmatrix}2\\-3\\4\end{pmatrix},\begin{pmatrix}0\\4\\0\end{pmatrix},\begin{pmatrix}-1\\5\\-2\end{pmatrix}\right\}.$$

Moreover, the first two columns are linearly independent.

Pivot Columns and Dimension. Let d be the number of pivot columns in the matrix

$$A=\begin{pmatrix}|&|&&|\\v_1&v_2&\cdots&v_k\\|&|&&|\end{pmatrix}.$$

- If $d=1$ then $\text{Span}\{v_1,v_2,\ldots,v_k\}$ is a line.
- If $d=2$ then $\text{Span}\{v_1,v_2,\ldots,v_k\}$ is a plane.
- If $d=3$ then $\text{Span}\{v_1,v_2,\ldots,v_k\}$ is a 3-space.
- Et cetera.

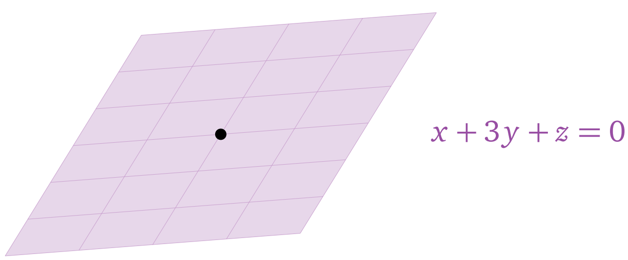
The number d is called the dimension. We discussed this notion in this important note in Section 2.4 and this important note in Section 2.4. We will define this concept rigorously in Section 2.7.

2.6 Subspaces

Objectives

1. Learn the definition of a subspace.
2. Learn to determine whether or not a subset is a subspace.
3. Learn the most important examples of subspaces.
4. Learn to write a given subspace as a column space or null space.
5. *Recipe:* compute a spanning set for a null space.
6. *Picture:* whether a subset of $\mathbf{R}^2$ or $\mathbf{R}^3$ is a subspace or not.
7. *Vocabulary words:* **subspace**, **column space**, **null space**.

In this section we discuss *subspaces* of $\mathbf{R}^n$. A subspace turns out to be exactly the same thing as a span, except we don't have a particular set of spanning vectors in mind. This change in perspective is quite useful, as it is easy to produce subspaces that are not obviously spans. For example, the solution set of the equation $x + 3y + z = 0$ is a span because the equation is homogeneous, but we would have to compute the parametric vector form in order to write it as a span.



(A subspace also turns out to be the same thing as the solution set of a homogeneous system of equations.)

2.6.1 Subspaces: Definition and Examples

Definition. A **subset** of $\mathbf{R}^n$ is any collection of points of $\mathbf{R}^n$.

For instance, the unit circle

$$C = \left\{(x, y) \text{ in } \mathbf{R}^2 \;\middle|\; x^2 + y^2 = 1\right\}$$

is a subset of $\mathbf{R}^2$.

Above we expressed C in set builder notation: in English, it reads "C is the set of all ordered pairs (x, y) in $\mathbf{R}^2$ such that $x^2 + y^2 = 1$."

Definition. A **subspace** of $\mathbf{R}^n$ is a subset V of $\mathbf{R}^n$ satisfying:

1. ***Non-emptiness:*** The zero vector is in V.
2. ***Closure under addition:*** If u and v are in V, then $u+v$ is also in V.
3. ***Closure under scalar multiplication:*** If v is in V and c is in $\mathbf{R}$, then cv is also in V.

As a consequence of these properties, we see:

- If v is a vector in V, then all scalar multiples of v are in V by the third property. In other words the line through any nonzero vector in V is also contained in V.
- If u, v are vectors in V and c, d are scalars, then cu, dv are also in V by the third property, so $cu + dv$ is in V by the second property. Therefore, all of Span$\{u, v\}$ is contained in V
- Similarly, if $v_1, v_2, \ldots, v_n$ are all in V, then Span$\{v_1, v_2, \ldots, v_n\}$ is contained in V. In other words, *a subspace contains the span of any vectors in it.*

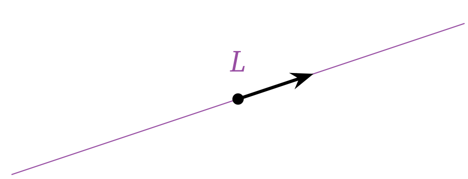
If you choose enough vectors, then eventually their span will fill up V, so we already see that a *subspace is a span*. See this theorem below for a precise statement.

Remark. Suppose that V is a non-empty subset of $\mathbf{R}^n$ that satisfies properties 2 and 3. Let v be any vector in V. Then $0v = 0$ is in V by the third property, so V automatically satisfies property 1. It follows that the only subset of $\mathbf{R}^n$ that satisfies properties 2 and 3 but *not* property 1 is the empty subset $\{\}$. This is why we call the first property "non-emptiness".

Example. The set $\mathbf{R}^n$ is a subspace of itself: indeed, it contains zero, and is closed under addition and scalar multiplication.

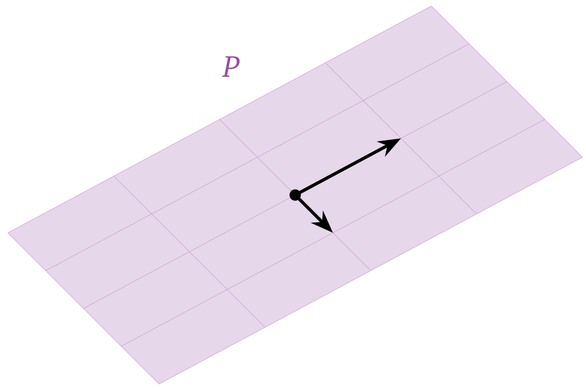
Example. The set $\{0\}$ containing only the zero vector is a subspace of $\mathbf{R}^n$: it contains zero, and if you add zero to itself or multiply it by a scalar, you always get zero.

Example (A line through the origin)**.** A line L through the origin is a subspace.



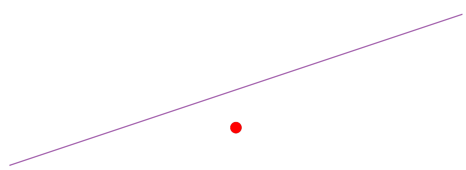
Indeed, L contains zero, and is easily seen to be closed under addition and scalar multiplication.

Example (A plane through the origin)**.** A plane P through the origin is a subspace.

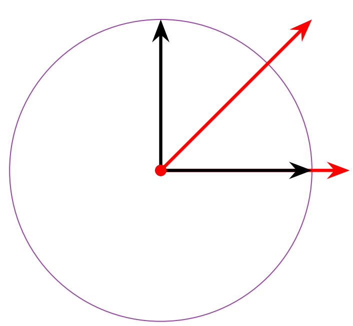


Indeed, P contains zero; the sum of two vectors in P is also in P; and any scalar multiple of a vector in P is also in P.

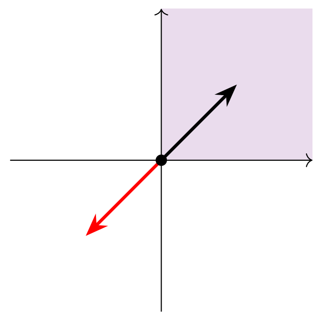
Non-example (A line not containing the origin). A line L (or any other subset) that does not contain the origin is not a subspace. It fails the first defining property: *every subspace contains the origin* by definition.



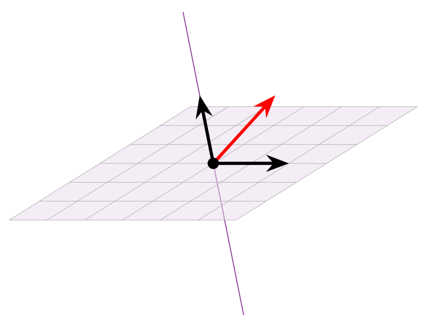
Non-example (A circle). The unit circle C is not a subspace. It fails all three defining properties: it does not contain the origin, it is not closed under addition, and it is not closed under scalar multiplication. In the picture, one red vector is the sum of the two black vectors (which are contained in C), and the other is a scalar multiple of a black vector.



Non-example (The first quadrant). The first quadrant in $\mathbf{R}^2$ is not a subspace. It contains the origin and is closed under addition, but it is not closed under scalar multiplication (by negative numbers).



Non-example (A line union a plane). The union of a line and a plane in $\mathbf{R}^3$ is not a subspace. It contains the origin and is closed under scalar multiplication, but it is not closed under addition: the sum of a vector on the line and a vector on the plane is not contained in the line or in the plane.



Subsets versus Subspaces. A sub*set* of $\mathbf{R}^n$ is any collection of vectors whatsoever. For instance, the unit circle

$$C = \left\{ (x, y) \text{ in } \mathbf{R}^2 \;\middle|\; x^2 + y^2 = 1 \right\}$$

is a subset of $\mathbf{R}^2$, but it is not a subspace. In fact, all of the non-examples above are still subsets of $\mathbf{R}^n$. A subspace is a subset that happens to satisfy the three additional defining properties.

In order to verify that a subset of $\mathbf{R}^n$ is in fact a subspace, one has to check the three defining properties. That is, unless the subset has already been verified to be a subspace: see this important note below.

Example (Verifying that a subset is a subspace)**.** Let

$$V = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \text{ in } \mathbf{R}^2 \;\middle|\; 2a = 3b \right\}.$$

Verify that V is a subspace.

Solution. First we point out that the condition "$2a = 3b$" defines whether or not a vector is in V: that is, to say $\binom{a}{b}$ is in V *means* that $2a = 3b$. In other words, a vector is in V if twice its first coordinate equals three times its second coordinate. So for instance, $\binom{3}{2}$ and $\binom{1/2}{1/3}$ are in V, but $\binom{2}{3}$ is not because $2 \cdot 2 \neq 3 \cdot 3$.

Let us check the first property. The subset V does contain the zero vector $\binom{0}{0}$, because $2 \cdot 0 = 3 \cdot 0$.

Next we check the second property. To show that V is closed under addition, we have to check that for *any* vectors $u = \binom{a}{b}$ and $v = \binom{c}{d}$ in V, the sum $u + v$ is in V. Since we cannot assume anything else about u and v, we must treat them as *unknowns*.

We have

$$\begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a + c \\ b + d \end{pmatrix}.$$

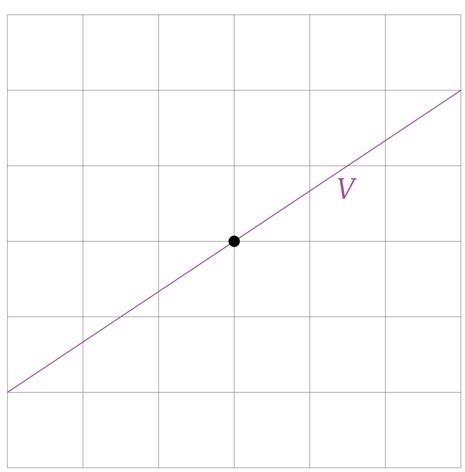
To say that $\binom{a+c}{b+d}$ is contained in V means that $2(a + c) = 3(b + d)$, or $2a + 2c = 3b + 3d$. The one thing we are allowed to assume about u and v is that $2a = 3b$ and $2c = 3d$, so we see that $u + v$ is indeed contained in V.

Next we check the third property. To show that V is closed under scalar multiplication, we have to check that for any vector $v = \binom{a}{b}$ in V and any scalar c in $\mathbf{R}$, the product cv is in V. Again, we must treat v and c as unknowns. We have

$$c\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ca \\ cb \end{pmatrix}.$$

To say that $\binom{ca}{cb}$ is contained in V means that $2(ca) = 3(cb)$, i.e., that $c \cdot 2a = c \cdot 3b$. The one thing we are allowed to assume about v is that $2a = 3b$, so cv is indeed contained in V.

Since V satisfies all three defining properties, it is a subspace. In fact, it is the line through the origin with slope $2/3$.



Example (Showing that a subset is not a subspace)**.** Let

$$V = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \text{ in } \mathbf{R}^2 \;\middle|\; ab = 0 \right\}.$$

Is V a subspace?

Solution. First we point out that the condition "$ab = 0$" defines whether or not a vector is in V: that is, to say $\binom{a}{b}$ is in V *means* that $ab = 0$. In other words, a vector is in V if the product of its coordinates is zero, i.e., if one (or both) of its coordinates are zero. So for instance, $\binom{1}{0}$ and $\binom{0}{2}$ are in V, but $\binom{1}{1}$ is not because $1 \cdot 1 \neq 0$.

Let us check the first property. The subset V does contain the zero vector $\binom{0}{0}$, because $0 \cdot 0 = 0$.

Next we check the third property. To show that V is closed under scalar multiplication, we have to check that for *any* vector $v = \binom{a}{b}$ in V and *any* scalar c in $\mathbf{R}$, the product cv is in V. Since we cannot assume anything else about v and c, we must treat them as *unknowns*.

We have

$$c\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ca \\ cb \end{pmatrix}.$$

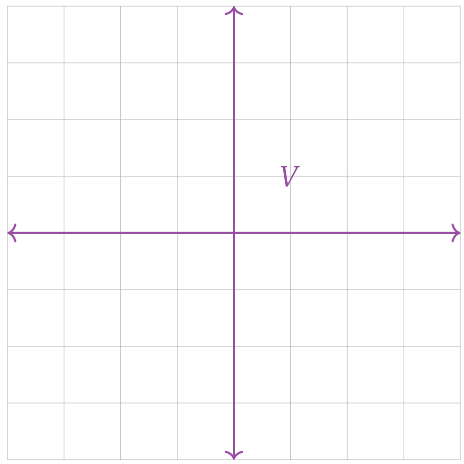
To say that $\binom{ca}{cb}$ is contained in V means that $(ca)(cb) = 0$. Rewriting, this means $c^2(ab) = 0$. The one thing we are allowed to assume about v is that $ab = 0$, so we see that cv is indeed contained in V.

Next we check the second property. It turns out that V is not closed under addition; to verify this, we must show that there exists *some* vectors u, v in V such that $u + v$ is not contained in V. The easiest way to do so is to produce examples of such vectors. We can take $u = \binom{1}{0}$ and $v = \binom{0}{1}$; these are contained in V because the products of their coordinates are zero, but

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

is not contained in V because $1 \cdot 1 \neq 0$.

Since V does not satisfy the second property (it is not closed under addition), we conclude that V is *not* a subspace. Indeed, it is the union of the two coordinate axes, which is not a span.



2.6.2 Common Types of Subspaces

Theorem (Spans are Subspaces and Subspaces are Spans). *If $v_1, v_2, \ldots, v_p$ are any vectors in $\mathbf{R}^n$, then* $\mathrm{Span}\{v_1, v_2, \ldots, v_p\}$ *is a subspace of $\mathbf{R}^n$. Moreover, any subspace of $\mathbf{R}^n$ can be written as a span of a set of p linearly independent vectors in $\mathbf{R}^n$ for $p \le n$.*

Proof. To show that $\mathrm{Span}\{v_1, v_2, \ldots, v_p\}$ is a subspace, we have to verify the three defining properties.

1. The zero vector $0 = 0v_1 + 0v_2 + \cdots + 0v_p$ is in the span.

2. If $u = a_1v_1 + a_2v_2 + \cdots + a_pv_p$ and $v = b_1v_1 + b_2v_2 + \cdots + b_pv_p$ are in $\mathrm{Span}\{v_1, v_2, \ldots, v_p\}$, then

$$u + v = (a_1 + b_1)v_1 + (a_2 + b_2)v_2 + \cdots + (a_p + b_p)v_p$$

   is also in $\mathrm{Span}\{v_1, v_2, \ldots, v_p\}$.

3. If $v = a_1v_1 + a_2v_2 + \cdots + a_pv_p$ is in $\mathrm{Span}\{v_1, v_2, \ldots, v_p\}$ and c is a scalar, then

$$cv = ca_1v_1 + ca_2v_2 + \cdots + ca_pv_p$$

   is also in $\mathrm{Span}\{v_1, v_2, \ldots, v_p\}$.

Since $\mathrm{Span}\{v_1, v_2, \ldots, v_p\}$ satisfies the three defining properties of a subspace, it is a subspace.

Now let V be a subspace of $\mathbf{R}^n$. If V is the zero subspace, then it is the span of the empty set, so we may assume V is nonzero. Choose a nonzero vector v_1 in V. If $V = \mathrm{Span}\{v_1\}$, then we are done. Otherwise, there exists a vector v_2 that is in V but not in $\mathrm{Span}\{v_1\}$. Then $\mathrm{Span}\{v_1, v_2\}$ is contained in V, and by the increasing span criterion in Section 2.5, the set $\{v_1, v_2\}$ is linearly independent. If $V = \mathrm{Span}\{v_1, v_2\}$ then we are done. Otherwise, we continue in this fashion until we have written $V = \mathrm{Span}\{v_1, v_2, \ldots, v_p\}$ for some linearly independent set $\{v_1, v_2, \ldots, v_p\}$. This process terminates after at most n steps by this important note in Section 2.5. □

If $V = \mathrm{Span}\{v_1, v_2, \ldots, v_p\}$, we say that V is the subspace **spanned by** or **generated by** the vectors $v_1, v_2, \ldots, v_p$. We call $\{v_1, v_2, \ldots, v_p\}$ a **spanning set** for V.

Any matrix naturally gives rise to *two* subspaces.

Definition. Let A be an $m \times n$ matrix.

- The **column space** of A is the subspace of $\mathbf{R}^m$ spanned by the columns of A. It is written $\mathrm{Col}(A)$.
- The **null space** of A is the subspace of $\mathbf{R}^n$ consisting of all solutions of the homogeneous equation $Ax = 0$:

$$\mathrm{Nul}(A) = \left\{ x \text{ in } \mathbf{R}^n \;\middle|\; Ax = 0 \right\}.$$

The column space is defined to be a span, so it is a subspace by the above theorem. We need to verify that the null space is really a subspace. In Section 2.4 we already saw that the set of solutions of $Ax = 0$ is always a span, so the fact that the null spaces is a subspace should not come as a surprise.

Proof. We have to verify the three defining properties. To say that a vector v is in $\mathrm{Nul}(A)$ *means* that $Av = 0$.

1. The zero vector is in $\mathrm{Nul}(A)$ because $A0 = 0$.
2. Suppose that u, v are in $\mathrm{Nul}(A)$. This means that $Au = 0$ and $Av = 0$. Hence

$$A(u + v) = Au + Av = 0 + 0 = 0$$

by the linearity of the matrix-vector product in Section 2.3. Therefore, $u + v$ is in $\mathrm{Nul}(A)$.

3. Suppose that v is in $\mathrm{Nul}(A)$ and c is a scalar. Then

$$A(cv) = cAv = c \cdot 0 = 0$$

by the linearity of the matrix-vector product in Section 2.3, so cv is also in $\mathrm{Nul}(A)$.

Since $\mathrm{Nul}(A)$ satisfies the three defining properties of a subspace, it is a subspace. □

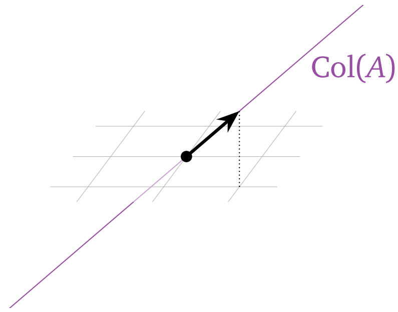
Example. Describe the column space and the null space of

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Solution. The column space is the span of the columns of A:

$$\mathrm{Col}(A) = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

This is a line in $\mathbf{R}^3$.



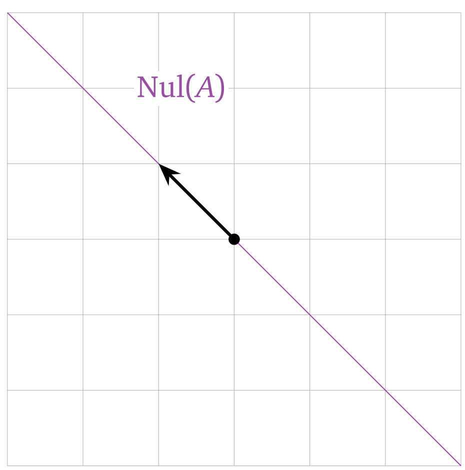
The null space is the solution set of the homogeneous system $Ax = 0$. To compute this, we need to row reduce A. Its reduced row echelon form is

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This gives the equation $x + y = 0$, or

$$\begin{cases} x = -y \\ y = y \end{cases} \xrightarrow{\text{parametric vector form}} \begin{pmatrix} x \\ y \end{pmatrix} = y \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Hence the null space is $\mathrm{Span}\left\{\binom{-1}{1}\right\}$, which is a line in $\mathbf{R}^2$.



Notice that the column space is a subspace of $\mathbf{R}^3$, whereas the null space is a subspace of $\mathbf{R}^2$. This is because A has three rows and two columns.

The column space and the null space of a matrix are both subspaces, so they are both spans. The column space of a matrix A is defined to be the span of the columns of A. The null space is defined to be the solution set of $Ax = 0$, so this is a good example of a kind of subspace that we can define *without any spanning set in mind.* In other words, it is easier to show that the null space is a subspace than to show it is a span—see the proof above. In order to do computations, however, it is usually necessary to find a spanning set.

Null Spaces are Solution Sets. The null space of a matrix is the solution set of a homogeneous system of equations. For example, the null space of the matrix

$$A = \begin{pmatrix} 1 & 7 & 2 \\ -2 & 1 & 3 \\ 4 & -2 & -3 \end{pmatrix}$$

is the solution set of $Ax = 0$, i.e., the solution set of the system of equations

$$\begin{cases} x + 7y + 2z = 0 \\ -2x + y + 3z = 0 \\ 4x - 2y - 3z = 0. \end{cases}$$

Conversely, the solution set of any homogeneous system of equations is precisely the null space of the corresponding coefficient matrix.

To find a spanning set for the null space, one has to solve a system of homogeneous equations.

Recipe: Compute a spanning set for a null space. To find a spanning set for Nul(A), compute the parametric vector form of the solutions to the homogeneous equation $Ax = 0$. The vectors attached to the free variables form a spanning set for Nul(A).

Example (Two free variables)**.** Find a spanning set for the null space of the matrix

$$A = \begin{pmatrix} 2 & 3 & -8 & -5 \\ -1 & 2 & -3 & -8 \end{pmatrix}.$$

Solution. We compute the parametric vector form of the solutions of $Ax = 0$. The reduced row echelon form of A is

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

The free variables are x_3 and x_4; the parametric form of the solution set is

$$\begin{cases} x_1 = x_3 - 2x_4 \\ x_2 = 2x_3 + 3x_4 \\ x_3 = x_3 \\ x_4 = x_4 \end{cases} \xrightarrow[\text{vector form}]{\text{parametric}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = x_3 \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ 3 \\ 0 \\ 1 \end{pmatrix}.$$

Therefore,

$$\text{Nul}(A) = \text{Span}\left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Example (No free variables)**.** Find a spanning set for the null space of the matrix

$$A=\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix}.$$

Solution. We compute the parametric vector form of the solutions of $Ax = 0$. The reduced row echelon form of A is

$$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

There are no free variables; hence the only solution of $Ax = 0$ is the trivial solution. In other words,

$$\mathrm{Nul}(A) = \{0\} = \mathrm{Span}\{0\}.$$

It is natural to define $\mathrm{Span}\{\} = \{0\}$, so that we can take our spanning set to be empty. This is consistent with the definition of dimension in Section 2.7.

Writing a subspace as a column space or a null space A subspace can be given to you in many different forms. In practice, computations involving subspaces are much easier if your subspace is the column space or null space of a matrix. The simplest example of such a computation is finding a spanning set: a column space is by definition the span of the columns of a matrix, and we showed above how to compute a spanning set for a null space using parametric vector form. For this reason, it is useful to rewrite a subspace as a column space or a null space before trying to answer questions about it.

> When asking questions about a subspace, it is usually best to rewrite the subspace as a column space or a null space.

This also applies to the question "is my subset a subspace?" If your subset is a column space or null space of a matrix, then the answer is yes.

Example. Let

$$V = \left\{\begin{pmatrix}a\\ b\end{pmatrix} \text{ in } \mathbf{R}^2 \;\middle|\; 2a = 3b\right\}$$

be the subset of a previous example. The subset V is exactly the solution set of the homogeneous equation $2x - 3y = 0$. Therefore,

$$V = \mathrm{Nul}\begin{pmatrix}2 & -3\end{pmatrix}.$$

In particular, it is a subspace. The reduced row echelon form of $\begin{pmatrix}2 & -3\end{pmatrix}$ is $\begin{pmatrix}1 & -3/2\end{pmatrix}$, so the parametric form of V is $x = 3/2y$, so the parametric vector form is $\binom{x}{y} = y\binom{3/2}{1}$, and hence $\left\{\binom{3/2}{1}\right\}$ spans V.

Example. Let V be the plane in $\mathbf{R}^3$ defined by

$$V = \left\{ \begin{pmatrix} 2x+y \\ x-y \\ 3x-2y \end{pmatrix} \;\middle|\; x, y \text{ are in } \mathbf{R} \right\}.$$

Write V as the column space of a matrix.

Solution. Since

$$\begin{pmatrix} 2x+y \\ x-y \\ 3x-2y \end{pmatrix} = x \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + y \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix},$$

we notice that V is exactly the span of the vectors

$$\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}.$$

Hence

$$V = \operatorname{Col} \begin{pmatrix} 2 & 1 \\ 1 & -1 \\ 3 & -2 \end{pmatrix}.$$

2.7 Basis and Dimension

Objectives

1. Understand the definition of a basis of a subspace.
2. Understand the basis theorem.
3. *Recipes:* basis for a column space, basis for a null space, basis of a span.
4. *Picture:* basis of a subspace of $\mathbf{R}^2$ or $\mathbf{R}^3$.
5. *Theorem:* basis theorem.
6. *Essential vocabulary words:* **basis**, **dimension**.

2.7.1 Basis of a Subspace

As we discussed in Section 2.6, a subspace is the same as a span, except we do not have a set of spanning vectors in mind. There are infinitely many choices of spanning sets for a nonzero subspace; to avoid reduncancy, usually it is most convenient to choose a spanning set with the *minimal* number of vectors in it. This is the idea behind the notion of a basis.

Essential Definition. Let V be a subspace of $\mathbf{R}^n$. A **basis** of V is a set of vectors $\{v_1, v_2, \dots, v_m\}$ in V such that:

1. $V = \text{Span}\{v_1, v_2, \dots, v_m\}$, and
2. the set $\{v_1, v_2, \dots, v_m\}$ is linearly independent.

Recall that a set of vectors is *linearly independent* if and only if, when you remove any vector from the set, the span shrinks (Theorem 2.5.10). In other words, if $\{v_1, v_2, \dots, v_m\}$ is a basis of a subspace V, then no proper subset of $\{v_1, v_2, \dots, v_m\}$ will span V: it is a *minimal* spanning set. Any subspace admits a basis by this theorem in Section 2.6.

A nonzero subspace has *infinitely many* different bases, but they all contain the same number of vectors.

We leave it as an exercise to prove that any two bases have the same number of vectors; one might want to wait until after learning the invertible matrix theorem in Section 3.5.

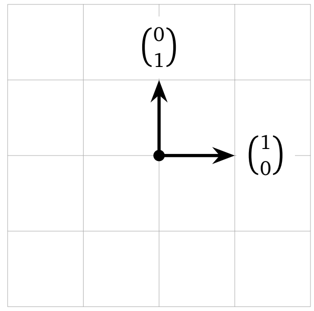
Essential Definition. Let V be a subspace of $\mathbf{R}^n$. The number of vectors in any basis of V is called the **dimension** of V, and is written $\dim V$.

Example (A basis of $\mathbf{R}^2$)**.** Find a basis of $\mathbf{R}^2$.

Solution. We need to find two vectors in $\mathbf{R}^2$ that span $\mathbf{R}^2$ and are linearly independent. One such basis is $\left\{\binom{1}{0}, \binom{0}{1}\right\}$:

1. They span because any vector $\binom{a}{b}$ can be written as a linear combination of $\binom{1}{0}, \binom{0}{1}$:
$$\begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$
2. They are linearly independent: if
$$x\begin{pmatrix} 1 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$
then $x = y = 0$.

This shows that the plane $\mathbf{R}^2$ has dimension 2.



Example (All bases of $\mathbf{R}^2$)**.** Find all bases of $\mathbf{R}^2$.

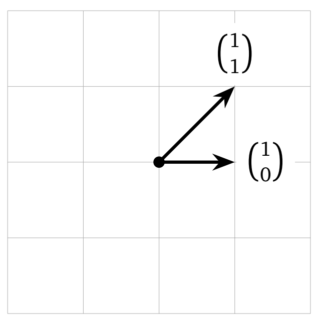
Solution. We know from the previous example that $\mathbf{R}^2$ has dimension 2, so any basis of $\mathbf{R}^2$ has two vectors in it. Let v_1, v_2 be vectors in $\mathbf{R}^2$, and let A be the matrix with columns v_1, v_2.

1. To say that $\{v_1, v_2\}$ spans $\mathbf{R}^2$ means that A has a pivot in every *row*: see this theorem in Section 2.3.
2. To say that $\{v_1, v_2\}$ is linearly independent means that A has a pivot in every *column*: see this important note in Section 2.5.

Since A is a 2×2 matrix, it has a pivot in every row exactly when it has a pivot in every column. Hence any two noncollinear vectors form a basis of $\mathbf{R}^2$. For example,

$$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right\}$$

is a basis.



Example (The standard basis of $\mathbf{R}^n$)**.** One shows exactly as in the above example that the standard coordinate vectors

$$e_1 = \begin{pmatrix}1\\0\\\vdots\\0\\0\end{pmatrix}, \quad e_2 = \begin{pmatrix}0\\1\\\vdots\\0\\0\end{pmatrix}, \quad \ldots, \quad e_{n-1} = \begin{pmatrix}0\\0\\\vdots\\1\\0\end{pmatrix}, \quad e_n = \begin{pmatrix}0\\0\\\vdots\\0\\1\end{pmatrix}$$

form a basis for $\mathbf{R}^n$. This is sometimes known as the **standard basis**.

In particular, $\mathbf{R}^n$ has dimension n.

Example. The previous example implies that any basis for $\mathbf{R}^n$ has n vectors in it. Let $v_1, v_2, \ldots, v_n$ be vectors in $\mathbf{R}^n$, and let A be the $n \times n$ matrix with columns $v_1, v_2, \ldots, v_n$.

1. To say that $\{v_1, v_2, \ldots, v_n\}$ spans $\mathbf{R}^n$ means that A has a pivot position in every *row*: see this theorem in Section 2.3.

2. To say that $\{v_1, v_2, \ldots, v_n\}$ is linearly independent means that A has a pivot position in every *column*: see this important note in Section 2.5.

Since A is a square matrix, it has a pivot in every row if and only if it has a pivot in every column. We will see in Section 3.5 that the above two conditions are equivalent to the *invertibility* of the matrix A.

Example. Let

$$V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \text{ in } \mathbf{R}^3 \;\middle|\; x + 3y + z = 0 \right\} \qquad \mathcal{B} = \left\{ \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix} \right\}.$$

Verify that V is a subspace, and show directly that $\mathcal{B}$ is a basis for V.

Solution. First we observe that V is the solution set of the homogeneous equation $x + 3y + z = 0$, so it is a subspace: see this important note in Section 2.6. To show that $\mathcal{B}$ is a basis, we really need to verify three things:

1. Both vectors are in V because

$$\begin{aligned} (-3) + 3(1) + (0) &= 0 \\ (0) + 3(1) + (-3) &= 0. \end{aligned}$$

2. *Span:* suppose that $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is in V. Since $x+3y+z = 0$ we have $y = -\frac{1}{3}(x+z)$, so

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ -\frac{1}{3}(x+z) \\ z \end{pmatrix} = -\frac{x}{3} \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} - \frac{z}{3} \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix}.$$

   Hence $\mathcal{B}$ spans V.

3. *Linearly independent:*

$$c_1 \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ -3 \end{pmatrix} = 0 \implies \begin{pmatrix} -3c_1 \\ c_1 + c_2 \\ -3c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \implies c_1 = c_2 = 0.$$

   Alternatively, one can observe that the two vectors are not collinear.

Since V has a basis with two vectors, it has dimension two: it is a *plane*.

Use this link to view the online demo

A picture of the plane V and its basis $\mathcal{B} = \{v_1, v_2\}$. Note that $\mathcal{B}$ spans V and is linearly independent.

This example is somewhat contrived, in that we will learn systematic methods for verifying that a subset is a basis. The intention is to illustrate the defining properties of a basis.

2.7.2 Computing a Basis for a Subspace

Now we show how to find bases for the column space of a matrix and the null space of a matrix. In order to find a basis for a given subspace, it is usually best to rewrite the subspace as a column space or a null space first: see this important note in Section 2.6.

A basis for the column space First we show how to compute a basis for the column space of a matrix.

Theorem. *The pivot columns of a matrix A form a basis for* $\mathrm{Col}(A)$.

Proof. This is a restatement of a theorem in Section 2.5. □

The above theorem is referring to the pivot columns in the *original* matrix, not its reduced row echelon form. Indeed, a matrix and its reduced row echelon form generally have different column spaces. For example, in the matrix A below:

$$A = \begin{pmatrix} \boxed{\begin{matrix}1\\-2\\2\end{matrix}} & \boxed{\begin{matrix}2\\-3\\4\end{matrix}} & \begin{matrix}0\\4\\0\end{matrix} & \begin{matrix}-1\\5\\-2\end{matrix} \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

pivot columns = basis ⟵ pivot columns in RREF

the pivot columns are the first two columns, so a basis for Col(A) is

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix} \right\}.$$

The first two columns of the reduced row echelon form certainly span a different subspace, as

$$\text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} = \left\{ \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \,\middle|\, a, b \text{ in } \mathbf{R} \right\} = (xy\text{-plane}),$$

but Col(A) contains vectors whose last coordinate is nonzero.

Corollary. *The dimension of* Col(A) *is the number of pivots of* A*.*

A basis of a span Computing a basis for a span is the same as computing a basis for a column space. Indeed, the span of finitely many vectors $v_1, v_2, \ldots, v_m$ *is* the column space of a matrix, namely, the matrix A whose columns are $v_1, v_2, \ldots, v_m$:

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_m \\ | & | & & | \end{pmatrix}.$$

Example (A basis of a span)**.** Find a basis of the subspace

$$V = \text{Span}\left\{ \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \\ -2 \end{pmatrix} \right\}.$$

Solution. The subspace V is the column space of the matrix

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ -2 & -3 & 4 & 5 \\ 2 & 4 & 0 & -2 \end{pmatrix}.$$

The reduced row echelon form of this matrix is

$$\begin{pmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The first two columns are pivot columns, so a basis for V is

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix} \right\}.$$

Use this link to view the online demo

A picture of the plane V and its basis $\mathcal{B} = \{v_1, v_2\}$.

Example (Another basis of the same span)**.** Find a basis of the subspace

$$V = \text{Span}\left\{ \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \\ -2 \end{pmatrix} \right\}$$

which does not consist of the first two vectors, as in the previous example.

Solution. The point of this example is that the above theorem gives *one* basis for V; as always, there are infinitely more.

Reordering the vectors, we can express V as the column space of

$$A' = \begin{pmatrix} 0 & -1 & 1 & 2 \\ 4 & 5 & -2 & -3 \\ 0 & -2 & 2 & 4 \end{pmatrix}.$$

The reduced row echelon form of this matrix is

$$\begin{pmatrix} 1 & 0 & 3/4 & 7/4 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The first two columns are pivot columns, so a basis for V is

$$\left\{ \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \\ -2 \end{pmatrix} \right\}.$$

These are the *last* two vectors in the given spanning set.

Use this link to view the online demo

A picture of the plane V and its basis $\mathcal{B} = \{v_1, v_2\}$.

A basis for the null space In order to compute a basis for the null space of a matrix, one has to find the parametric vector form of the solutions of the homogeneous equation $Ax = 0$.

Theorem. *The vectors attached to the free variables in the parametric vector form of the solution set of $Ax = 0$ form a basis of* $\mathrm{Nul}(A)$*.*

The proof of the theorem has two parts. The first part is that every solution lies in the span of the given vectors. This is automatic: the vectors are exactly chosen so that every solution is a linear combination of those vectors. The second part is that the vectors are linearly independent. This part was discussed in this example in Section 2.5.

A basis for a general subspace As mentioned at the beginning of this subsection, when given a subspace written in a different form, in order to compute a basis it is usually best to rewrite it as a column space or null space of a matrix.

Example (A basis of a subspace)**.** Let V be the subspace defined by

$$V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; x + 2y = z \right\}.$$

Find a basis for V. What is $\dim(V)$?

Solution. First we notice that V is exactly the solution set of the homogeneous linear equation $x + 2y - z = 0$. Hence $V = \mathrm{Nul}\begin{pmatrix} 1 & 2 & -1 \end{pmatrix}$. This matrix is in reduced row echelon form; the parametric form of the general solution is $x = -2y + z$, so the parametric vector form is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

It follows that a basis is

$$\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Since V has a basis with two vectors, its dimension is 2: it is a plane.

2.7.3 The Basis Theorem

Recall that $\{v_1, v_2, \ldots, v_n\}$ forms a basis for $\mathbf{R}^n$ if and only if the matrix A with columns $v_1, v_2, \ldots, v_n$ has a pivot in every row and column (see this example). Since A is an $n \times n$ matrix, these two conditions are equivalent: the vectors span if and only if they are linearly independent. The basis theorem is an abstract version of the preceding statement, that applies to any subspace.

Basis Theorem. *Let V be a subspace of dimension m. Then:*

- *Any m linearly independent vectors in V form a basis for V.*
- *Any m vectors that span V form a basis for V.*

Proof. Suppose that $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ is a set of linearly independent vectors in V. In order to show that $\mathcal{B}$ is a basis for V, we must prove that $V = \text{Span}\{v_1, v_2, \ldots, v_m\}$. If not, then there exists some vector v_{m+1} in V that is not contained in $\text{Span}\{v_1, v_2, \ldots, v_m\}$. By the increasing span criterion in Section 2.5, the set $\{v_1, v_2, \ldots, v_m, v_{m+1}\}$ is also linearly independent. Continuing in this way, we keep choosing vectors until we eventually do have a linearly independent spanning set: say $V = \text{Span}\{v_1, v_2, \ldots, v_m, \ldots, v_{m+k}\}$. Then $\{v_1, v_2, \ldots, v_{m+k}\}$ is a basis for V, which implies that $\dim(V) = m + k > m$. But we were assuming that V has dimension m, so $\mathcal{B}$ must have already been a basis.

Now suppose that $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ spans V. If $\mathcal{B}$ is not linearly independent, then by this theorem in Section 2.5, we can remove some number of vectors from $\mathcal{B}$ without shrinking its span. After reordering, we can assume that we removed the last k vectors without shrinking the span, and that we cannot remove any more. Now $V = \text{Span}\{v_1, v_2, \ldots, v_{m-k}\}$, and $\{v_1, v_2, \ldots, v_{m-k}\}$ is a basis for V because it is linearly independent. This implies that $\dim V = m - k < m$. But we were assuming that $\dim V = m$, so $\mathcal{B}$ must have already been a basis. □

In other words, if you *already* know that $\dim V = m$, and if you have a set of m vectors $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ in V, then you only have to check *one* of:

1. $\mathcal{B}$ is linearly independent, *or*
2. $\mathcal{B}$ spans V,

in order for $\mathcal{B}$ to be a basis of V. If you did not already know that $\dim V = m$, then you would have to check *both* properties.

To put it yet another way, suppose we have a set of vectors $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ in a subspace V. Then if any two of the following statements is true, the third must also be true:

1. $\mathcal{B}$ is linearly independent,
2. $\mathcal{B}$ spans V, and
3. $\dim V = m$.

For example, if V is a plane, then any two noncollinear vectors in V form a basis.

Example (Two noncollinear vectors form a basis of a plane)**.** Find a basis of the subspace

$$V = \mathrm{Span}\left\{\begin{pmatrix}1\\-2\\2\end{pmatrix}, \begin{pmatrix}2\\-3\\4\end{pmatrix}, \begin{pmatrix}0\\4\\0\end{pmatrix}, \begin{pmatrix}-1\\5\\-2\end{pmatrix}\right\}$$

which is different from the bases in this example and this example.

Solution. We know from the previous examples that $\dim V = 2$. By the basis theorem, it suffices to find any two noncollinear vectors in V. We write two linear combinations of the four given spanning vectors, chosen at random:

$$w_1 = \begin{pmatrix}1\\-2\\2\end{pmatrix} + \begin{pmatrix}2\\-3\\4\end{pmatrix} = \begin{pmatrix}3\\-5\\6\end{pmatrix} \qquad w_2 = -\begin{pmatrix}2\\-3\\4\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0\\4\\0\end{pmatrix} = \begin{pmatrix}-2\\5\\-4\end{pmatrix}.$$

Since w_1, w_2 are not collinear, $\mathcal{B} = \{w_1, w_2\}$ is a basis for V.

Use this link to view the online demo

A picture of the plane V and its basis $\mathcal{B} = \{w_1, w_2\}$.

Example (Finding a basis by inspection)**.** Find a basis for the plane

$$V = \left\{\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \,\middle|\, x_1 + x_2 = x_3\right\}$$

by inspection. (This plane is expressed in set builder notation.)

Solution. First note that V is the null space of the matrix $\begin{pmatrix}1 & 1 & -1\end{pmatrix}$; this matrix is in reduced row echelon form and has two free variables, so V is indeed a plane. We write down two vectors satisfying $x_1 + x_2 = x_3$:

$$v_1 = \begin{pmatrix}1\\0\\1\end{pmatrix} \qquad v_2 = \begin{pmatrix}0\\1\\1\end{pmatrix}.$$

Since v_1 and v_2 are not collinear, they are linearly independent; since $\dim(V) = 2$, the basis theorem implies that $\{v_1, v_2\}$ is a basis for V.

2.8 Bases as Coordinate Systems

Objectives

1. Learn to view a basis as a coordinate system on a subspace.
2. *Recipes:* compute the $\mathcal{B}$-coordinates of a vector, compute the usual coordinates of a vector from its $\mathcal{B}$-coordinates.
3. *Picture:* the $\mathcal{B}$-coordinates of a vector using its location on a nonstandard coordinate grid.
4. *Vocabulary word:* $\mathcal{B}$-**coordinates**.

In this section, we interpret a basis of a subspace V as a *coordinate system* on V, and we learn how to write a vector in V in that coordinate system.

Fact. *If $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ is a basis for a subspace V, then any vector x in V can be written as a linear combination*

$$x = c_1v_1 + c_2v_2 + \cdots + c_mv_m$$

in exactly one way.

Proof. Recall that to say $\mathcal{B}$ is a *basis* for V means that $\mathcal{B}$ spans V and $\mathcal{B}$ is linearly independent. Since $\mathcal{B}$ spans V, we can write any x in V as a linear combination of $v_1, v_2, \ldots, v_m$. For uniqueness, suppose that we had two such expressions:

$$\begin{aligned} x &= c_1v_1 + c_2v_2 + \cdots + c_mv_m \\ x &= c_1'v_1 + c_2'v_2 + \cdots + c_m'v_m. \end{aligned}$$

Subtracting the first equation from the second yields

$$0 = x - x = (c_1 - c_1')v_1 + (c_2 - c_2')v_2 + \cdots + (c_m - c_m')v_m.$$

Since $\mathcal{B}$ is linearly independent, the only solution to the above equation is the trivial solution: all the coefficients must be zero. It follows that $c_i - c_i'$ for all i, which proves that $c_1 = c_1'$, $c_2 = c_2'$, ..., $c_m = c_m'$. ☐

Example. Consider the standard basis of $\mathbf{R}^3$ from this example in Section 2.7:

$$e_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad e_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad e_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}.$$

According to the above fact, every vector in $\mathbf{R}^3$ can be written as a linear combination of e_1, e_2, e_3, with unique coefficients. For example,

$$v = \begin{pmatrix}3\\5\\-2\end{pmatrix} = 3\begin{pmatrix}1\\0\\0\end{pmatrix} + 5\begin{pmatrix}0\\1\\0\end{pmatrix} - 2\begin{pmatrix}0\\0\\1\end{pmatrix} = 3e_1 + 5e_2 - 2e_3.$$

In this case, *the coordinates of v are exactly the coefficients of* e_1, e_2, e_3.

What exactly are coordinates, anyway? One way to think of coordinates is that they give directions for how to get to a certain point from the origin. In the above example, the linear combination $3e_1 + 5e_2 - 2e_3$ can be thought of as the following list of instructions: start at the origin, travel 3 units north, then travel 5 units east, then 2 units down.

Definition. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ be a basis of a subspace V, and let

$$x = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m$$

be a vector in V. The coefficients $c_1, c_2, \ldots, c_m$ are the **coordinates of x with respect to $\mathcal{B}$**. The **$\mathcal{B}$-coordinate vector of x** is the vector

$$[x]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} \quad \text{in } \mathbf{R}^m.$$

If we change the basis, then we can still give instructions for how to get to the point $(3, 5, -2)$, but the instructions will be different. Say for example we take the basis

$$v_1 = e_1 + e_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

We can write $(3, 5, -2)$ in this basis as $3v_1 + 2v_2 - 2v_3$. In other words: start at the origin, travel northeast 3 times as far as v_1, then 2 units east, then 2 units down. In this situation, we can say that "3 is the v_1-coordinate of $(3, 5, -2)$, 2 is the v_2-coordinate of $(3, 5, -2)$, and -2 is the v_3-coordinate of $(3, 5, -2)$."

The above definition gives a way of using $\mathbf{R}^m$ to *label* the points of a subspace of dimension m: a point is simply labeled by its $\mathcal{B}$-coordinate vector. For instance, if we choose a basis for a plane, we can label the points of that plane with the points of $\mathbf{R}^2$.

Example (A nonstandard coordinate system on $\mathbf{R}^2$)**.** Define

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathcal{B} = \{v_1, v_2\}.$$

1. Verify that $\mathcal{B}$ is a basis for $\mathbf{R}^2$.
2. If $[w]_{\mathcal{B}} = \binom{1}{2}$, then what is w?
3. Find the $\mathcal{B}$-coordinates of $v = \binom{5}{3}$.

Solution.

1. By the basis theorem in Section 2.7, any two linearly independent vectors form a basis for $\mathbf{R}^2$. Clearly v_1, v_2 are not multiples of each other, so they are linearly independent.

2. To say $[w]_{\mathcal{B}} = \binom{1}{2}$ means that 1 is the v_1-coordinate of w, and that 2 is the v_2-coordinate:

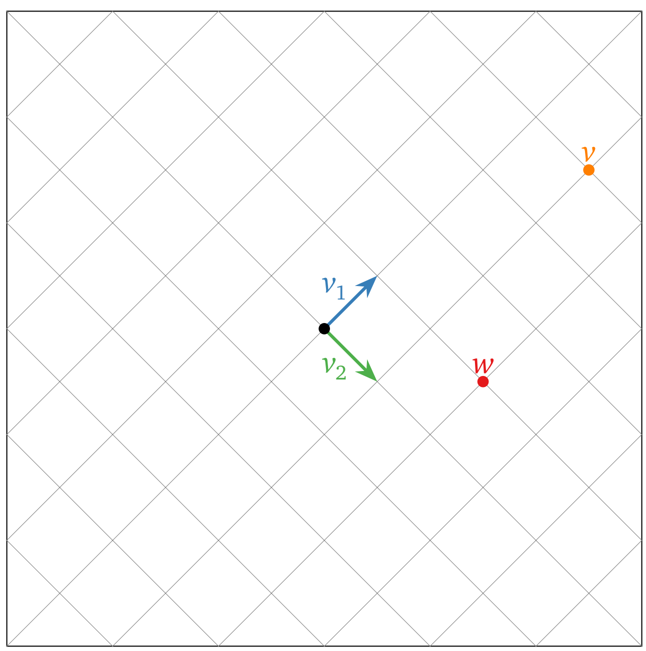
$$w = v_1 + 2v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

3. We have to solve the vector equation $v = c_1v_1 + c_2v_2$ in the unknowns c_1, c_2. We form an augmented matrix and row reduce:

$$\left(\begin{array}{cc|c} 1 & 1 & 5 \\ 1 & -1 & 3 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 4 \\ 0 & 1 & 1 \end{array}\right).$$

   We have $c_1 = 4$ and $c_2 = 1$, so $v = 4v_1 + v_2$ and $[v]_{\mathcal{B}} = \binom{4}{1}$.

In the following picture, we indicate the coordinate system defined by $\mathcal{B}$ by drawing lines parallel to the "v_1-axis" and "v_2-axis". Using this grid it is easy to see that the $\mathcal{B}$-coordinates of v are $\binom{5}{1}$, and that the $\mathcal{B}$-coordinates of w are $\binom{1}{2}$.



This picture could be the grid of streets in Palo Alto, California. Residents of Palo Alto refer to northwest as "north" and to northeast as "east". There is a reason for this: the old road to San Francisco is called El Camino Real, and that road runs from the southeast to the northwest in Palo Alto. So when a Palo Alto resident says "go south two blocks and east one block", they are giving directions from the origin to the Whole Foods at w.

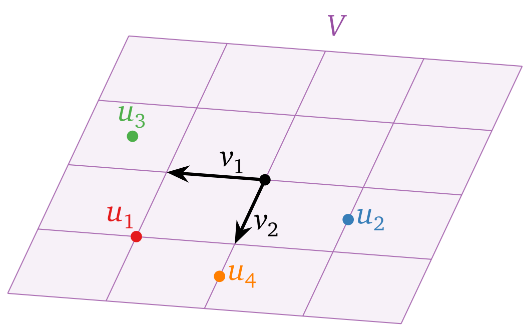
Use this link to view the online demo

A picture of the basis $\mathcal{B} = \{v_1, v_2\}$ of $\mathbf{R}^2$. The grid indicates the coordinate system defined by the basis $\mathcal{B}$; one set of lines measures the v_1-coordinate, and the other set measures the v_2-coordinate. Use the sliders to find the $\mathcal{B}$-coordinates of w.

Example. Let

$$v_1 = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

These form a basis $\mathcal{B}$ for a plane $V = \text{Span}\{v_1, v_2\}$ in $\mathbf{R}^3$. We indicate the coordinate system defined by $\mathcal{B}$ by drawing lines parallel to the "v_1-axis" and "v_2-axis":



We can see from the picture that the v_1-coordinate of u_1 is equal to 1, as is the v_2-coordinate, so $[u_1]_{\mathcal{B}} = \binom{1}{1}$. Similarly, we have

$$[u_2]_{\mathcal{B}} = \begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix} \qquad [u_3]_{\mathcal{B}} = \begin{pmatrix} \frac{3}{2} \\ -\frac{1}{2} \end{pmatrix} \qquad [u_4]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \frac{3}{2} \end{pmatrix}.$$

Use this link to view the online demo

Left: the $\mathcal{B}$-coordinates of a vector x. Right: the vector x. The violet grid on the right is a picture of the coordinate system defined by the basis $\mathcal{B}$; one set of lines measures the v_1-coordinate, and the other set measures the v_2-coordinate. Drag the heads of the vectors x and $[x]_{\mathcal{B}}$ to understand the correspondence between x and its $\mathcal{B}$-coordinate vector.

Example (A coordinate system on a plane)**.** Define

$$v_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \; v_2 = \begin{pmatrix}1\\1\\1\end{pmatrix}, \qquad \mathcal{B} = \{v_1, v_2\}, \quad V = \mathrm{Span}\{v_1, v_2\}.$$

1. Verify that $\mathcal{B}$ is a basis for V.
2. If $[w]_{\mathcal{B}} = \binom{5}{2}$, then what is w?
3. Find the $\mathcal{B}$-coordinates of $v = \begin{pmatrix}5\\3\\5\end{pmatrix}$.

Solution.

1. We need to verify that $\mathcal{B}$ spans V, and that it is linearly independent. By definition, V is the span of $\mathcal{B}$; since v_1 and v_2 are not multiples of each other, they are linearly independent. This shows in particular that V is a *plane*.
2. To say $[w]_{\mathcal{B}} = \binom{5}{2}$ means that 5 is the v_1-coordinate of w, and that 2 is the v_2-coordinate:
$$w = 5v_1 + 2v_2 = 5\begin{pmatrix}1\\0\\1\end{pmatrix} + 2\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}7\\2\\7\end{pmatrix}.$$
3. We have to solve the vector equation $v = c_1v_1 + c_2v_2$ in the unknowns c_1, c_2. We form an augmented matrix and row reduce:
$$\left(\begin{array}{cc|c}1 & 1 & 5\\0 & 1 & 3\\1 & 1 & 5\end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c}1 & 0 & 2\\0 & 1 & 3\\0 & 0 & 0\end{array}\right).$$
We have $c_1 = 2$ and $c_2 = 3$, so $v = 2v_1 + 3v_2$ and $[v]_{\mathcal{B}} = \binom{2}{3}$.

Use this link to view the online demo

A picture of the plane V and the basis $\mathcal{B} = \{v_1, v_2\}$. The violet grid is a picture of the coordinate system defined by the basis $\mathcal{B}$; one set of lines measures the v_1-coordinate, and the other set measures the v_2-coordinate. Use the sliders to find the $\mathcal{B}$-coordinates of v.

Example (A coordinate system on another plane)**.** Define

$$v_1 = \begin{pmatrix}2\\3\\2\end{pmatrix}, \; v_2 = \begin{pmatrix}-1\\1\\1\end{pmatrix}, \; v_3 = \begin{pmatrix}2\\8\\6\end{pmatrix}, \qquad V = \mathrm{Span}\{v_1, v_2, v_3\}.$$

1. Find a basis $\mathcal{B}$ for V.

2. Find the $\mathcal{B}$-coordinates of $x = \begin{pmatrix} 4 \\ 11 \\ 8 \end{pmatrix}$.

Solution.

1. We write V as the column space of a matrix A, then row reduce to find the pivot columns, as in this example in Section 2.7.

$$A = \begin{pmatrix} 2 & -1 & 2 \\ 3 & 1 & 8 \\ 2 & 1 & 6 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

   The first two columns are pivot columns, so we can take $\mathcal{B} = \{v_1, v_2\}$ as our basis for V.

2. We have to solve the vector equation $x = c_1v_1 + c_2v_2$. We form an augmented matrix and row reduce:

$$\left(\begin{array}{cc|c} 2 & -1 & 4 \\ 3 & 1 & 11 \\ 2 & 1 & 8 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right).$$

   We have $c_1 = 3$ and $c_2 = 2$, so $x = 3v_1 + 2v_2$, and thus $[x]_{\mathcal{B}} = \binom{3}{2}$.

Use this link to view the online demo

A picture of the plane V and the basis $\mathcal{B} = \{v_1, v_2\}$. The violet grid is a picture of the coordinate system defined by the basis $\mathcal{B}$; one set of lines measures the v_1-coordinate, and the other set measures the v_2-coordinate. Use the sliders to find the $\mathcal{B}$-coordinates of x.

Recipes: $\mathcal{B}$-coordinates. If $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ is a basis for a subspace V and x is in V, then

$$[x]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} \quad \text{means} \quad x = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m.$$

Finding the $\mathcal{B}$-coordinates of x means solving the vector equation

$$x = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m$$

in the unknowns $c_1, c_2, \ldots, c_m$. This generally means row reducing the augmented matrix

$$\left(\begin{array}{cccc|c} | & | & & | & | \\ v_1 & v_2 & \cdots & v_m & x \\ | & | & & | & | \end{array}\right).$$

Remark. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ be a basis of a subspace V. Finding the $\mathcal{B}$-coordinates of a vector x means solving the vector equation

$$x = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m.$$

If x is *not* in V, then this equation has no solution, as x is not in $V = \text{Span}\{v_1, v_2, \ldots, v_m\}$. In other words, the above equation is *inconsistent* when x is not in V.

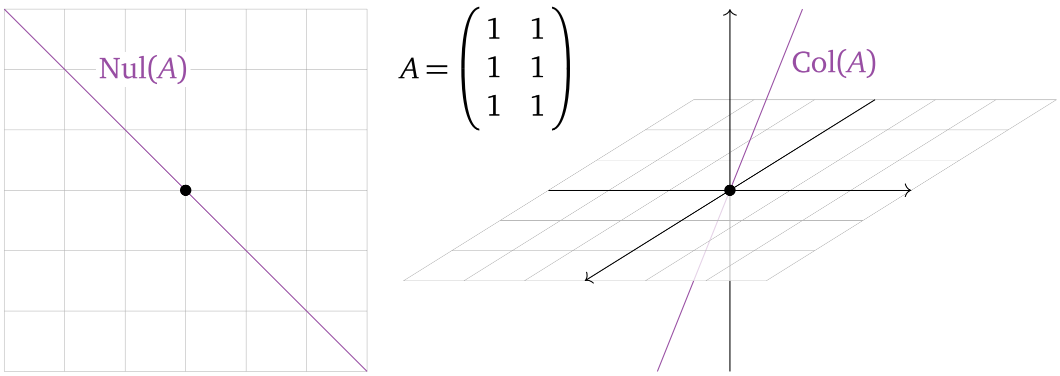
2.9 The Rank Theorem

Objectives

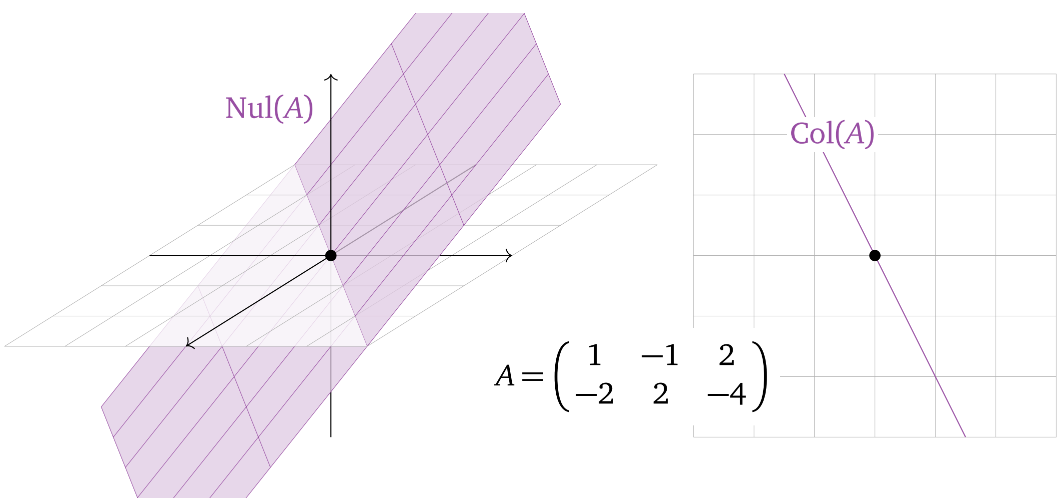
1. Learn to understand and use the rank theorem.
2. *Picture:* the rank theorem.
3. *Theorem:* rank theorem.
4. *Vocabulary words:* **rank**, **nullity**.

In this section we present the rank theorem, which is the culmination of all of the work we have done so far.

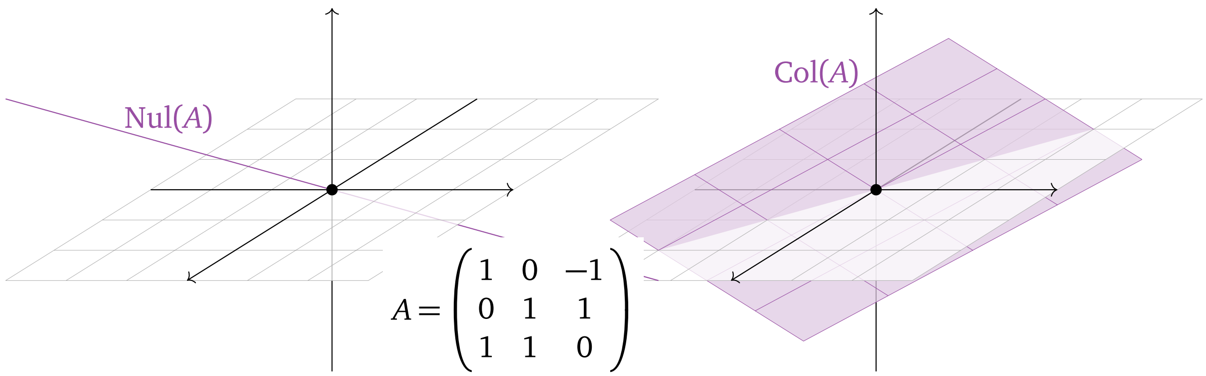
The reader may have observed a relationship between the column space and the null space of a matrix. In this example in Section 2.6, the column space and the null space of a 3×2 matrix are both lines, in $\mathbf{R}^2$ and $\mathbf{R}^3$, respectively:



In this example in Section 2.4, the null space of the 2×3 matrix $\begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{pmatrix}$ is a plane in $\mathbf{R}^3$, and the column space the line in $\mathbf{R}^2$ spanned by $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$:



In this example in Section 2.4, the null space of a 3×3 matrix is a line in $\mathbf{R}^3$, and the column space is a plane in $\mathbf{R}^3$:



$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

In all examples, the dimension of the column space plus the dimension of the null space is equal to the number of *columns* of the matrix. This is the content of the rank theorem.

Definition. The **rank** of a matrix A, written $\text{rank}(A)$, is the dimension of the column space $\text{Col}(A)$.

The **nullity** of a matrix A, written $\text{nullity}(A)$, is the dimension of the null space $\text{Nul}(A)$.

The rank of a matrix A gives us important information about the solutions to $Ax = b$. Recall from this note in Section 2.3 that $Ax = b$ is consistent exactly when b is in the span of the columns of A, in other words when b is in the column space of A. Thus, $\text{rank}(A)$ is the dimension of the set of b with the property that $Ax = b$ is consistent.

We know that the rank of A is equal to the number of pivot columns (see this theorem in Section 2.7), and the nullity of A is equal to the number of free variables (see this theorem in Section 2.7), which is the number of columns without pivots. To summarize:

$$\begin{aligned} \text{rank}(A) = \dim \text{Col}(A) &= \text{the number of columns with pivots} \\ \text{nullity}(A) = \dim \text{Nul}(A) &= \text{the number of free variables} \\ &= \text{the number of columns without pivots} \end{aligned}$$

Clearly

$$\#(\text{columns with pivots}) + \#(\text{columns without pivots}) = \#(\text{columns}),$$

so we have proved the following theorem.

Rank Theorem. *If A is a matrix with n columns, then*

$$\text{rank}(A) + \text{nullity}(A) = n.$$

In other words, for any consistent system of linear equations,

$$(\text{dim of column span}) + (\text{dim of solution set}) = (\text{number of variables}).$$

The rank theorem theorem is really the culmination of this chapter, as it gives a strong relationship between the null space of a matrix (the solution set of $Ax = 0$) with the column space (the set of vectors b making $Ax = b$ consistent), our two primary objects of interest. The more freedom we have in choosing x the less freedom we have in choosing b and vice versa.

Example (Rank and nullity)**.** Here is a concrete example of the rank theorem and the interplay between the degrees of freedom we have in choosing x and b in a matrix equation $Ax = b$.

Consider the matrices

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If we multiply a vector (x, y, z) in $\mathbf{R}^3$ by A and B we obtain the vectors $Ax = (x, y, 0)$ and $Bx = (0, 0, z)$. So we can think of multiplication by A as giving the latitude and longitude of a point in $\mathbf{R}^3$ and we can think of multiplication by B as giving the height of a point in $\mathbf{R}^3$. The rank of A is 2 and the nullity is 1. Similarly, the rank of B is 1 and the nullity is 2.

These facts have natural interpretations. For the matrix A: the set of all latitudes and longitudes in $\mathbf{R}^3$ is a plane, and the set of points with the same latitude and longitude in $\mathbf{R}^3$ is a line; and for the matrix B: the set of all heights in $\mathbf{R}^3$ is a line, and the set of points at a given height in $\mathbf{R}^3$ is a plane. As the rank theorem tells us, we "trade off" having more choices for x for having more choices for b, and vice versa.

The rank theorem is a prime example of how we use the theory of linear algebra to say something qualitative about a system of equations without ever solving it. This is, in essence, the power of the subject.

Example (The rank is 2 and the nullity is 2)**.** Consider the following matrix and its reduced row echelon form:

$$A = \begin{pmatrix} 1 & 2 & 0 & -1 \\ -2 & -3 & 4 & 5 \\ 2 & 4 & 0 & -2 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -8 & -7 \\ 0 & 1 & 4 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

basis of Col(A) (first two columns of A)   free variables (last two columns of the RREF)

A basis for Col(A) is given by the pivot columns:

$$\left\{\begin{pmatrix}1\\-2\\2\end{pmatrix}, \begin{pmatrix}2\\-3\\4\end{pmatrix}\right\},$$

so $\text{rank}(A) = \dim \text{Col}(A) = 2$.

Since there are two free variables x_3, x_4, the null space of A has two vectors (see this theorem in Section 2.7):

$$\left\{\begin{pmatrix}8\\-4\\1\\0\end{pmatrix}, \begin{pmatrix}7\\-3\\0\\1\end{pmatrix}\right\},$$

so $\text{nullity}(A) = 2$.

In this case, the rank theorem says that $2 + 2 = 4$, where 4 is the number of columns.

Interactive: Rank is 1, nullity is 2.

Use this link to view the online demo

This 3×3 matrix has rank 1 and nullity 2. The violet plane on the left is the null space, and the violet line on the right is the column space.

Interactive: Rank is 2, nullity is 1.

Use this link to view the online demo

This 3×3 matrix has rank 2 and nullity 1. The violet line on the left is the null space, and the violet plane on the right is the column space.

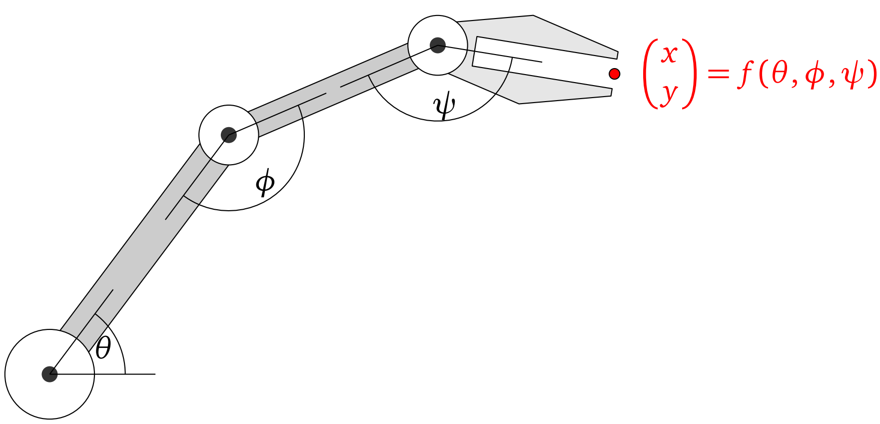
Chapter 3

Linear Transformations and Matrix Algebra

Primary Goal. Learn about linear transformations and their relationship to matrices.

In practice, one is often lead to ask questions about the geometry of a *transformation*: a function that takes an input and produces an output. This kind of question can be answered by linear algebra if the transformation can be expressed by a matrix.

Example. Suppose you are building a robot arm with three joints that can move its hand around a plane, as in the following picture.



Define a transformation f as follows: $f(\theta, \phi, \psi)$ is the (x, y) position of the hand when the joints are rotated by angles θ, ϕ, ψ, respectively. The output of f tells you where the hand will be on the plane when the joints are set at the given input angles.

Unfortunately, this kind of function does not come from a matrix, so one cannot use linear algebra to answer questions about this function. In fact, these functions are rather complicated; their study is the subject of inverse kinematics.

In this chapter, we will be concerned with the relationship between matrices and transformations. In Section 3.1, we will consider the equation $b = Ax$ as a function with independent variable x and dependent variable b, and we draw pictures accordingly. We spend some time studying transformations in the abstract, and asking questions about a transformation, like whether it is one-to-one and/or onto (Section 3.2). In Section 3.3 we will answer the question: "when exactly can a transformation be expressed by a matrix?" We then present matrix multiplication as a special case of composition of transformations (Section 3.4). This leads to the study of *matrix algebra*: that is, to what extent one can do arithmetic with matrices in the place of numbers. With this in place, we learn to solve matrix equations by *dividing* by a matrix in Section 3.5.

3.1 Matrix Transformations

Objectives

1. Learn to view a matrix geometrically as a function.
2. Learn examples of matrix transformations: reflection, dilation, rotation, shear, projection.
3. Understand the vocabulary surrounding transformations: domain, codomain, range.
4. Understand the domain, codomain, and range of a matrix transformation.
5. *Pictures:* common matrix transformations.
6. *Vocabulary words:* **transformation** / **function**, **domain**, **codomain**, **range**, **identity transformation**, **matrix transformation**.

In this section we learn to understand matrices geometrically as *functions*, or *transformations*. We briefly discuss transformations in general, then specialize to matrix transformations, which are transformations that come from matrices.

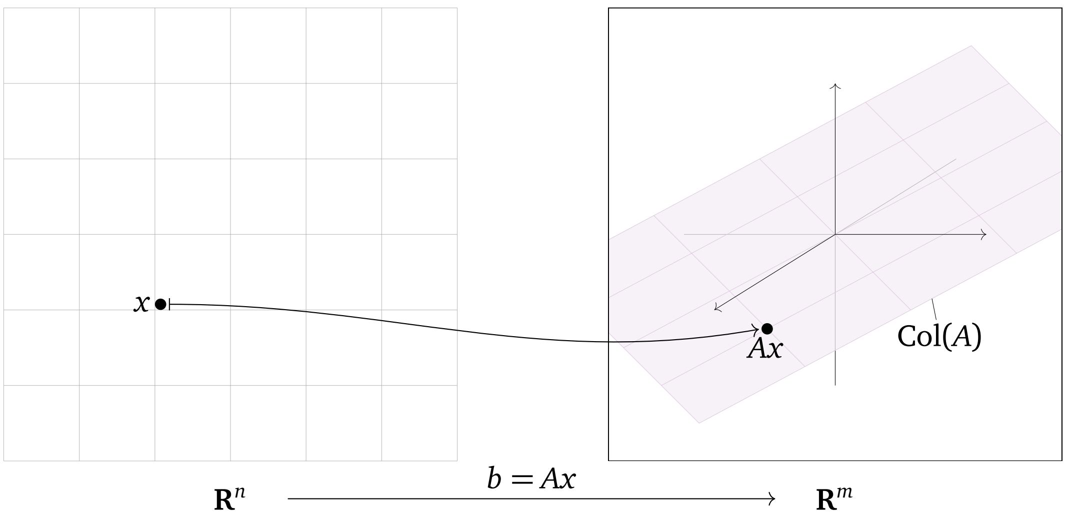
3.1.1 Matrices as Functions

Informally, a *function* is a rule that accepts inputs and produces outputs. For instance, $f(x) = x^2$ is a function that accepts one number x as its input, and outputs the square of that number: $f(2) = 4$. In this subsection, we interpret matrices as functions.

Let A be a matrix with m rows and n columns. Consider the matrix equation $b = Ax$ (we write it this way instead of $Ax = b$ to remind the reader of the notation $y = f(x)$). If we vary x, then b will also vary; in this way, we think of A as a function with independent variable x and dependent variable b.

- The independent variable (the input) is x, which is a vector in $\mathbf{R}^n$.
- The dependent variable (the output) is b, which is a vector in $\mathbf{R}^m$.

The set of all possible output vectors are the vectors b such that $Ax = b$ has some solution; this is the same as the column space of A by this note in Section 2.3.



Interactive: A 2×3 matrix.

Use this link to view the online demo

A picture of a 2×3 matrix, regarded as a function. The input vector is x, which is a vector in $\mathbf{R}^3$, and the output vector is $b = Ax$, which is a vector in $\mathbf{R}^2$. The violet line on the right is the column space; as you vary x, the output b is constrained to lie on this line.

Interactive: A 3×2 matrix.

Use this link to view the online demo

A picture of a 3×2 matrix, regarded as a function. The input vector is x, which is a vector in $\mathbf{R}^2$, and the output vector is $b = Ax$, which is a vector in $\mathbf{R}^3$. The violet plane on the right is the column space; as you vary x, the output b is constrained to lie on this plane.

Example (Projection onto the xy-plane)**.** Let

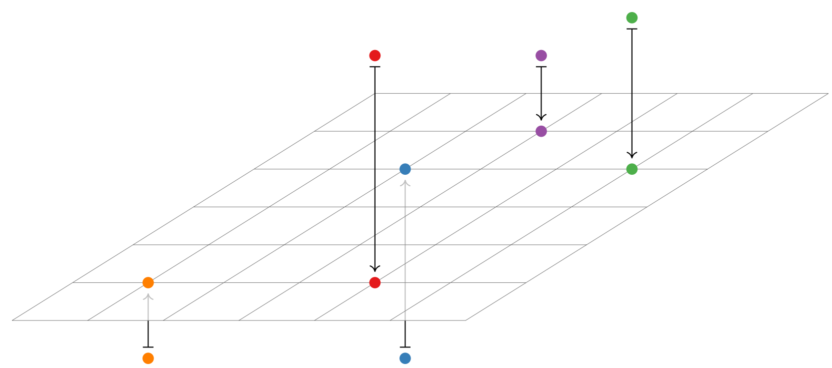
$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Describe the function $b = Ax$ geometrically.

Solution. In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbf{R}^3$. First we multiply A by a vector to see what it does:

$$A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.$$

Multiplication by A simply sets the z-coordinate equal to zero: it *projects vertically onto the xy-plane*.



Use this link to view the online demo

Multiplication by the matrix A projects a vector onto the xy-plane. Move the input vector x to see how the output vector b changes.

Example (Reflection)**.** Let

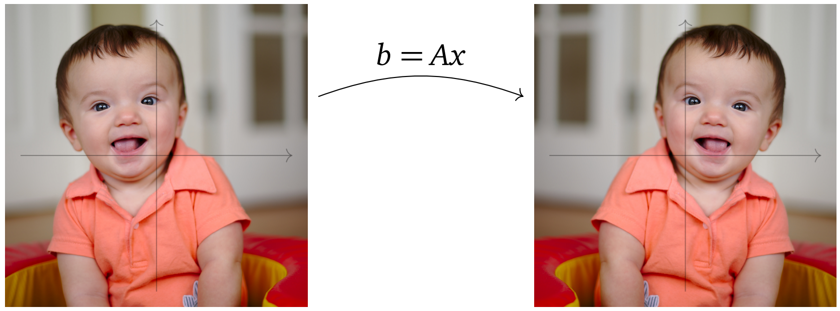
$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Describe the function $b = Ax$ geometrically.

Solution. In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbf{R}^2$. First we multiply A by a vector to see what it does:

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}.$$

Multiplication by A negates the x-coordinate: it *reflects over the y-axis*.



Use this link to view the online demo

Multiplication by the matrix A reflects over the y-axis. Move the input vector x to see how the output vector b changes.

Example (Dilation)**.** Let

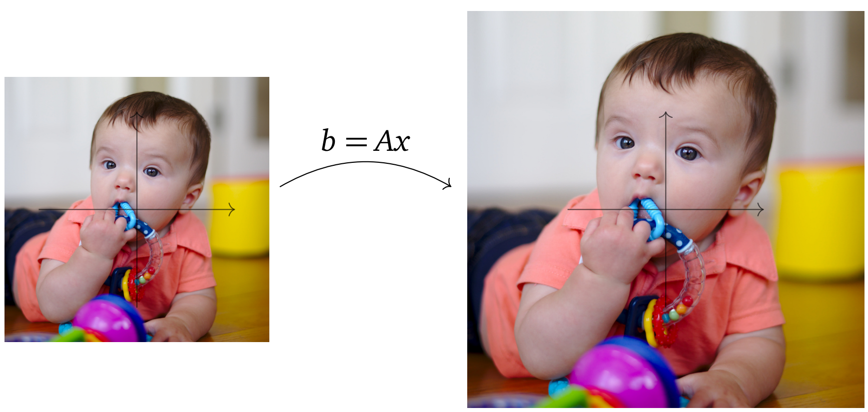
$$A = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix}.$$

Describe the function $b = Ax$ geometrically.

Solution. In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbf{R}^2$. First we multiply A by a vector to see what it does:

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1.5x \\ 1.5y \end{pmatrix} = 1.5\begin{pmatrix} x \\ y \end{pmatrix}.$$

Multiplication by A is the same as scalar multiplication by 1.5: it *scales* or *dilates* the plane by a factor of 1.5.



Use this link to view the online demo

Multiplication by the matrix A dilates the plane by a factor of 1.5*. Move the input vector x to see how the output vector b changes.*

Example (Identity)**.** Let

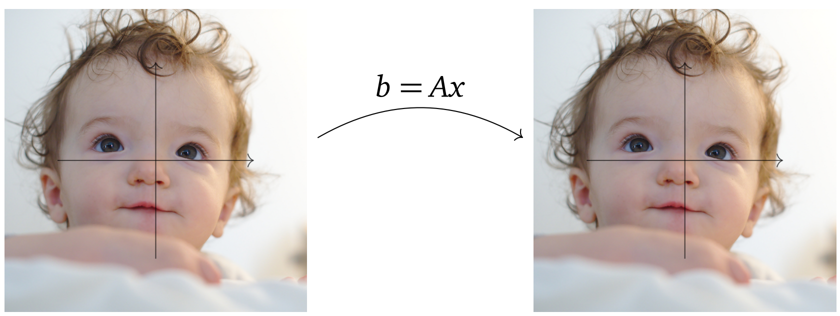
$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Describe the function $b = Ax$ geometrically.

Solution. In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbf{R}^2$. First we multiply A by a vector to see what it does:

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Multiplication by A does not change the input vector at all: it is the *identity transformation which does nothing*.



Use this link to view the online demo

Multiplication by the matrix A does not move the vector x: that is, $b = Ax = x$. Move the input vector x to see how the output vector b changes.

Example (Rotation)**.** Let

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

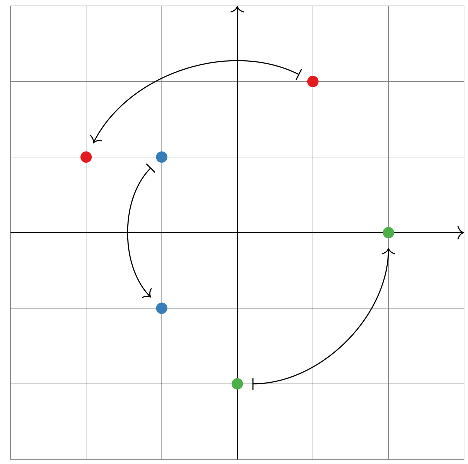
Describe the function $b = Ax$ geometrically.

Solution. In the equation $Ax = b$, the input vector x and the output vector b are both in $\mathbf{R}^2$. First we multiply A by a vector to see what it does:

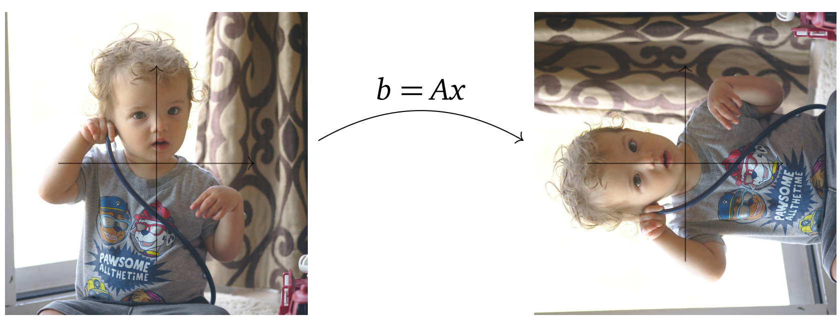
$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix}.$$

We substitute a few test points in order to understand the geometry of the transformation:

$$A\begin{pmatrix}1\\2\end{pmatrix}=\begin{pmatrix}-2\\1\end{pmatrix}$$
$$A\begin{pmatrix}-1\\1\end{pmatrix}=\begin{pmatrix}-1\\-1\end{pmatrix}$$
$$A\begin{pmatrix}0\\-2\end{pmatrix}=\begin{pmatrix}2\\0\end{pmatrix}$$



Multiplication by A is *counterclockwise rotation by* $90°$.



Use this link to view the online demo

Multiplication by the matrix A rotates the vector x counterclockwise by $90°$. Move the input vector x to see how the output vector b changes.

Example (Shear)**.** Let

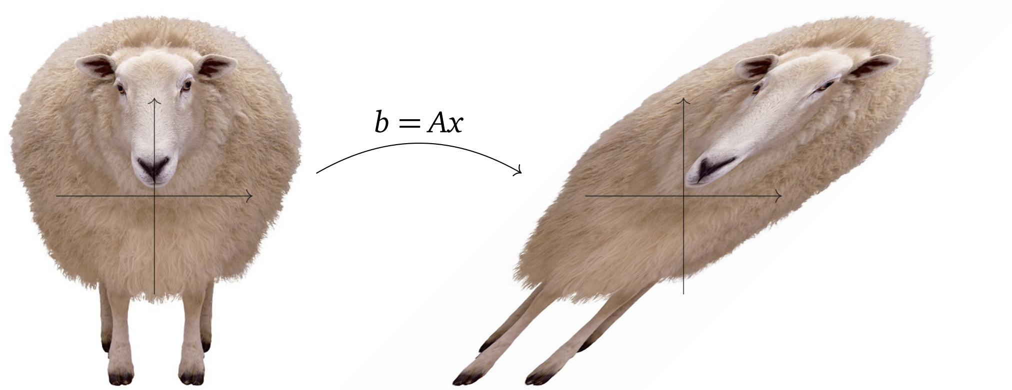
$$A=\begin{pmatrix}1&1\\0&1\end{pmatrix}.$$

Describe the function $b=Ax$ geometrically.

Solution. In the equation $Ax=b$, the input vector x and the output vector b are both in $\mathbf{R}^2$. First we multiply A by a vector to see what it does:

$$A\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x+y\\y\end{pmatrix}.$$

Multiplication by A adds the y-coordinate to the x-coordinate; this is called a *shear in the x-direction*.



Use this link to view the online demo

Multiplication by the matrix A adds the y-coordinate to the x-coordinate. Move the input vector x to see how the output vector b changes.

3.1.2 Transformations

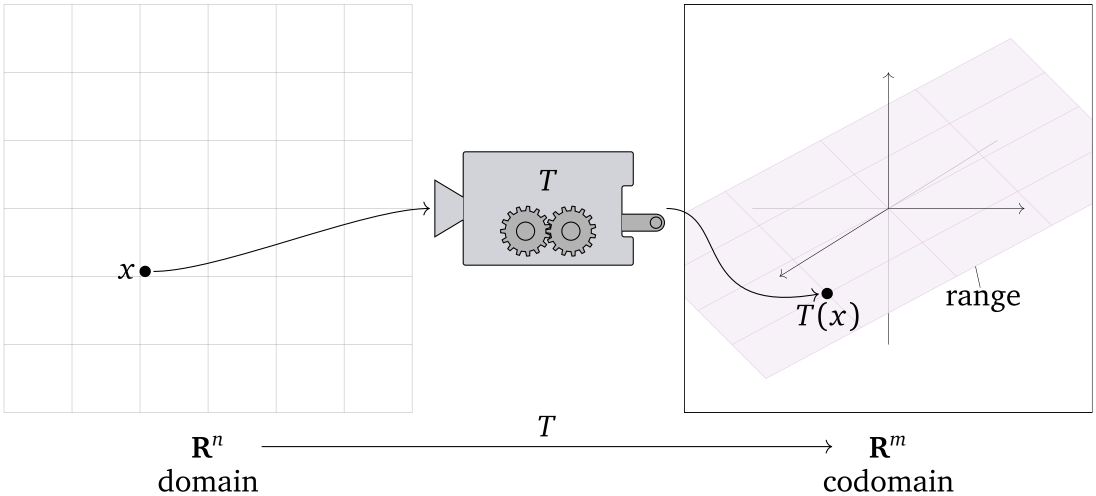
At this point it is convenient to fix our ideas and terminology regarding functions, which we will call *transformations* in this book. This allows us to systematize our discussion of matrices as functions.

Definition. A **transformation** from $\mathbf{R}^n$ to $\mathbf{R}^m$ is a rule T that assigns to each vector x in $\mathbf{R}^n$ a vector $T(x)$ in $\mathbf{R}^m$.

- $\mathbf{R}^n$ is called the **domain** of T.
- $\mathbf{R}^m$ is called the **codomain** of T.
- For x in $\mathbf{R}^n$, the vector $T(x)$ in $\mathbf{R}^m$ is the **image** of x under T.
- The set of all images $\{T(x) \mid x \text{ in } \mathbf{R}^n\}$ is the **range** of T.

The notation $T: \mathbf{R}^n \longrightarrow \mathbf{R}^m$ means "T is a transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$."

It may help to think of T as a "machine" that takes x as an input, and gives you $T(x)$ as the output.



The points of the domain $\mathbf{R}^n$ are the *inputs* of T: this simply means that it makes sense to evaluate T on vectors with n entries, i.e., lists of n numbers. Likewise, the points of the codomain $\mathbf{R}^m$ are the *outputs* of T: this means that the result of evaluating T is always a vector with m entries.

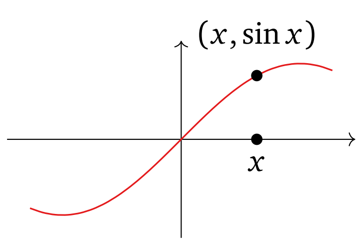
The *range* of T is the set of all vectors in the codomain that actually arise as outputs of the function T, for some input. In other words, the range is all vectors b in the codomain such that $T(x) = b$ has a solution x in the domain.

Example (A Function of one variable)**.** Most of the functions you may have seen previously have domain and codomain equal to $\mathbf{R} = \mathbf{R}^1$. For example,

$$\sin\colon \mathbf{R} \longrightarrow \mathbf{R} \qquad \sin(x) = \begin{pmatrix} \text{the length of the opposite} \\ \text{edge over the hypotenuse of} \\ \text{a right triangle with angle } x \\ \text{in radians} \end{pmatrix}.$$

Notice that we have defined sin by a rule: a function is defined by specifying what the output of the function is for any possible input.

You may be used to thinking of such functions in terms of their graphs:



In this case, the horizontal axis is the domain, and the vertical axis is the codomain. This is useful when the domain and codomain are $\mathbf{R}$, but it is hard to do when, for instance, the domain is $\mathbf{R}^2$ and the codomain is $\mathbf{R}^3$. The graph of such a function is a subset of $\mathbf{R}^5$, which is difficult to visualize. For this reason, we will rarely graph a transformation.

Note that the *range* of sin is the interval $[-1,1]$: this is the set of all possible outputs of the sin function.

Example (Functions of several variables)**.** Here is an example of a function from $\mathbf{R}^2$ to $\mathbf{R}^3$:

$$f\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}x+y\\\cos(y)\\y-x^2\end{pmatrix}.$$

The inputs of f each have two entries, and the outputs have three entries. In this case, we have defined f by a formula, so we evaluate f by substituting values for the variables:

$$f\begin{pmatrix}2\\3\end{pmatrix}=\begin{pmatrix}2+3\\\cos(3)\\3-2^2\end{pmatrix}=\begin{pmatrix}5\\\cos(3)\\-1\end{pmatrix}.$$

Here is an example of a function from $\mathbf{R}^3$ to $\mathbf{R}^3$:

$$f(v)=\begin{pmatrix}\text{the counterclockwise rotation}\\\text{of } v \text{ by an angle of } 42^\circ \text{ about}\\\text{the } z\text{-axis}\end{pmatrix}.$$

In other words, f takes a vector with three entries, then rotates it; hence the ouput of f also has three entries. In this case, we have defined f by a geometric rule.

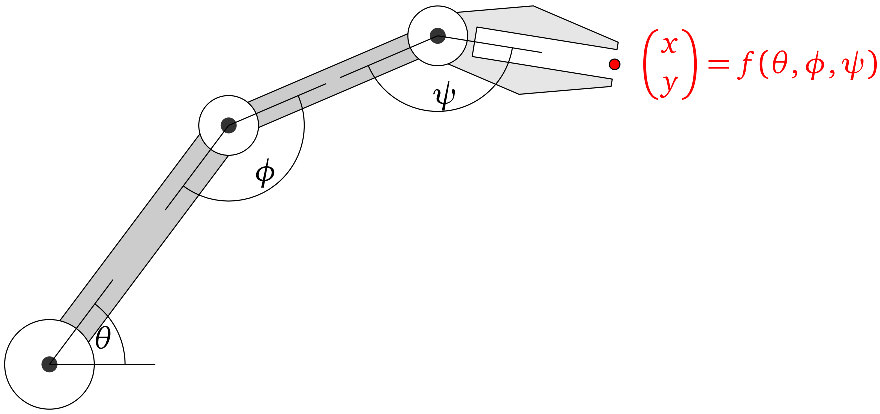
Definition. The **identity transformation** $\mathrm{Id}_{\mathbf{R}^n}\colon \mathbf{R}^n \to \mathbf{R}^n$ is the transformation defined by the rule

$$\mathrm{Id}_{\mathbf{R}^n}(x)=x \qquad \text{for all } x \text{ in } \mathbf{R}^n.$$

In other words, the identity transformation does not move its input vector: the output is the same as the input. Its domain and codomain are both $\mathbf{R}^n$, and its range is $\mathbf{R}^n$ as well, since every vector in $\mathbf{R}^n$ is the output of itself.

Example (A real-word transformation: robotics)**.** The definition of *transformation* and its associated vocabulary may seem quite abstract, but transformations are extremely common in real life. Here is an example from the fields of robotics and computer graphics.

Suppose you are building a robot arm with three joints that can move its hand around a plane, as in the following picture.



Define a transformation $f : \mathbf{R}^3 \to \mathbf{R}^2$ as follows: $f(\theta, \phi, \psi)$ is the (x, y) position of the hand when the joints are rotated by angles θ, ϕ, ψ, respectively. Evaluating f tells you where the hand will be on the plane when the joints are set at the given angles.

It is relatively straightforward to find a formula for $f(\theta, \phi, \psi)$ using some basic trigonometry. If you want the robot to fetch your coffee cup, however, you have to *find* the angles θ, ϕ, ψ that will put the hand at the position of your beverage. It is not at all obvious how to do this, and it is not even clear if the answer is unique! You can ask yourself: "which positions on the table can my robot arm reach?" or "what is the arm's range of motion?" This is the same as asking: "what is the range of f?"

Unfortunately, this kind of function does not come from a matrix, so one cannot use linear algebra to answer these kinds of questions. In fact, these functions are rather complicated; their study is the subject of inverse kinematics.

3.1.3 Matrix Transformations

Now we specialize the general notions and vocabulary from the previous subsection to the functions defined by matrices that we considered in the first subsection.

Definition. Let A be an $m \times n$ matrix. The **matrix transformation** associated to A is the transformation

$$T : \mathbf{R}^n \longrightarrow \mathbf{R}^m \qquad \text{defined by} \qquad T(x) = Ax.$$

This is the transformation that takes a vector x in $\mathbf{R}^n$ to the vector Ax in $\mathbf{R}^m$.

If A has n columns, then it only makes sense to multiply A by vectors with n entries. This is why the domain of $T(x) = Ax$ is $\mathbf{R}^n$. If A has n rows, then Ax has m entries for any vector x in $\mathbf{R}^n$; this is why the codomain of $T(x) = Ax$ is $\mathbf{R}^m$.

The definition of a matrix transformation T tells us how to evaluate T on any given vector: we multiply the input vector by a matrix. For instance, let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

and let $T(x) = Ax$ be the associated matrix transformation. Then

$$T\begin{pmatrix}-1\\-2\\-3\end{pmatrix} = A\begin{pmatrix}-1\\-2\\-3\end{pmatrix} = \begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}\begin{pmatrix}-1\\-2\\-3\end{pmatrix} = \begin{pmatrix}-14\\-32\end{pmatrix}.$$

Suppose that A has columns $v_1, v_2, \ldots, v_n$. If we multiply A by a general vector x, we get

$$Ax = \begin{pmatrix}|&|&&|\\v_1&v_2&\cdots&v_n\\|&|&&|\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix} = x_1v_1 + x_2v_2 + \cdots + x_nv_n.$$

This is just a general linear combination of $v_1, v_2, \ldots, v_n$. Therefore, the outputs of $T(x) = Ax$ are exactly the linear combinations of the columns of A: the *range* of T is the column space of A. See this note in Section 2.3.

Let A be an $m \times n$ matrix, and let $T(x) = Ax$ be the associated matrix transformation.

- The *domain* of T is $\mathbf{R}^n$, where n is the number of *columns* of A.
- The *codomain* of T is $\mathbf{R}^m$, where m is the number of *rows* of A.
- The *range* of T is the *column space* of A.

Interactive: A 2×3 matrix: reprise. Let

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

and define $T(x) = Ax$. The domain of T is $\mathbf{R}^3$, and the codomain is $\mathbf{R}^2$. The range of T is the column space; since all three columns are collinear, the range is a line in $\mathbf{R}^2$.

Use this link to view the online demo

A picture of the matrix transformation T. The input vector is x, which is a vector in $\mathbf{R}^3$, and the output vector is $b = T(x) = Ax$, which is a vector in $\mathbf{R}^2$. The violet line on the right is the range of T; as you vary x, the output b is constrained to lie on this line.

Interactive: A 3×2 matrix: reprise. Let

$$A = \begin{pmatrix}1&0\\0&1\\1&0\end{pmatrix},$$

and define $T(x) = Ax$. The domain of T is $\mathbf{R}^2$, and the codomain is $\mathbf{R}^3$. The range of T is the column space; since A has two columns which are not collinear, the range is a plane in $\mathbf{R}^3$.

Use this link to view the online demo

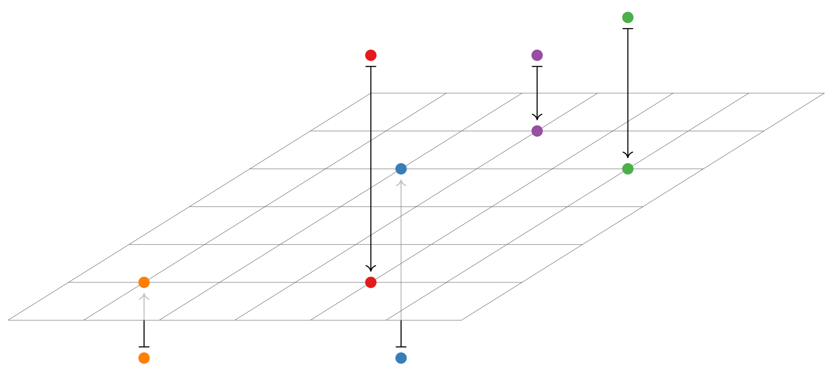
A picture of the matrix transformation T. The input vector is x, which is a vector in $\mathbf{R}^2$, and the output vector is $b = T(x) = Ax$, which is a vector in $\mathbf{R}^3$. The violet plane on the right is the range of T; as you vary x, the output b is constrained to lie on this plane.

Example (Projection onto the xy-plane: reprise)**.** Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and let $T(x) = Ax$. What are the domain, the codomain, and the range of T?

Solution. Geometrically, the transformation T projects a vector directly “down” onto the xy-plane in $\mathbf{R}^3$.



The inputs and outputs have three entries, so the domain and codomain are both $\mathbf{R}^3$. The possible outputs all lie on the xy-plane, and every point on the xy-plane is an output of T (with itself as the input), so the range of T is the xy-plane.

Be careful not to confuse the codomain with the range here. The range is a plane, but it is a plane *in* $\mathbf{R}^3$, so the codomain is still $\mathbf{R}^3$. The outputs of T all have three entries; the last entry is simply always zero.

In the case of an $n \times n$ square matrix, the domain and codomain of $T(x) = Ax$ are both $\mathbf{R}^n$. In this situation, one can regard T as *operating on* $\mathbf{R}^n$: it moves the vectors around in the same space.

Example (Matrix transformations of $\mathbf{R}^2$)**.** In the first subsection we discussed the transformations defined by several 2×2 matrices, namely:

$$\begin{aligned}
\text{Reflection:}\quad & A=\begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}\\
\text{Dilation:}\quad & A=\begin{pmatrix}1.5 & 0\\ 0 & 1.5\end{pmatrix}\\
\text{Identity:}\quad & A=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\\
\text{Rotation:}\quad & A=\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\\
\text{Shear:}\quad & A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}.
\end{aligned}$$

In each case, the associated matrix transformation $T(x)=Ax$ has domain and codomain equal to $\mathbf{R}^2$. The range is also $\mathbf{R}^2$, as can be seen geometrically (what is the input for a given output?), or using the fact that the columns of A are not collinear (so they form a basis for $\mathbf{R}^2$).

Example (Questions about a [matrix)**.** transformation] Let

$$A=\begin{pmatrix}1 & 1\\ 0 & 1\\ 1 & 1\end{pmatrix},$$

and let $T(x)=Ax$, so $T:\mathbf{R}^2\to\mathbf{R}^3$ is a matrix transformation.

1. Evaluate $T(u)$ for $u=\begin{pmatrix}3\\4\end{pmatrix}$.

2. Let

   $$b=\begin{pmatrix}7\\5\\7\end{pmatrix}.$$

   Find a vector v in $\mathbf{R}^2$ such that $T(v)=b$. Is there more than one?

3. Does there exist a vector w in $\mathbf{R}^3$ such that there is more than one v in $\mathbf{R}^2$ with $T(v)=w$?

4. Find a vector w in $\mathbf{R}^3$ which is not in the range of T.

Note: all of the above questions are intrinsic to the transformation T: they make sense to ask whether or not T is a matrix transformation. See the next example. As T is in fact a matrix transformation, all of these questions will translate into questions about the corresponding matrix A.

Solution.

1. We evaluate $T(u)$ by substituting the definition of T in terms of matrix multiplication:

$$T\begin{pmatrix}3\\4\end{pmatrix} = \begin{pmatrix}1 & 1\\0 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}3\\4\end{pmatrix} = \begin{pmatrix}7\\4\\7\end{pmatrix}.$$

2. We want to find a vector v such that $b = T(v) = Av$. In other words, we want to solve the matrix equation $Av = b$. We form an augmented matrix and row reduce:

$$\begin{pmatrix}1 & 1\\0 & 1\\1 & 1\end{pmatrix} v = \begin{pmatrix}7\\5\\7\end{pmatrix} \xrightarrow[\text{matrix}]{\text{augmented}} \left(\begin{array}{cc|c}1 & 1 & 7\\0 & 1 & 5\\1 & 1 & 7\end{array}\right) \xrightarrow[\text{reduce}]{\text{row}} \left(\begin{array}{cc|c}1 & 0 & 2\\0 & 1 & 5\\0 & 0 & 0\end{array}\right).$$

   This gives $x = 2$ and $y = 5$, so that there is a unique vector

$$v = \begin{pmatrix}2\\5\end{pmatrix}$$

   such that $T(v) = b$.

3. *Translation:* is there any vector w in $\mathbf{R}^3$ such that the solution set of $Av = w$ has more than one vector in it? The solution set of $Ax = w$, if non-empty, is a translate of the solution set of $Av = b$ above, which has one vector in it. See this key observation in Section 2.4. It follows that the solution set of $Av = w$ can have at most one vector.

4. *Translation:* find a vector w such that the matrix equation $Av = w$ is not consistent. Notice that if we take

$$w = \begin{pmatrix}1\\2\\3\end{pmatrix},$$

   then the matrix equation $Av = w$ translates into the system of equations

$$\begin{cases} x + y = 1\\ \phantom{x + {}} y = 2\\ x + y = 3, \end{cases}$$

   which is clearly inconsistent.

Example (Questions about a [non-matrix). transformation] Define a transformation $T: \mathbf{R}^2 \to \mathbf{R}^3$ by the formula

$$T\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\ln(x)\\\cos(y)\\\ln(x)\end{pmatrix}.$$

1. Evaluate $T(u)$ for $u = \begin{pmatrix} 1 \\ \pi \end{pmatrix}$.

2. Let
$$b = \begin{pmatrix} 7 \\ 1 \\ 7 \end{pmatrix}.$$
Find a vector v in $\mathbf{R}^2$ such that $T(v) = b$. Is there more than one?

3. Does there exist a vector w in $\mathbf{R}^3$ such that there is more than one v in $\mathbf{R}^2$ with $T(v) = w$?

4. Find a vector w in $\mathbf{R}^3$ which is not in the range of T.

Note: we asked (almost) the exact same questions about a matrix transformation in the previous example. The point of this example is to illustrate the fact that the questions make sense for a transformation that has no hope of coming from a matrix. In this case, these questions do not translate into questions about a matrix; they have to be answered in some other way.

Solution.

1. We evaluate $T(u)$ using the defining formula:
$$T\begin{pmatrix} 1 \\ \pi \end{pmatrix} = \begin{pmatrix} \ln(1) \\ \cos(\pi) \\ \ln(1) \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}.$$

2. We have
$$T\begin{pmatrix} e^7 \\ 2\pi n \\ e^7 \end{pmatrix} = \begin{pmatrix} \ln(e^7) \\ \cos(2\pi n) \\ \ln(e^7) \end{pmatrix} = \begin{pmatrix} 7 \\ 1 \\ 7 \end{pmatrix}$$
for any whole number n. Hence there are infinitely many such vectors.

3. The vector b from the previous part is an example of such a vector.

4. Since $\cos(y)$ is always between -1 and 1, the vector
$$w = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$$
is not in the range of T.

3.2 One-to-one and Onto Transformations

Objectives

1. Understand the definitions of one-to-one and onto transformations.
2. *Recipes:* verify whether a matrix transformation is one-to-one and/or onto.
3. *Pictures:* examples of matrix transformations that are/are not one-to-one and/or onto.
4. *Vocabulary words:* **one-to-one**, **onto**.

In this section, we discuss two of the most basic questions one can ask about a transformation: whether it is *one-to-one* and/or *onto*. For a matrix transformation, we translate these questions into the language of matrices.

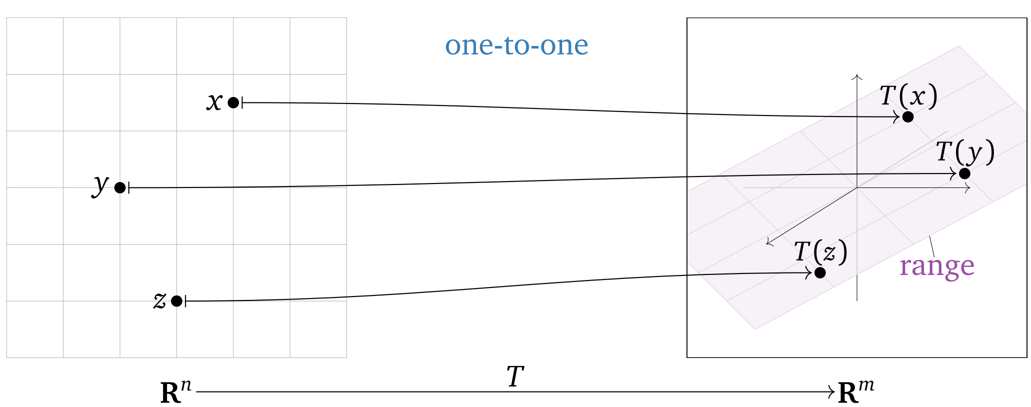
3.2.1 One-to-one Transformations

Definition (One-to-one transformations)**.** A transformation $T: \mathbf{R}^n \to \mathbf{R}^m$ is **one-to-one** if, for every vector b in $\mathbf{R}^m$, the equation $T(x) = b$ has *at most one* solution x in $\mathbf{R}^n$.

Remark. Another word for *one-to-one* is *injective*.

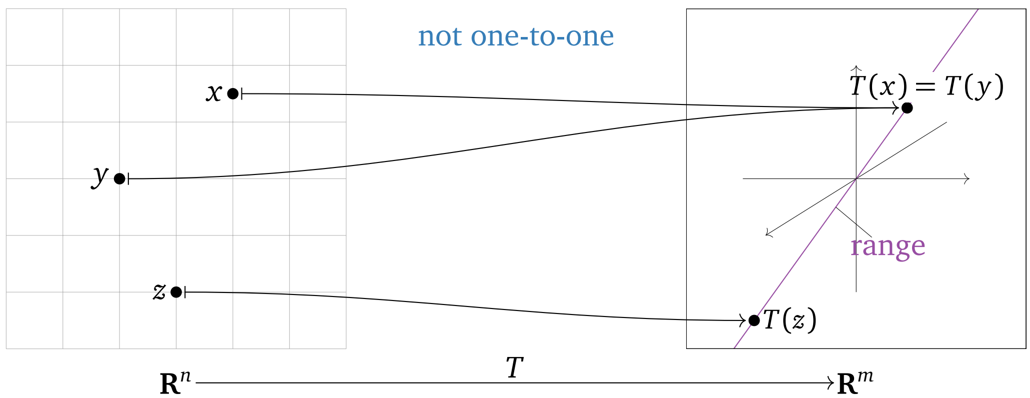
Here are some equivalent ways of saying that T is one-to-one:

- For every vector b in $\mathbf{R}^m$, the equation $T(x) = b$ has *zero or one* solution x in $\mathbf{R}^n$.
- Different inputs of T have different outputs.
- If $T(u) = T(v)$ then $u = v$.



Here are some equivalent ways of saying that T is *not* one-to-one:

- There exists some vector b in $\mathbf{R}^m$ such that the equation $T(x) = b$ has *more than one* solution x in $\mathbf{R}^n$.
- There are two different inputs of T with the same output.
- There exist vectors u, v such that $u \neq v$ but $T(u) = T(v)$.



Example (Functions of one variable)**.** The function $\sin\colon \mathbf{R} \to \mathbf{R}$ is not one-to-one. Indeed, $\sin(0) = \sin(\pi) = 0$, so the inputs 0 and π have the same output 0. In fact, the equation $\sin(x) = 0$ has infinitely many solutions $\ldots, -2\pi, -\pi, 0, \pi, 2\pi, \ldots$.

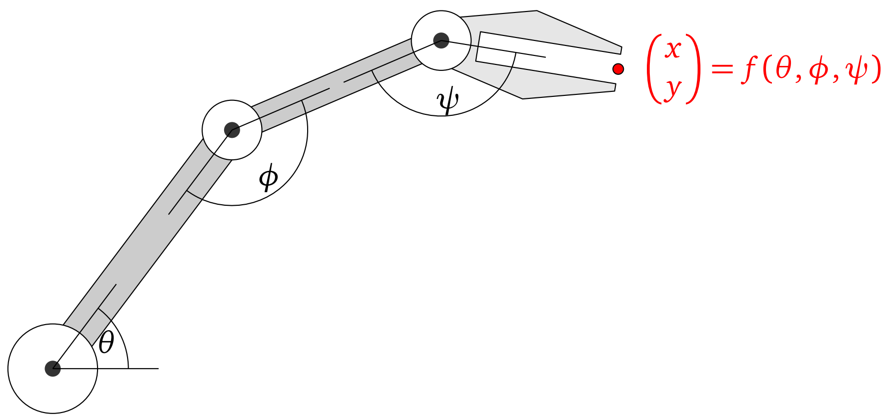
The function $\exp\colon \mathbf{R} \to \mathbf{R}$ defined by $\exp(x) = e^x$ is one-to-one. Indeed, if $T(x) = T(y)$, then $e^x = e^y$, so $\ln(e^x) = \ln(e^y)$, and hence $x = y$. The equation $T(x) = C$ has one solution $x = \ln(C)$ if $C > 0$, and it has zero solutions if $C \leq 0$.

The function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^3$ is one-to-one. Indeed, if $f(x) = f(y)$ then $x^3 = y^3$; taking cube roots gives $x = y$. In other words, the only solution of $f(x) = C$ is $x = \sqrt[3]{C}$.

The function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^3 - x$ is not one-to-one. Indeed, $f(0) = f(1) = f(-1) = 0$, so the inputs $0, 1, -1$ all have the same output 0. The solutions of the equation $x^3 - x = 0$ are exactly the roots of $f(x) = x(x-1)(x+1)$, and this equation has three roots.

The function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^2$ is not one-to-one. Indeed, $f(1) = 1 = f(-1)$, so the inputs 1 and -1 have the same outputs. The function $g\colon \mathbf{R} \to \mathbf{R}$ defined by $g(x) = |x|$ is not one-to-one for the same reason.

Example (A real-word transformation: robotics)**.** Suppose you are building a robot arm with three joints that can move its hand around a plane, as in this example in Section 3.1.



Define a transformation $f : \mathbf{R}^3 \to \mathbf{R}^2$ as follows: $f(\theta, \phi, \psi)$ is the (x, y) position of the hand when the joints are rotated by angles θ, ϕ, ψ, respectively. Asking whether f is one-to-one is the same as asking whether there is more than one way to move the arm in order to reach your coffee cup. (There is.)

Theorem (One-to-one matrix transformations)**.** *Let A be an $m \times n$ matrix, and let $T(x) = Ax$ be the associated matrix transformation. The following statements are equivalent:*

1. *T is one-to-one.*
2. *For every b in $\mathbf{R}^m$, the equation $T(x) = b$ has at most one solution.*
3. *For every b in $\mathbf{R}^m$, the equation $Ax = b$ has a unique solution or is inconsistent.*
4. *$Ax = 0$ has* only *the trivial solution.*
5. *The columns of A are linearly independent.*
6. *A has a pivot in every column.*
7. *The range of T has dimension n.*

Proof. Statements 1, 2, and 3 are translations of each other. The equivalence of 3 and 4 follows from this key observation in Section 2.4: if $Ax = 0$ has only one solution, then $Ax = b$ has only one solution as well, or it is inconsistent. The equivalence of 4, 5, and 6 is a consequence of this important note in Section 2.5, and the equivalence of 6 and 7 follows from the fact that the rank of a matrix is equal to the number of columns with pivots. □

Recall that *equivalent* means that, for a given matrix, either all of the statements are true simultaneously, or they are all false.

Example (A matrix transformation that is one-to-one)**.** Let A be the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and define $T : \mathbf{R}^2 \to \mathbf{R}^3$ by $T(x) = Ax$. Is T one-to-one?

Solution. The reduced row echelon form of A is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Hence A has a pivot in every column, so T is one-to-one.

Use this link to view the online demo

A picture of the matrix transformation T. As you drag the input vector on the left side, you see that different input vectors yield different output vectors on the right side.

Example (A matrix transformation that is not one-to-one)**.** Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

and define $T : \mathbf{R}^3 \to \mathbf{R}^3$ by $T(x) = Ax$. Is T one-to-one? If not, find two different vectors u, v such that $T(u) = T(v)$.

Solution. The matrix A is already in reduced row echelon form. It does not have a pivot in every column, so T is not one-to-one. Therefore, we know from the theorem that $Ax = 0$ has nontrivial solutions. If v is a nontrivial (i.e., nonzero) solution of $Av = 0$, then $T(v) = Av = 0 = A0 = T(0)$, so 0 and v are different vectors with the same output. For instance,

$$T\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 = T\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Geometrically, T is projection onto the xy-plane. Any two vectors that lie on the same vertical line will have the same projection. For b on the xy-plane, the solution set of $T(x) = b$ is the entire vertical line containing b. In particular, $T(x) = b$ has *infinitely many* solutions.

Use this link to view the online demo

A picture of the matrix transformation T. The transformation T projects a vector onto the xy-plane. The violet line is the solution set of $T(x) = 0$. If you drag x along the violet line, the output $T(x) = Ax$ does not change. This demonstrates that $T(x) = 0$ has more than one solution, so T is not one-to-one.

Example (A matrix transformation that is not one-to-one)**.** Let A be the matrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

and define $T: \mathbf{R}^3 \to \mathbf{R}^2$ by $T(x) = Ax$. Is T one-to-one? If not, find two different vectors u, v such that $T(u) = T(v)$.

Solution. The reduced row echelon form of A is

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

There is not a pivot in every column, so T is not one-to-one. Therefore, we know from the theorem that $Ax = 0$ has nontrivial solutions. If v is a nontrivial (i.e., nonzero) solution of $Av = 0$, then $T(v) = Av = 0 = A0 = T(0)$, so 0 and v are different vectors with the same output. In order to find a nontrivial solution, we find the parametric form of the solutions of $Ax = 0$ using the reduced matrix above:

$$\begin{cases} x \phantom{{}+y} - z = 0 \\ \phantom{x+{}} y + z = 0 \end{cases} \implies \begin{cases} x = z \\ y = -z \end{cases}$$

The free variable is z. Taking $z = 1$ gives the nontrivial solution

$$T\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = 0 = T\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Use this link to view the online demo

A picture of the matrix transformation T. The violet line is the null space of A, i.e., solution set of $T(x) = 0$. If you drag x along the violet line, the output $T(x) = Ax$ does not change. This demonstrates that $T(x) = 0$ has more than one solution, so T is not one-to-one.

Example (A matrix transformation that is not one-to-one)**.** Let

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

and define $T: \mathbf{R}^3 \to \mathbf{R}^2$ by $T(x) = Ax$. Is T one-to-one? If not, find two different vectors u, v such that $T(u) = T(v)$.

Solution. The reduced row echelon form of A is

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

There is not a pivot in every column, so T is not one-to-one. Therefore, we know from the theorem that $Ax = 0$ has nontrivial solutions. If v is a nontrivial (i.e., nonzero) solution of $Av = 0$, then $T(v) = Av = 0 = A0 = T(0)$, so 0 and v are different vectors with the same output. In order to find a nontrivial solution, we find the parametric form of the solutions of $Ax = 0$ using the reduced matrix above:

$$x - y + 2z = 0 \quad \Longrightarrow \quad x = y - 2z.$$

The free variables are y and z. Taking $y = 1$ and $z = 0$ gives the nontrivial solution

$$T\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}1 & -1 & 2\\ -2 & 2 & -4\end{pmatrix}\begin{pmatrix}1\\1\\0\end{pmatrix} = 0 = T\begin{pmatrix}0\\0\\0\end{pmatrix}.$$

Use this link to view the online demo

A picture of the matrix transformation T. The violet plane is the solution set of $T(x) = 0$. If you drag x along the violet plane, the output $T(x) = Ax$ does not change. This demonstrates that $T(x) = 0$ has more than one solution, so T is not one-to-one.

The previous three examples can be summarized as follows. Suppose that $T(x) = Ax$ is a matrix transformation that is *not* one-to-one. By the theorem, there is a nontrivial solution of $Ax = 0$. This means that the null space of A is not the zero space. All of the vectors in the null space are solutions to $T(x) = 0$. If you compute a nonzero vector v in the null space (by row reducing and finding the parametric form of the solution set of $Ax = 0$, for instance), then v and 0 both have the same output: $T(v) = Av = 0 = T(0)$.

Wide matrices do not have one-to-one transformations. If $T: \mathbf{R}^n \to \mathbf{R}^m$ is a one-to-one matrix transformation, what can we say about the relative sizes of n and m?

The matrix associated to T has n columns and m rows. Each row and each column can only contain one pivot, so in order for A to have a pivot in every column, it must have *at least as many rows* as columns: $n \le m$.

This says that, for instance, $\mathbf{R}^3$ is "too big" to admit a one-to-one linear transformation into $\mathbf{R}^2$.

Note that there exist tall matrices that are not one-to-one: for example,

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

does not have a pivot in every column.

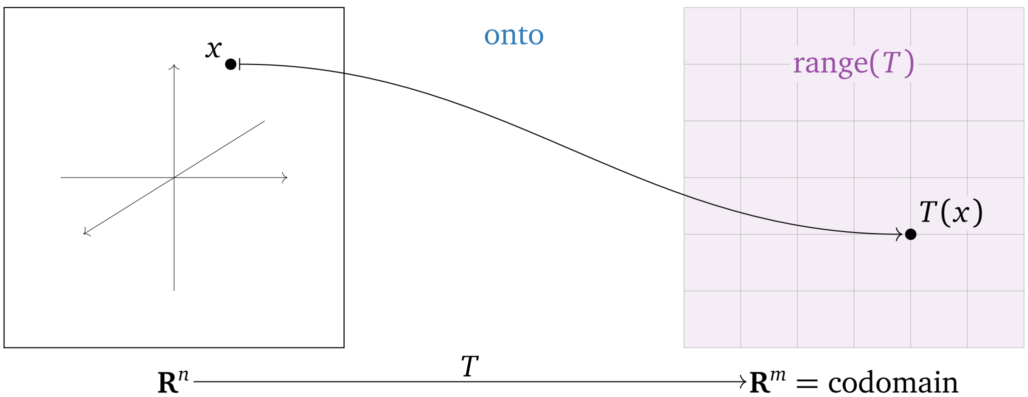
3.2.2 Onto Transformations

Definition (Onto transformations)**.** A transformation $T : \mathbf{R}^n \to \mathbf{R}^m$ is **onto** if, for every vector b in $\mathbf{R}^m$, the equation $T(x) = b$ has *at least one* solution x in $\mathbf{R}^n$.

Remark. Another word for *onto* is *surjective*.

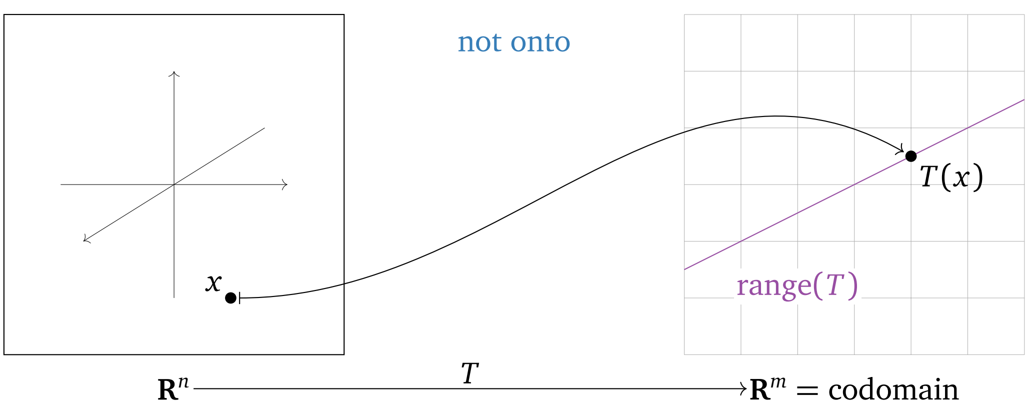
Here are some equivalent ways of saying that T is onto:

- The range of T is equal to the codomain of T.
- Every vector in the codomain is the output of some input vector.



Here are some equivalent ways of saying that T is *not* onto:

- The range of T is smaller than the codomain of T.
- There exists a vector b in $\mathbf{R}^m$ such that the equation $T(x) = b$ does not have a solution.
- There is a vector in the codomain that is not the output of any input vector.



Example (Functions of one variable)**.** The function $\sin: \mathbf{R} \to \mathbf{R}$ is not onto. Indeed, taking $b = 2$, the equation $\sin(x) = 2$ has no solution. The range of sin is the closed interval $[-1, 1]$, which is smaller than the codomain $\mathbf{R}$.

The function $\exp: \mathbf{R} \to \mathbf{R}$ defined by $\exp(x) = e^x$ is not onto. Indeed, taking $b = -1$, the equation $\exp(x) = e^x = -1$ has no solution. The range of exp is the set $(0, \infty)$ of all *positive* real numbers.

The function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^3$ is onto. Indeed, the equation $f(x) = x^3 = b$ always has the solution $x = \sqrt[3]{b}$.

The function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^3 - x$ is onto. Indeed, the solutions of the equation $f(x) = x^3 - x = b$ are the roots of the polynomial $x^3 - x - b$; as this is a cubic polynomial, it has at least one real root.

Example (A real-word transformation: robotics)**.** The robot arm transformation of this example is not onto. The robot cannot reach objects that are very far away.

Theorem (Onto matrix transformations)**.** *Let A be an $m \times n$ matrix, and let $T(x) = Ax$ be the associated matrix transformation. The following statements are equivalent:*

1. *T is onto.*
2. *$T(x) = b$ has at least one solution for every b in $\mathbf{R}^m$.*
3. *$Ax = b$ is consistent for every b in $\mathbf{R}^m$.*
4. *The columns of A span $\mathbf{R}^m$.*
5. *A has a pivot in every row.*
6. *The range of T has dimension m.*

Proof. Statements 1, 2, and 3 are translations of each other. The equivalence of 3, 4, 5, and 6 follows from this theorem in Section 2.3. □

Example (A matrix transformation that is onto)**.** Let A be the matrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

and define $T: \mathbf{R}^3 \to \mathbf{R}^2$ by $T(x) = Ax$. Is T onto?

Solution. The reduced row echelon form of A is

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Hence A has a pivot in every row, so T is onto.

Use this link to view the online demo

A picture of the matrix transformation T. Every vector on the right side is the output of T for a suitable input. If you drag b, the demo will find an input vector x with output b.

Example (A matrix transformation that is not onto)**.** Let A be the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and define $T : \mathbf{R}^2 \to \mathbf{R}^3$ by $T(x) = Ax$. Is T onto? If not, find a vector b in $\mathbf{R}^3$ such that $T(x) = b$ has no solution.

Solution. The reduced row echelon form of A is

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Hence A does not have a pivot in every row, so T is not onto. In fact, since

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \\ x \end{pmatrix},$$

we see that for every output vector of T, the third entry is equal to the first. Therefore,

$$b = (1, 2, 3)$$

is not in the range of T.

Use this link to view the online demo

A picture of the matrix transformation T. The range of T is the violet plane on the right; this is smaller than the codomain $\mathbf{R}^3$. If you drag b off of the violet plane, then the equation $Ax = b$ becomes inconsistent; this means $T(x) = b$ has no solution.

Example (A matrix transformation that is not onto)**.** Let

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

and define $T : \mathbf{R}^3 \to \mathbf{R}^2$ by $T(x) = Ax$. Is T onto? If not, find a vector b in $\mathbf{R}^2$ such that $T(x) = b$ has no solution.

Solution. The reduced row echelon form of A is

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

There is not a pivot in every row, so T is not onto. The range of T is the column space of A, which is equal to

$$\text{Span}\left\{\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -4 \end{pmatrix}\right\} = \text{Span}\left\{\begin{pmatrix} 1 \\ -2 \end{pmatrix}\right\},$$

since all three columns of A are collinear. Therefore, any vector not on the line through $\binom{1}{-2}$ is not in the range of T. For instance, if $b = \binom{1}{1}$ then $T(x) = b$ has no solution.

Use this link to view the online demo

A picture of the matrix transformation T. The range of T is the violet line on the right; this is smaller than the codomain $\mathbf{R}^2$. If you drag b off of the violet line, then the equation $Ax = b$ becomes inconsistent; this means $T(x) = b$ has no solution.

The previous two examples illustrate the following observation. Suppose that $T(x) = Ax$ is a matrix transformation that is *not* onto. This means that range$(T) =$ Col(A) is a subspace of $\mathbf{R}^m$ of dimension less than m: perhaps it is a line in the plane, or a line in 3-space, or a plane in 3-space, etc. Whatever the case, the range of T is *very small* compared to the codomain. To find a vector not in the range of T, choose a random nonzero vector b in $\mathbf{R}^m$; you have to be extremely unlucky to choose a vector that is in the range of T. Of course, to check whether a given vector b is in the range of T, you have to solve the matrix equation $Ax = b$ to see whether it is consistent.

Tall matrices do not have onto transformations. If $T : \mathbf{R}^n \to \mathbf{R}^m$ is an onto matrix transformation, what can we say about the relative sizes of n and m?

The matrix associated to T has n columns and m rows. Each row and each column can only contain one pivot, so in order for A to have a pivot in every row, it must have *at least as many columns* as rows: $m \leq n$.

This says that, for instance, $\mathbf{R}^2$ is "too small" to admit an onto linear transformation to $\mathbf{R}^3$.

Note that there exist wide matrices that are not onto: for example,

$$\begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{pmatrix}$$

does not have a pivot in every row.

3.2.3 Comparison

The above expositions of one-to-one and onto transformations were written to mirror each other. However, "one-to-one" and "onto" are complementary notions: neither one implies the other. Below we have provided a chart for comparing the two. In the chart, A is an $m\times n$ matrix, and $T\colon \mathbf{R}^n \to \mathbf{R}^m$ is the matrix transformation $T(x) = Ax$.

T **is one-to-one**	T **is onto**
$T(x) = b$ has *at most one solution* for every b.	$T(x) = b$ has *at least one solution* for every b.
The columns of A are linearly independent.	The columns of A span $\mathbf{R}^m$.
A has a pivot in every column.	A has a pivot in every row.
The range of T has dimension n.	The range of T has dimension m.

Example (Functions of one variable)**.** The function $\sin\colon \mathbf{R} \to \mathbf{R}$ is neither one-to-one nor onto.

The function $\exp\colon \mathbf{R} \to \mathbf{R}$ defined by $\exp(x) = e^x$ is one-to-one but not onto.
The function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^3$ is one-to-one and onto.
The function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^3 - x$ is onto but not one-to-one.

Example (A matrix transformation that is neither one-to-one nor onto)**.** Let

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

and define $T\colon \mathbf{R}^3 \to \mathbf{R}^2$ by $T(x) = Ax$. This transformation is neither one-to-one nor onto, as we saw in this example and this example.

Use this link to view the online demo

A picture of the matrix transformation T. The violet plane is the solution set of $T(x) = 0$. If you drag x along the violet plane, the output $T(x) = Ax$ does not change. This demonstrates that $T(x) = 0$ has more than one solution, so T is not one-to-one. The range of T is the violet line on the right; this is smaller than the codomain $\mathbf{R}^2$. If you drag b off of the violet line, then the equation $Ax = b$ becomes inconsistent; this means $T(x) = b$ has no solution.

Example (A matrix transformation that is one-to-one but not onto)**.** Let A be the matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and define $T: \mathbf{R}^2 \to \mathbf{R}^3$ by $T(x) = Ax$. This transformation is one-to-one but not onto, as we saw in this example and this example.

Use this link to view the online demo

A picture of the matrix transformation T. The range of T is the violet plane on the right; this is smaller than the codomain $\mathbf{R}^3$. If you drag b off of the violet plane, then the equation $Ax = b$ becomes inconsistent; this means $T(x) = b$ has no solution. However, for b lying on the violet plane, there is a unique vector x such that $T(x) = b$.

Example (A matrix transformation that is onto but not one-to-one)**.** Let A be the matrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

and define $T: \mathbf{R}^3 \to \mathbf{R}^2$ by $T(x) = Ax$. This transformation is onto but not one-to-one, as we saw in this example and this example.

Use this link to view the online demo

A picture of the matrix transformation T. Every vector on the right side is the output of T for a suitable input. If you drag b, the demo will find an input vector x with output b. The violet line is the null space of A, i.e., solution set of $T(x) = 0$. If you drag x along the violet line, the output $T(x) = Ax$ does not change. This demonstrates that $T(x) = 0$ has more than one solution, so T is not one-to-one.

Example (Matrix transformations that are both one-to-one and onto)**.** In this subsection in Section 3.1, we discussed the transformations defined by several 2×2 matrices, namely:

$$\begin{aligned}
\text{Reflection:} &\qquad A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\
\text{Dilation:} &\qquad A = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix} \\
\text{Identity:} &\qquad A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\text{Rotation:} &\qquad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\
\text{Shear:} &\qquad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.
\end{aligned}$$

In each case, the associated matrix transformation $T(x) = Ax$ is both one-to-one and onto. A 2×2 matrix A has a pivot in every row if and only if it has a pivot in every column (if and only if it has two pivots), so in this case, the transformation T is one-to-one if and only if it is onto. One can see geometrically that they are onto (what is the input for a given output?), or that they are one-to-one using the fact that the columns of A are not collinear.

Use this link to view the online demo

Counterclockwise rotation by 90° is a matrix transformation. This transformation is onto (if b is a vector in $\mathbf{R}^2$, then it is the output vector for the input vector which is b rotated clockwise *by 90°), and it is one-to-one (different vectors rotate to different vectors).*

One-to-one is the same as onto for square matrices. We observed in the previous example that a square matrix has a pivot in every row if and only if it has a pivot in every column. Therefore, a matrix transformation T from $\mathbf{R}^n$ to itself is one-to-one if and only if it is onto: in this case, the two notions are equivalent.

Conversely, by this note and this note, if a matrix transformation $T: \mathbf{R}^m \to \mathbf{R}^n$ is both one-to-one and onto, then $m = n$.

Note that in general, a transformation T is both one-to-one and onto if and only if $T(x) = b$ has *exactly one* solution for all b in $\mathbf{R}^m$.

3.3 Linear Transformations

Objectives

1. Learn how to verify that a transformation is linear, or prove that a transformation is not linear.
2. Understand the relationship between linear transformations and matrix transformations.
3. *Recipe:* compute the matrix of a linear transformation.
4. *Theorem:* linear transformations and matrix transformations.
5. *Notation:* the **standard coordinate vectors** $e_1, e_2, \ldots$.
6. *Vocabulary words:* **linear transformation**, **standard matrix**, **identity matrix**.

In Section 3.1, we studied the geometry of matrices by regarding them as functions, i.e., by considering the associated *matrix transformations*. We defined some vocabulary (domain, codomain, range), and asked a number of natural questions about a transformation. For a matrix transformation, these translate into questions about matrices, which we have many tools to answer.

In this section, we make a change in perspective. Suppose that we are given a *transformation* that we would like to study. If we can prove that our transformation is a matrix transformation, then we can use linear algebra to study it. This raises two important questions:

1. How can we tell if a transformation is a matrix transformation?
2. If our transformation is a matrix transformation, how do we find its matrix?

For example, we saw in this example in Section 3.1 that the matrix transformation

$$T\colon \mathbf{R}^2 \longrightarrow \mathbf{R}^2 \qquad T(x) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x$$

is a counterclockwise rotation of the plane by 90°. However, we could have *defined* T in this way:

$$T\colon \mathbf{R}^2 \longrightarrow \mathbf{R}^2 \qquad T(x) = \text{the counterclockwise rotation of } x \text{ by } 90^\circ.$$

Given this definition, it is not at all obvious that T is a matrix transformation, or what matrix it is associated to.

3.3.1 Linear Transformations: Definition

In this section, we introduce the class of transformations that come from matrices.

Definition. A **linear transformation** is a transformation $T\colon \mathbf{R}^n \to \mathbf{R}^m$ satisfying

$$\begin{aligned} T(u+v) &= T(u) + T(v) \\ T(cu) &= cT(u) \end{aligned}$$

for all vectors u, v in $\mathbf{R}^n$ and all scalars c.

Let $T\colon \mathbf{R}^n \to \mathbf{R}^m$ be a matrix transformation: $T(x) = Ax$ for an $m \times n$ matrix A. By this proposition in Section 2.3, we have

$$\begin{aligned} T(u+v) &= A(u+v) = Au + Av = T(u) + T(v) \\ T(cu) &= A(cu) = cAu = cT(u) \end{aligned}$$

for all vectors u, v in $\mathbf{R}^n$ and all scalars c. Since a matrix transformation satisfies the two defining properties, it is a linear transformation

We will see in the next subsection that the opposite is true: every linear transformation is a matrix transformation; we just haven't computed its matrix yet.

Facts about linear transformations. *Let $T\colon \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Then:*

1. $T(0) = 0$.

2. *For any vectors $v_1, v_2, \ldots, v_k$ in $\mathbf{R}^n$ and scalars $c_1, c_2, \ldots, c_k$, we have*

$$T\bigl(c_1v_1 + c_2v_2 + \cdots + c_kv_k\bigr) = c_1T(v_1) + c_2T(v_2) + \cdots + c_kT(v_k).$$

Proof.

1. Since $0 = -0$, we have

$$T(0) = T(-0) = -T(0)$$

   by the second defining property. The only vector w such that $w = -w$ is the zero vector.

2. Let us suppose for simplicity that $k = 2$. Then

$$\begin{aligned} T(c_1v_1 + c_2v_2) &= T(c_1v_1) + T(c_2v_2) && \text{first property} \\ &= c_1T(v_1) + c_2T(v_2) && \text{second property.} \end{aligned}$$

□

In engineering, the second fact is called the *superposition principle*; it should remind you of the distributive property. For example, $T(cu+dv) = cT(u)+dT(v)$ for any vectors u, v and any scalars c, d. To restate the first fact:

A linear transformation necessarily takes the zero vector to the zero vector.

Example (A non-linear transformation)**.** Define $T : \mathbf{R} \to \mathbf{R}$ by $T(x) = x + 1$. Is T a linear transformation?

Solution. We have $T(0) = 0+1 = 1$. Since any linear transformation necessarily takes zero to zero by the above important note, we conclude that T is *not* linear (even though its graph is a line).

Note: in this case, it was not necessary to check explicitly that T does not satisfy both defining properties: since $T(0) = 0$ is a consequence of these properties, at least one of them must not be satisfied. (In fact, this T satisfies neither.)

Example (Verifying linearity: dilation)**.** Define $T : \mathbf{R}^2 \to \mathbf{R}^2$ by $T(x) = 1.5x$. Verify that T is linear.

Solution. We have to check the defining properties for *all* vectors u, v and *all* scalars c. In other words, we have to treat u, v, and c as *unknowns*. The only thing we are allowed to use is the definition of T.

$$\begin{aligned} T(u+v) &= 1.5(u+v) = 1.5u + 1.5v = T(u) + T(v) \\ T(cu) &= 1.5(cu) = c(1.5u) = cT(u). \end{aligned}$$

Since T satisfies both defining properties, T is linear.

Note: we know from this example in Section 3.1 that T is a matrix transformation: in fact,

$$T(x) = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix} x.$$

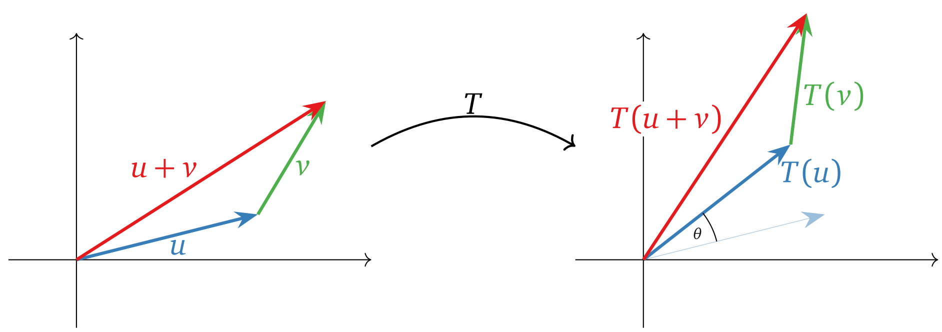
Since a matrix transformation is a linear transformation, this is another proof that T is linear.

Example (Verifying linearity: rotation)**.** Define $T : \mathbf{R}^2 \to \mathbf{R}^2$ by

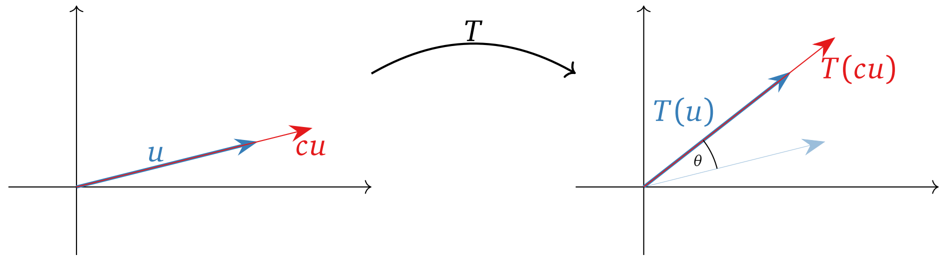
$$T(x) = \text{ the vector } x \text{ rotated counterclockwise by the angle } \theta.$$

Verify that T is linear.

Solution. Since T is defined geometrically, we give a geometric argument. For the first property, $T(u) + T(v)$ is the sum of the vectors obtained by rotating u and v by θ. On the other side of the equation, $T(u + v)$ is the vector obtained by rotating the sum of the vectors u and v. But it does not matter whether we sum or rotate first, as the following picture shows.



For the second property, $cT(u)$ is the vector obtained by rotating u by the angle θ, then changing its length by a factor of c (reversing direction of $c < 0$. On the other hand, $T(cu)$ first changes the length of c, then rotates. But it does not matter in which order we do these two operations.



This verifies that T is a linear transformation. We will find its matrix in the next subsection. Note however that it is not at all obvious that T can be expressed as multiplication by a matrix.

Example (A transformation defined by a formula)**.** Define $T\colon \mathbf{R}^2 \to \mathbf{R}^3$ by the formula

$$T\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}3x-y\\y\\x\end{pmatrix}.$$

Verify that T is linear.

Solution. We have to check the defining properties for *all* vectors u, v and *all* scalars c. In other words, we have to treat u, v, and c as *unknowns*; the only thing we are allowed to use is the definition of T. Since T is defined in terms of the coordinates of u, v, we need to give those names as well; say $u = \binom{x_1}{y_1}$ and $v = \binom{x_2}{y_2}$. For the first property, we have

$$\begin{aligned} T\left(\begin{pmatrix}x_1\\y_1\end{pmatrix} + \begin{pmatrix}x_2\\y_2\end{pmatrix}\right) = T\begin{pmatrix}x_1+x_2\\y_1+y_2\end{pmatrix} &= \begin{pmatrix}3(x_1+x_2)-(y_1+y_2)\\y_1+y_2\\x_1+x_2\end{pmatrix} \\ &= \begin{pmatrix}(3x_1-y_1)+(3x_2-y_2)\\y_1+y_2\\x_1+x_2\end{pmatrix} \\ &= \begin{pmatrix}3x_1-y_1\\y_1\\x_1\end{pmatrix} + \begin{pmatrix}3x_2-y_2\\y_2\\x_2\end{pmatrix} = T\begin{pmatrix}x_1\\y_1\end{pmatrix} + T\begin{pmatrix}x_2\\y_2\end{pmatrix}. \end{aligned}$$

For the second property,

$$\begin{aligned} T\left(c\begin{pmatrix}x_1\\y_1\end{pmatrix}\right) = T\begin{pmatrix}cx_1\\cy_1\end{pmatrix} &= \begin{pmatrix}3(cx_1)-(cy_1)\\cy_1\\cx_1\end{pmatrix} \\ &= \begin{pmatrix}c(3x_1-y_1)\\cy_1\\cx_1\end{pmatrix} = c\begin{pmatrix}3x_1-y_1\\y_1\\x_1\end{pmatrix} = cT\begin{pmatrix}x_1\\y_1\end{pmatrix}. \end{aligned}$$

Since T satisfies the defining properties, T is a linear transformation.

Note: we will see in this example below that

$$T\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}3 & -1\\0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}.$$

Hence T is in fact a matrix transformation.

One can show that, if a transformation is defined by formulas in the coordinates as in the above example, then the transformation is linear if and only if each coordinate is a linear expression in the variables with no constant term.

Example (A translation)**.** Define $T : \mathbf{R}^3 \to \mathbf{R}^3$ by

$$T(x) = x + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

This kind of transformation is called a **translation**. As in a previous example, this T is not linear, because $T(0)$ is not the zero vector.

Example (More non-linear transformations)**.** Verify that the following transformations from $\mathbf{R}^2$ to $\mathbf{R}^2$ are not linear:

$$T_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} |x| \\ y \end{pmatrix} \qquad T_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} xy \\ y \end{pmatrix} \qquad T_3\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+1 \\ x-2y \end{pmatrix}.$$

Solution. In order to verify that a transformation T is *not* linear, we have to show that T does not satisfy *at least one* of the two defining properties. For the first, the negation of the statement "$T(u+v) = T(u)+T(v)$ for all vectors u, v" is "there exists at least one pair of vectors u, v such that $T(u+v) \neq T(u)+T(v)$." In other words, it suffices to find *one example* of a pair of vectors u, v such that $T(u+v) \neq T(u)+T(v)$. Likewise, for the second, the negation of the statement "$T(cu) = cT(u)$ for all vectors u and all scalars c" is "there exists some vector u and some scalar c such that $T(cu) \neq cT(u)$." In other words, it suffices to find *one* vector u and *one* scalar c such that $T(cu) \neq cT(u)$.

For the first transformation, we note that

$$T_1\left(-\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = T_1\begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} |-1| \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

but that

$$-T_1\begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\begin{pmatrix} |1| \\ 0 \end{pmatrix} = -\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

Therefore, this transformation does not satisfy the second property.

For the second transformation, we note that

$$T_2\left(2\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = T_2\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2\cdot 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

but that

$$2T_2\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1\cdot 1 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Therefore, this transformation does not satisfy the second property.

For the third transformation, we observe that

$$T_3\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2(0)+1 \\ 0-2(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Since T_3 does not take the zero vector to the zero vector, it cannot be linear.

When deciding whether a transformation T is linear, generally the first thing to do is to check whether $T(0) = 0$; if not, T is automatically not linear. Note however that the non-linear transformations T_1 and T_2 of the above example do take the zero vector to the zero vector.

Challenge. Find an example of a transformation that satisfies the first property of linearity but not the second.

3.3.2 The Standard Coordinate Vectors

In the next subsection, we will present the relationship between linear transformations and matrix transformations. Before doing so, we need the following important notation.

Standard coordinate vectors. The **standard coordinate vectors in $\mathbf{R}^n$** are the n vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \ldots, \quad e_{n-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

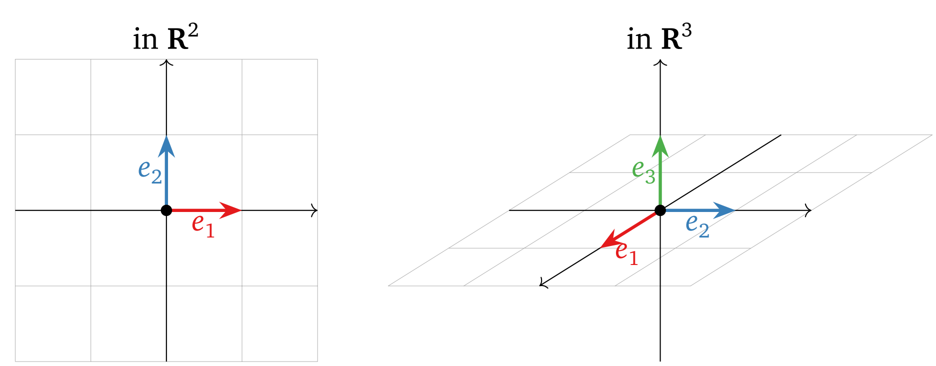
The ith entry of e_i is equal to 1, and the other entries are zero.
From now on, for the rest of the book, we will use the symbols $e_1, e_2, \ldots$ to denote the standard coordinate vectors.

There is an ambiguity in this notation: one has to know from context that e_1 is meant to have n entries. That is, the vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

may both be denoted e_1, depending on whether we are discussing vectors in $\mathbf{R}^2$ or in $\mathbf{R}^3$.

The standard coordinate vectors in $\mathbf{R}^2$ and $\mathbf{R}^3$ are pictured below.



These are the vectors of length 1 that point in the positive directions of each of the axes.

Multiplying a matrix by the standard coordinate vectors. *If A is an $m \times n$ matrix with columns $v_1, v_2, \ldots, v_m$, then $Ae_i = v_i$ for each $i = 1, 2, \ldots, n$:*

$$\begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} e_i = v_i.$$

In other words, multiplying a matrix by e_i simply selects its ith column.

For example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}.$$

Definition. The $n \times n$ **identity matrix** is the matrix I_n whose columns are the n standard coordinate vectors in $\mathbf{R}^n$:

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

We will see in this example below that the identity matrix is the matrix of the identity transformation.

3.3.3 The Matrix of a Linear Transformation

Now we can prove that every linear transformation is a matrix transformation, and we will show how to compute the matrix.

Theorem (The matrix of a linear transformation)**.** *Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Let A be the $m \times n$ matrix*

$$A = \begin{pmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{pmatrix}.$$

Then T is the matrix transformation associated with A: that is, $T(x) = Ax$.

Proof. We suppose for simplicity that T is a transformation from $\mathbf{R}^3$ to $\mathbf{R}^2$. Let A be the matrix given in the statement of the theorem. Then

$$\begin{aligned} T\begin{pmatrix} x \\ y \\ z \end{pmatrix} &:= T\left(x\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \\ &= T\big(xe_1 + ye_2 + ze_3\big) \\ &= xT(e_1) + yT(e_2) + zT(e_3) \\ &= \begin{pmatrix} | & | & | \\ T(e_1) & T(e_2) & T(e_3) \\ | & | & | \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= A\begin{pmatrix} x \\ y \\ z \end{pmatrix}. \end{aligned}$$ □

The matrix A in the above theorem is called the **standard matrix** for T. The columns of A are the vectors obtained by evaluating T on the n standard coordinate vectors in $\mathbf{R}^n$. To summarize part of the theorem:

Matrix transformations are the same as linear transformations.

Dictionary. Linear transformations are the same as matrix transformations, which come from matrices. The correspondence can be summarized in the following dictionary.

$T: \mathbf{R}^n \to \mathbf{R}^m$
Linear transformation $\longrightarrow$ $m \times n$ matrix $A = \begin{pmatrix} | & | & & | \\ T(e_1) & T(e_2) & \cdots & T(e_n) \\ | & | & & | \end{pmatrix}$

$T: \mathbf{R}^n \to \mathbf{R}^m$
$T(x) = Ax$ $\longleftarrow$ $m \times n$ matrix A

Example (The matrix of a dilation)**.** Define $T: \mathbf{R}^2 \to \mathbf{R}^2$ by $T(x) = 1.5x$. Find the standard matrix A for T.

Solution. The columns of A are obtained by evaluating T on the standard coordinate vectors e_1, e_2.

$$\left.\begin{aligned} T(e_1) &= 1.5e_1 = \begin{pmatrix} 1.5 \\ 0 \end{pmatrix} \\ T(e_2) &= 1.5e_2 = \begin{pmatrix} 0 \\ 1.5 \end{pmatrix} \end{aligned}\right\} \implies A = \begin{pmatrix} 1.5 & 0 \\ 0 & 1.5 \end{pmatrix}.$$

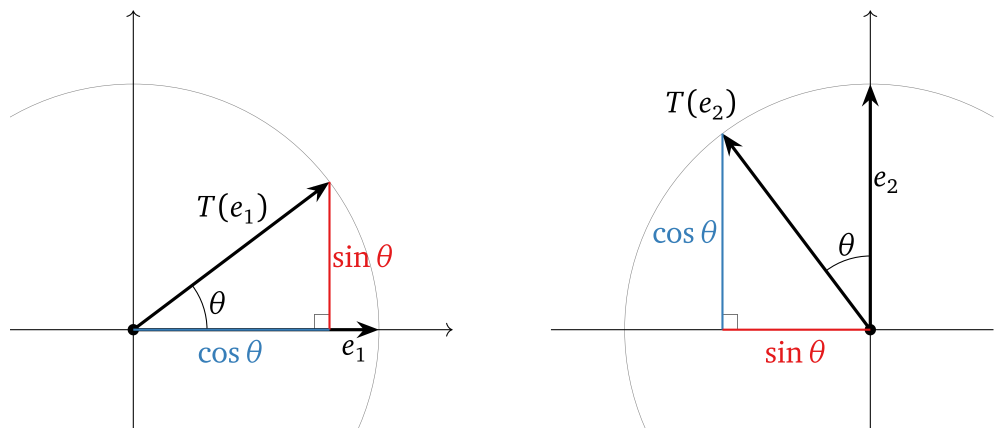
This is the matrix we started with in this example in Section 3.1.

Example (The matrix of a rotation)**.** Define $T: \mathbf{R}^2 \to \mathbf{R}^2$ by

$$T(x) = \text{ the vector } x \text{ rotated counterclockwise by the angle } \theta.$$

Find the standard matrix for T.

Solution. The columns of A are obtained by evaluating T on the standard coordinate vectors e_1, e_2. In order to compute the entries of $T(e_1)$ and $T(e_2)$, we have to do some trigonometry.



We see from the picture that

$$\left.\begin{aligned} T(e_1) &= \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \\ T(e_2) &= \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix} \end{aligned}\right\} \implies A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

We saw in the above example that the matrix for counterclockwise rotation of the plane by an angle of θ is

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Example (A transformation defined by a formula)**.** Define $T: \mathbf{R}^2 \to \mathbf{R}^3$ by the formula

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3x - y \\ y \\ x \end{pmatrix}.$$

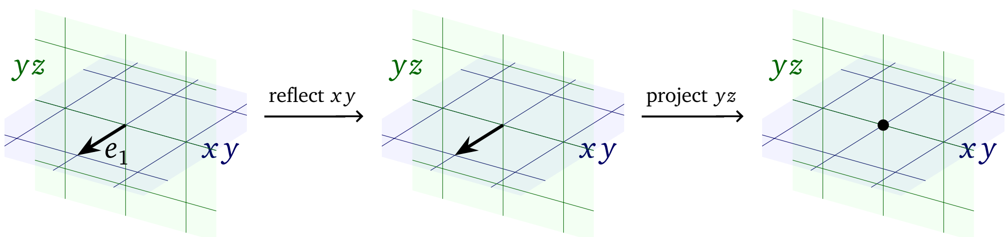
Find the standard matrix for T.

Solution. We substitute the standard coordinate vectors into the formula defining T:

$$\left.\begin{aligned} T(e_1) &= T\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}3(1)-0\\0\\1\end{pmatrix} = \begin{pmatrix}3\\0\\1\end{pmatrix} \\ T(e_2) &= T\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}3(0)-1\\1\\0\end{pmatrix} = \begin{pmatrix}-1\\1\\0\end{pmatrix} \end{aligned}\right\} \implies A = \begin{pmatrix}3 & -1\\0 & 1\\1 & 0\end{pmatrix}.$$

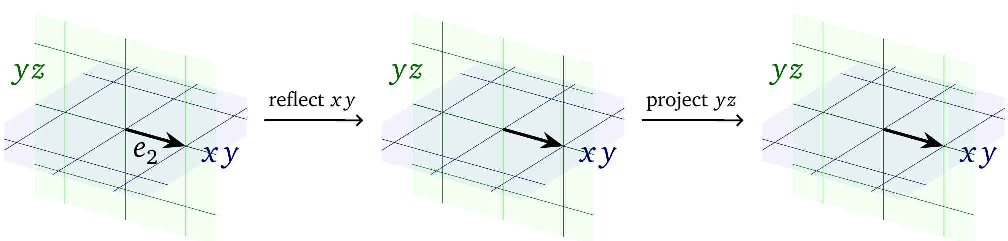
Example (A transformation defined in steps)**.** Let $T: \mathbf{R}^3 \to \mathbf{R}^3$ be the linear transformation that reflects over the xy-plane and then projects onto the yz-plane. What is the standard matrix for T?

Solution. This transformation is described geometrically, in two steps. To find the columns of A, we need to follow the standard coordinate vectors through each of these steps.



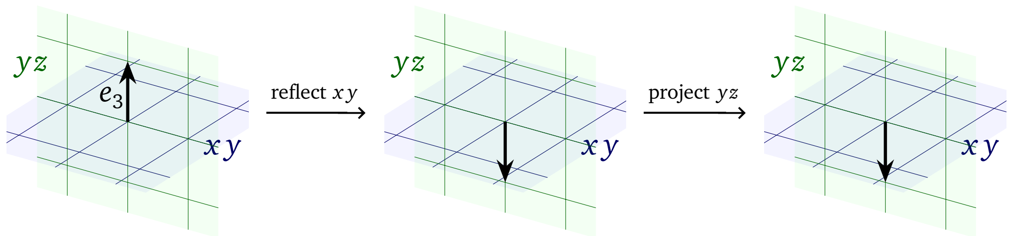
Since e_1 lies on the xy-plane, reflecting over the xy-plane does not move e_1. Since e_1 is perpendicular to the yz-plane, projecting e_1 onto the yz-plane sends it to zero. Therefore,

$$T(e_1) = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$



Since e_2 lies on the xy-plane, reflecting over the xy-plane does not move e_2. Since e_2 lies on the yz-plane, projecting onto the yz-plane does not move e_2 either. Therefore,

$$T(e_2) = e_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}.$$



Since e_3 is perpendicular to the xy-plane, reflecting over the xy-plane takes e_3 to its negative. Since $-e_3$ lies on the yz-plane, projecting onto the yz-plane does not move it. Therefore,

$$T(e_3) = -e_3 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

Now we have computed all three columns of A:

$$\left.\begin{aligned} T(e_1) &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ T(e_2) &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\ T(e_1) &= \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \end{aligned}\right\} \implies A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Use this link to view the online demo

Illustration of a transformation defined in steps. Click and drag the vector on the left.

Recall from this definition in Section 3.1 that the *identity transformation* is the transformation $\mathrm{Id}_{\mathbf{R}^n} : \mathbf{R}^n \to \mathbf{R}^n$ defined by $\mathrm{Id}_{\mathbf{R}^n}(x) = x$ for every vector x.

Example (The standard matrix of the identity transformation)**.** Verify that the identity transformation $\mathrm{Id}_{\mathbf{R}^n} : \mathbf{R}^n \to \mathbf{R}^n$ is linear, and compute its standard matrix.

Solution. We verify the two defining properties of linear transformations. Let u, v be vectors in $\mathbf{R}^n$. Then

$$\mathrm{Id}_{\mathbf{R}^n}(u + v) = u + v = \mathrm{Id}_{\mathbf{R}^n}(u) + \mathrm{Id}_{\mathbf{R}^n}(v).$$

If c is a scalar, then

$$\mathrm{Id}_{\mathbf{R}^n}(cu) = cu = c\,\mathrm{Id}_{\mathbf{R}^n}(u).$$

Since $\mathrm{Id}_{\mathbf{R}^n}$ satisfies the two defining properties, it is a linear transformation.

Now that we know that $\mathrm{Id}_{\mathbf{R}^n}$ is linear, it makes sense to compute its standard matrix. For each standard coordinate vector e_i, we have $\mathrm{Id}_{R^n}(e_i) = e_i$. In other words, the columns of the standard matrix of $\mathrm{Id}_{\mathbf{R}^n}$ are the standard coordinate vectors, so the standard matrix is the identity matrix

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

We computed in this example that the matrix of the identity transform is the identity matrix: for every x in $\mathbf{R}^n$,

$$x = \mathrm{Id}_{\mathbf{R}^n}(x) = I_n x.$$

Therefore, $I_n x = x$ for all vectors x: the product of the identity matrix and a vector is the same vector.

3.4 Matrix Multiplication

Objectives

1. Understand compositions of transformations.
2. Understand the relationship between matrix products and compositions of matrix transformations.
3. Become comfortable doing basic algebra involving matrices.
4. *Recipe:* matrix multiplication (two ways).
5. *Picture:* composition of transformations.
6. *Vocabulary word:* **composition**.

In this section, we study compositions of transformations. As we will see, composition is a way of chaining transformations together. The composition of matrix transformations corresponds to a notion of *multiplying* two matrices together. We also discuss addition and scalar multiplication of transformations and of matrices.

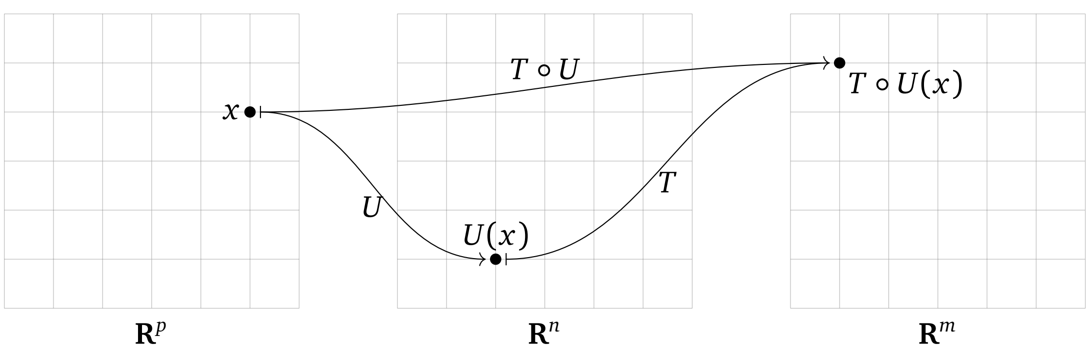
3.4.1 Composition of linear transformations

Composition means the same thing in linear algebra as it does in Calculus. Here is the definition.

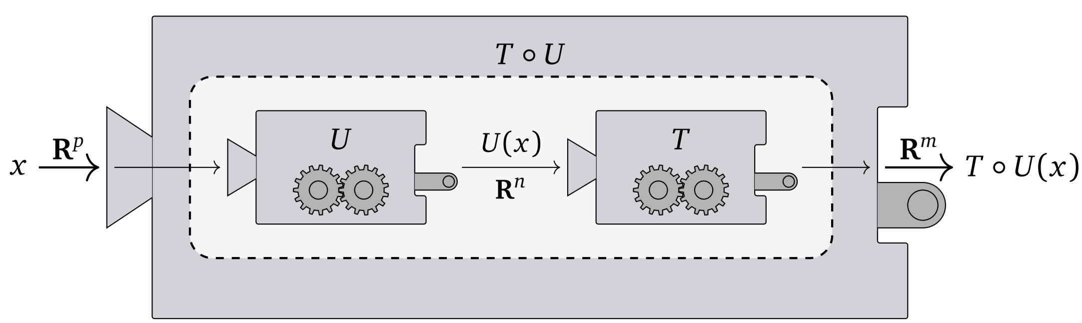
Definition. Let $T: \mathbf{R}^n \to \mathbf{R}^m$ and $U: \mathbf{R}^p \to \mathbf{R}^n$ be transformations. Their **composition** is the transformation $T \circ U: \mathbf{R}^p \to \mathbf{R}^m$ defined by

$$(T \circ U)(x) = T(U(x)).$$

Composing two transformations means chaining them together: $T \circ U$ is the transformation that first applies U, then applies T (note the order of operations). More precisely, to evaluate $T \circ U$ on an input vector x, first you evaluate $U(x)$, then you take this output vector of U and use it as an input vector of T: that is, $(T \circ U)(x) = T(U(x))$. Of course, this only makes sense when the outputs of U are valid inputs of T, that is, when the range of U is contained in the domain of T.



Here is a picture of the composition $T \circ U$ as a "machine" that first runs U, then takes its output and feeds it into T; there is a similar picture in this subsection in Section 3.1.



Domain and codomain of a composition.

- In order for $T \circ U$ to be defined, the codomain of U must equal the domain of T.
- The domain of $T \circ U$ is the domain of U.
- The codomain of $T \circ U$ is the codomain of T.

Example (Functions of one variable)**.** Define $f : \mathbf{R} \to \mathbf{R}$ by $f(x) = x^2$ and $g : \mathbf{R} \to \mathbf{R}$ by $g(x) = x^3$. The composition $f \circ g : \mathbf{R} \to \mathbf{R}$ is the transformation defined by the rule

$$f \circ g(x) = f(g(x)) = f(x^3) = (x^3)^2 = x^6.$$

For instance, $f \circ g(-2) = f(-8) = 64$.

Interactive: A composition of matrix transformations. Define $T : \mathbf{R}^3 \to \mathbf{R}^2$ and $U : \mathbf{R}^2 \to \mathbf{R}^3$ by

$$T(x) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} x \quad \text{and} \quad U(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} x.$$

Their composition is a transformation $T \circ U : \mathbf{R}^2 \to \mathbf{R}^2$; it turns out to be the matrix transformation associated to the matrix $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$.

Use this link to view the online demo

A composition of two matrix transformations, i.e., a transformation performed in two steps. On the left is the domain of U/the domain of $T \circ U$; in the middle is the codomain of U/the domain of T, and on the right is the codomain of T/the codomain of $T \circ U$. The vector x is the input of U and of $T \circ U$; the vector in the middle is the output of U/the input of T, and the vector on the right is the output of T/of $T \circ U$. Click and drag x.

Interactive: A transformation defined in steps. Let $S : \mathbf{R}^3 \to \mathbf{R}^3$ be the linear transformation that first reflects over the xy-plane and then projects onto the yz-plane, as in this example in Section 3.3. The transformation S is the composition $T \circ U$, where $U : \mathbf{R}^3 \to \mathbf{R}^3$ is the transformation that reflects over the xy-plane, and $T : \mathbf{R}^3 \to \mathbf{R}^3$ is the transformation that projects onto the yz-plane.

Use this link to view the online demo

Illustration of a transformation defined in steps. On the left is the domain of U/the domain of S; in the middle is the codomain of U/the domain of T, and on the right is the codomain of T/the codomain of S. The vector u is the input of U and of S; the vector in the middle is the output of U/the input of T, and the vector on the right is the output of T/of S. Click and drag u.

Interactive: A transformation defined in steps. Let $S\colon \mathbf{R}^3 \to \mathbf{R}^3$ be the linear transformation that first projects onto the xy-plane, and then projects onto the xz-plane. The transformation S is the composition $T \circ U$, where $U\colon \mathbf{R}^3 \to \mathbf{R}^3$ is the transformation that projects onto the xy-plane, and $T\colon \mathbf{R}^3 \to \mathbf{R}^3$ is the transformation that projects onto the xz-plane.

Use this link to view the online demo

Illustration of a transformation defined in steps. Note that projecting onto the xy-plane, followed by projecting onto the xz-plane, is the projection onto the x-axis.

Recall from this definition in Section 3.1 that the *identity transformation* is the transformation $\mathrm{Id}_{\mathbf{R}^n}\colon \mathbf{R}^n \to \mathbf{R}^n$ defined by $\mathrm{Id}_{\mathbf{R}^n}(x) = x$ for every vector x.

Properties of composition. Let S, T, U be transformations and let c be a scalar. Suppose that $T\colon \mathbf{R}^n \to \mathbf{R}^m$, and that in each of the following identities, the domains and the codomains are compatible when necessary for the composition to be defined. The following properties are easily verified:

$$\begin{aligned} S\circ(T+U) &= S\circ T + S\circ U & (S+T)\circ U &= S\circ U + T\circ U \\ c(T\circ U) &= (cT)\circ U & c(T\circ U) &= T\circ(cU) \quad \text{if } T \text{ is linear} \\ T\circ \mathrm{Id}_{\mathbf{R}^n} &= T & \mathrm{Id}_{\mathbf{R}^m}\circ T &= T \\ & & S\circ(T\circ U) &= (S\circ T)\circ U \end{aligned}$$

The final property is called **associativity**. Unwrapping both sides, it says:

$$S\circ(T\circ U)(x) = S(T\circ U(x)) = S(T(U(x))) = S\circ T(U(x)) = (S\circ T)\circ U(x).$$

In other words, both $S\circ(T\circ U)$ and $(S\circ T)\circ U$ are the transformation defined by first applying U, then T, then S.

Composition of transformations is *not* commutative in general. That is, in general, $T\circ U \neq U\circ T$, even when both compositions are defined.

Example (Functions of one variable)**.** Define $f\colon \mathbf{R}\to\mathbf{R}$ by $f(x) = x^2$ and $g\colon \mathbf{R}\to\mathbf{R}$ by $g(x) = e^x$. The composition $f\circ g\colon \mathbf{R}\to\mathbf{R}$ is the transformation defined by the rule

$$f\circ g(x) = f(g(x)) = f(e^x) = (e^x)^2 = e^{2x}.$$

The composition $g \circ f : \mathbf{R} \to \mathbf{R}$ is the transformation defined by the rule

$$g \circ f(x) = g(f(x)) = g(x^2) = e^{x^2}.$$

Note that $e^{x^2} \neq e^{2x}$ in general; for instance, if $x = 1$ then $e^{x^2} = e$ and $e^{2x} = e^2$. Thus $f \circ g$ is not equal to $g \circ f$, and we can already see with functions of one variable that composition of functions is not commutative.

Example (Non-commutative composition of transformations)**.** Define matrix transformations $T, U : \mathbf{R}^2 \to \mathbf{R}^2$ by

$$T(x) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} x \quad \text{and} \quad U(x) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} x.$$

Geometrically, T is a shear in the x-direction, and U is a shear in the Y-direction. We evaluate

$$T \circ U \begin{pmatrix} 1 \\ 0 \end{pmatrix} = T \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

and

$$U \circ T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = U \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Since $T \circ U$ and $U \circ T$ have different outputs for the input vector $\binom{1}{0}$, they are different transformations. (See this example.)

Use this link to view the online demo

Illustration of the composition $T \circ U$.

Use this link to view the online demo

Illustration of the composition $U \circ T$.

3.4.2 Matrix multiplication

In this subsection, we introduce a seemingly unrelated operation on matrices, namely, matrix multiplication. As we will see in the next subsection, matrix multiplication exactly corresponds to the composition of the corresponding linear transformations. First we need some terminology.

Notation. Let A be an $m \times n$ matrix. We will generally write a_{ij} for the entry in the ith row and the jth column. It is called the i, j **entry** of the matrix.

$$\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix}$$

ith row

jth column

Definition (Matrix multiplication)**.** Let A be an $m \times n$ matrix and let B be an $n \times p$ matrix. Denote the columns of B by $v_1, v_2, \ldots, v_p$:

$$B = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_p \\ | & | & & | \end{pmatrix}.$$

The **product** AB is the $m \times p$ matrix with columns $Av_1, Av_2, \ldots, Av_p$:

$$AB = \begin{pmatrix} | & | & & | \\ Av_1 & Av_2 & \cdots & Av_p \\ | & | & & | \end{pmatrix}.$$

In other words, matrix multiplication is defined column-by-column, or "distributes over the columns of B."

Example.

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \left(\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right)$$
$$= \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

In order for the vectors $Av_1, Av_2, \ldots, Av_p$ to be defined, the numbers of rows of B has to equal the number of columns of A.

The sizes of the matrices in the matrix product.

- In order for AB to be defined, the number of rows of B has to equal the number of columns of A.
- The product of an $m \times n$ matrix and an $n \times p$ matrix is an $m \times p$ matrix.

If B has only one column, then AB also has one column. A matrix with one column is the same as a vector, so the definition of the matrix product generalizes the definition of the matrix-vector product from this definition in Section 2.3.

If A is a square matrix, then we can multiply it by itself; we define its **powers** to be

$$A^2 = AA \qquad A^3 = AAA \qquad \text{etc.}$$

The row-column rule for matrix multiplication Recall from this definition in Section 2.3 that the product of a row vector and a column vector is the scalar

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

The following procedure for finding the matrix product is much better adapted to computations by hand; the previous definition is more suitable for proving theorems, such as this theorem below.

Recipe: The row-column rule for matrix multiplication. Let A be an $m \times n$ matrix, let B be an $n \times p$ matrix, and let $C = AB$. Then the ij entry of C is the ith row of A times the jth column of B:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

Here is a diagram:

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ik} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{k1} & \cdots & b_{kj} & \cdots & b_{kp} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{np} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & c_{1j} & \cdots & c_{1p} \\ \vdots & & \vdots & & \vdots \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{ip} \\ \vdots & & \vdots & & \vdots \\ c_{m1} & \cdots & c_{mj} & \cdots & c_{mp} \end{pmatrix}$$

ith row

jth column

ij entry

Proof. The row-column rule for matrix-vector multiplication in Section 2.3 says that if A has rows $r_1, r_2, \ldots, r_m$ and x is a vector, then

$$Ax = \begin{pmatrix} — r_1 — \\ — r_2 — \\ \vdots \\ — r_m — \end{pmatrix} x = \begin{pmatrix} r_1x \\ r_2x \\ \vdots \\ r_mx \end{pmatrix}.$$

The definition of matrix multiplication is

$$A \begin{pmatrix} | & | & & | \\ c_1 & c_2 & \cdots & c_p \\ | & | & & | \end{pmatrix} = \begin{pmatrix} | & | & & | \\ Ac_1 & Ac_2 & \cdots & Ac_p \\ | & | & & | \end{pmatrix}.$$

It follows that

$$\begin{pmatrix} — r_1 — \\ — r_2 — \\ \vdots \\ — r_m — \end{pmatrix} \begin{pmatrix} | & | & & | \\ c_1 & c_2 & \cdots & c_p \\ | & | & & | \end{pmatrix} = \begin{pmatrix} r_1c_1 & r_1c_2 & \cdots & r_1c_p \\ r_2c_1 & r_2c_2 & \cdots & r_2c_p \\ \vdots & \vdots & & \vdots \\ r_mc_1 & r_mc_2 & \cdots & r_mc_p \end{pmatrix}.$$

□

Example. The row-column rule allows us to compute the product matrix one entry at a time:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 2\cdot 2 + 3\cdot 3 & \square \\ \square & \square \end{pmatrix} = \begin{pmatrix} 14 & \square \\ \square & \square \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}\begin{pmatrix} 1 & -3 \\ 2 & -2 \\ 3 & -1 \end{pmatrix} = \begin{pmatrix} \square & \square \\ 4\cdot 1 + 5\cdot 2 + 6\cdot 3 & \square \end{pmatrix} = \begin{pmatrix} 14 & \square \\ 32 & \square \end{pmatrix}$$

You should try to fill in the other two boxes!

Although matrix multiplication satisfies many of the properties one would expect (see the end of the section), one must be careful when doing matrix arithmetic, as there are several properties that are not satisfied in general.

Matrix multiplication caveats.

- Matrix multiplication is not commutative: AB is not usually equal to BA, even when both products are defined and have the same size. See this example.
- Matrix multiplication does not satisfy the cancellation law: $AB = AC$ does not imply $B = C$, even when $A \neq 0$. For example,

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 5 & 6 \end{pmatrix}.$$

- It is possible for $AB = 0$, even when $A \neq 0$ and $B \neq 0$. For example,

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

While matrix multiplication is not commutative in general there are examples of matrices A and B with $AB = BA$. For example, this always works when A is the zero matrix, or when $A = B$. The reader is encouraged to find other examples.

Example (Non-commutative multiplication of matrices)**.** Consider the matrices

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

as in this example. The matrix AB is

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

whereas the matrix BA is

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

In particular, we have

$$AB \neq BA.$$

And so matrix multiplication is not always commutative. It is not a coincidence that this example agrees with the previous example; we are about to see that multiplication of matrices corresponds to composition of transformations.

Order of Operations. Let $T : \mathbf{R}^n \to \mathbf{R}^m$ and $U : \mathbf{R}^p \to \mathbf{R}^n$ be linear transformations, and let A and B be their standard matrices, respectively. Recall that $T \circ U(x)$ is the vector obtained by first applying U to x, and then T.

On the matrix side, the standard matrix of $T \circ U$ is the product AB, so $T \circ U(x) = (AB)x$. By associativity of matrix multiplication, we have $(AB)x = A(Bx)$, so the product $(AB)x$ can be computed by first multiplying x by B, *then* multipyling the product by A.

Therefore, matrix multiplication happens in the same order as composition of transformations. In other words, *both matrices and transformations are written in the order opposite from the order in which they act.* But matrix multiplication and composition of transformations are written in the same order as each other: the matrix for $T \circ U$ is AB.

3.4.3 Composition and Matrix Multiplication

The point of this subsection is to show that matrix multiplication corresponds to composition of transformations, that is, the standard matrix for $T \circ U$ is the product of the standard matrices for T and for U. It should be hard to believe that our complicated formula for matrix multiplication actually means something intuitive such as "chaining two transformations together"!

Theorem. *Let $T : \mathbf{R}^n \to \mathbf{R}^m$ and $U : \mathbf{R}^p \to \mathbf{R}^n$ be linear transformations, and let A and B be their standard matrices, respectively, so A is an $m \times n$ matrix and B is an $n \times p$ matrix. Then $T \circ U : \mathbf{R}^p \to \mathbf{R}^m$ is a linear transformation, and its standard matrix is the product AB.*

Proof. First we verify that $T \circ U$ is linear. Let u, v be vectors in $\mathbf{R}^p$. Then

$$\begin{aligned} T \circ U(u+v) &= T(U(u+v)) = T(U(u)+U(v)) \\ &= T(U(u)) + T(U(v)) = T \circ U(u) + T \circ U(v). \end{aligned}$$

If c is a scalar, then

$$T \circ U(cv) = T(U(cv)) = T(cU(v)) = cT(U(v)) = cT \circ U(v).$$

Since $T \circ U$ satisfies the two defining properties in Section 3.3, it is a linear transformation.

Now that we know that $T \circ U$ is linear, it makes sense to compute its standard matrix. Let C be the standard matrix of $T \circ U$, so $T(x) = Ax$, $U(x) = Bx$, and $T \circ U(x) = Cx$. By this theorem in Section 3.3, the first column of C is Ce_1, and the first column of B is Be_1. We have

$$T \circ U(e_1) = T(U(e_1)) = T(Be_1) = A(Be_1).$$

By definition, the first column of the product AB is the product of A with the first column of B, which is Be_1, so

$$Ce_1 = T \circ U(e_1) = A(Be_1) = (AB)e_1.$$

It follows that C has the same first column as AB. The same argument as applied to the ith standard coordinate vector e_i shows that C and AB have the same ith column; since they have the same columns, they are the same matrix. □

The theorem justifies our choice of definition of the matrix product. This is the one and only reason that matrix products are defined in this way. To rephrase:

Products and compositions. The matrix of the composition of two linear transformations is the product of the matrices of the transformations.

Example (Composition of rotations)**.** In this example in Section 3.3, we showed that the standard matrix for the counterclockwise rotation of the plane by an angle of θ is

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be counterclockwise rotation by 45°, and let $U : \mathbf{R}^2 \to \mathbf{R}^2$ be counterclockwise rotation by 90°. The matrices A and B for T and U are, respectively,

$$A = \begin{pmatrix} \cos(45°) & -\sin(45°) \\ \sin(45°) & \cos(45°) \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$
$$B = \begin{pmatrix} \cos(90°) & -\sin(90°) \\ \sin(90°) & \cos(90°) \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Here we used the trigonometric identities

$$\cos(45°) = \frac{1}{\sqrt{2}} \qquad \sin(45°) = \frac{1}{\sqrt{2}}$$
$$\cos(90°) = 0 \qquad \sin(90°) = 1.$$

The standard matrix of the composition $T \circ U$ is

$$AB = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}.$$

This is consistent with the fact that $T \circ U$ is counterclockwise rotation by $90°+45° = 135°$: we have

$$\begin{pmatrix} \cos(135°) & -\sin(135°) \\ \sin(135°) & \cos(135°) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}$$

because $\cos(135°) = -1/\sqrt{2}$ and $\sin(135°) = 1/\sqrt{2}$.

Challenge. Derive the trigonometric identities

$$\sin(\alpha \pm \beta) = \sin(\alpha)\cos(\beta) \pm \cos(\alpha)\sin(\beta)$$

and

$$\cos(\alpha \pm \beta) = \cos(\alpha)\cos(\beta) \mp \sin(\alpha)\sin(\beta)$$

using the above theorem as applied to rotation transformations, as in the previous example.

Interactive: A composition of matrix transformations. Define $T : \mathbf{R}^3 \to \mathbf{R}^2$ and $U : \mathbf{R}^2 \to \mathbf{R}^3$ by

$$T(x) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} x \quad \text{and} \quad U(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} x.$$

Their composition is a linear transformation $T \circ U : \mathbf{R}^2 \to \mathbf{R}^2$. By the theorem, its standard matrix is

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

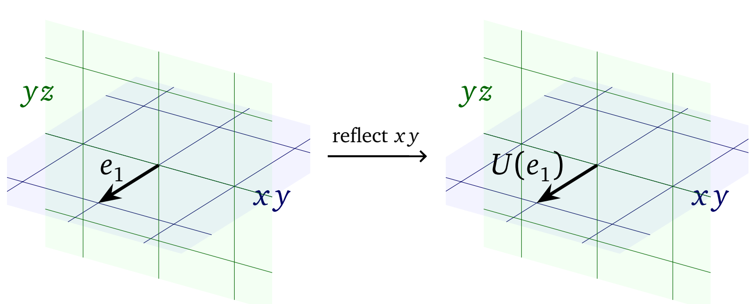
as we computed in the above example.

Use this link to view the online demo

The matrix of the composition $T \circ U$ is the product of the matrices for T and U.

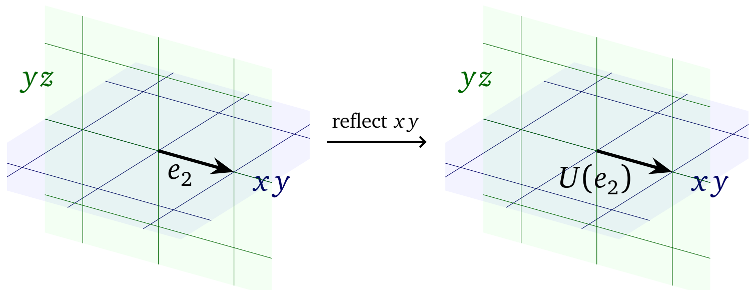
Interactive: A transformation defined in steps. Let $S : \mathbf{R}^3 \to \mathbf{R}^3$ be the linear transformation that first reflects over the xy-plane and then projects onto the yz-plane, as in this example in Section 3.3. The transformation S is the composition $T \circ U$, where $U : \mathbf{R}^3 \to \mathbf{R}^3$ is the transformation that reflects over the xy-plane, and $T : \mathbf{R}^3 \to \mathbf{R}^3$ is the transformation that projects onto the yz-plane.

Let us compute the matrix B for U.



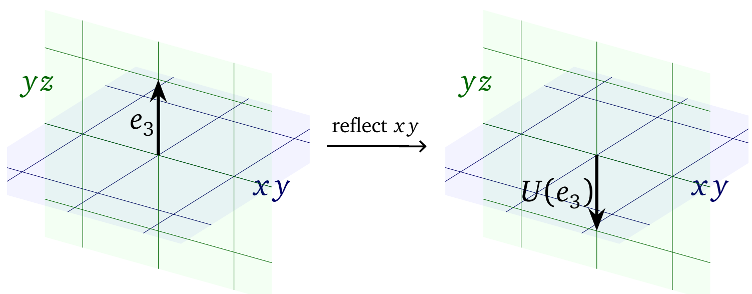
Since e_1 lies on the xy-plane, reflecting it over the xy-plane does not move it:

$$U(e_1) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$



Since e_2 lies on the xy-plane, reflecting over the xy-plane does not move it either:

$$U(e_2) = e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$



Since e_3 is perpendicular to the xy-plane, reflecting over the xy-plane takes e_3 to its negative:

$$U(e_3) = -e_3 = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}.$$

We have computed all of the columns of B:

$$B = \begin{pmatrix} | & | & | \\ U(e_1) & U(e_2) & U(e_3) \\ | & | & | \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

By a similar method, we find

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It follows that the matrix for $S = T \circ U$ is

$$\begin{aligned} AB &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\ &= \left(\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \;\; \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \;\; \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \right) \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \end{aligned}$$

as we computed in this example in Section 3.3.

Use this link to view the online demo

Recall from this definition in Section 3.3 that the *identity matrix* is the $n \times n$ matrix I_n whose columns are the standard coordinate vectors in $\mathbf{R}^n$. The identity matrix is the standard matrix of the identity transformation: that is, $x = \mathrm{Id}_{\mathbf{R}^n}(x) = I_n x$ for all vectors x in $\mathbf{R}^n$. For any linear transformation $T : \mathbf{R}^n \to \mathbf{R}^m$ we have

$$I_{\mathbf{R}^m} \circ T = T$$

and by the same token we have for any $m \times n$ matrix A we have

$$I_m A = A.$$

Similarly, we have $T \circ I_{\mathbf{R}^n} = T$ and $AI_n = A$.

3.4.4 The algebra of transformations and matrices

In this subsection we describe two more operations that one can perform on transformations: addition and scalar multiplication. We then translate these operations into the language of matrices. This is analogous to what we did for the composition of linear transformations, but much less subtle.

Definition.

- Let $T, U\colon \mathbf{R}^n \to \mathbf{R}^m$ be two transformations. Their **sum** is the transformation $T + U\colon \mathbf{R}^n \to \mathbf{R}^m$ defined by

$$(T + U)(x) = T(x) + U(x).$$

Note that addition of transformations is only defined when both transformations have the same domain and codomain.

- Let $T\colon \mathbf{R}^n \to \mathbf{R}^m$ be a transformation, and let c be a scalar. The **scalar product** of c with T is the transformation $cT\colon \mathbf{R}^n \to \mathbf{R}^m$ defined by

$$(cT)(x) = c \cdot T(x).$$

To emphasize, the sum of two transformations $T, U\colon \mathbf{R}^n \to \mathbf{R}^m$ is another transformation called $T + U$; its value on an input vector x is the sum of the outputs of T and U. Similarly, the product of T with a scalar c is another transformation called cT; its value on an input vector x is the vector $c \cdot T(x)$.

Example (Functions of one variable)**.** Define $f\colon \mathbf{R} \to \mathbf{R}$ by $f(x) = x^2$ and $g\colon \mathbf{R} \to \mathbf{R}$ by $g(x) = x^3$. The sum $f + g\colon \mathbf{R} \to \mathbf{R}$ is the transformation defined by the rule

$$(f + g)(x) = f(x) + g(x) = x^2 + x^3.$$

For instance, $(f + g)(-2) = (-2)^2 + (-2)^3 = -4$.

Define $\exp\colon \mathbf{R} \to \mathbf{R}$ by $\exp(x) = e^x$. The product $2\exp\colon \mathbf{R} \to \mathbf{R}$ is the transformation defined by the rule

$$(2\exp)(x) = 2 \cdot \exp(x) = 2e^x.$$

For instance, $(2\exp)(1) = 2 \cdot \exp(1) = 2e$.

Properties of addition and scalar multiplication for transformations. Let $S, T, U\colon \mathbf{R}^n \to \mathbf{R}^m$ be transformations and let c, d be scalars. The following properties are easily verified:

$$\begin{aligned} T + U &= U + T & S + (T + U) &= (S + T) + U \\ c(T + U) &= cT + cU & (c + d)T &= cT + dT \\ c(dT) &= (cd)T & T + 0 &= T \end{aligned}$$

In one of the above properties, we used 0 to denote the transformation $\mathbf{R}^n \to \mathbf{R}^m$ that is zero on every input vector: $0(x) = 0$ for all x. This is called the **zero transformation**.

We now give the analogous operations for matrices.

Definition.

- The **sum** of two $m \times n$ matrices is the matrix obtained by summing the entries of A and B individually:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & a_{13}+b_{13} \\ a_{21}+b_{21} & a_{22}+b_{22} & a_{23}+b_{23} \end{pmatrix}$$

  In other words, the i,j entry of $A+B$ is the sum of the i,j entries of A and B. Note that addition of matrices is only defined when both matrices have the same size.

- The **scalar product** of a scalar c with a matrix A is obtained by scaling all entries of A by c:

$$c \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} & ca_{13} \\ ca_{21} & ca_{22} & ca_{23} \end{pmatrix}$$

  In other words, the i,j entry of cA is c times the i,j entry of A.

Fact. *Let $T, U : \mathbf{R}^n \to \mathbf{R}^m$ be linear transformations with standard matrices A, B, respectively, and let c be a scalar.*

- *The standard matrix for $T + U$ is $A + B$.*
- *The standard matrix for cT is cA.*

In view of the above fact, the following properties are consequences of the corresponding properties of transformations. They are easily verified directly from the definitions as well.

Properties of addition and scalar multiplication for matrices. Let A, B, C be $m \times n$ matrices and let c, d be scalars. Then:

$$\begin{aligned} A+B &= B+A & C+(A+B) &= (C+A)+B \\ c(A+B) &= cA+cB & (c+d)A &= cA+dA \\ c(dA) &= (cd)A & A+0 &= A \end{aligned}$$

In one of the above properties, we used 0 to denote the $m \times n$ matrix whose entries are all zero. This is the standard matrix of the zero transformation, and is called the **zero matrix**.

We can also combine addition and scalar multiplication of matrices with multiplication of matrices. Since matrix multiplication corresponds to composition of transformations (theorem), the following properties are consequences of the corresponding properties of transformations.

Properties of matrix multiplication. Let A, B, C be matrices and let c be a scalar. Suppose that A is an $m \times n$ matrix, and that in each of the following identities, the sizes of B and C are compatible when necessary for the product to be defined. Then:

$$\begin{aligned} C(A+B) &= CA+CB & (A+B)C &= AC+BC \\ c(AB) &= (cA)B & c(AB) &= A(cB) \\ AI_n &= A & I_mA &= A \\ (AB)C &= A(BC) \end{aligned}$$

Most of the above properties are easily verified directly from the definitions. The *associativity* property $(AB)C = A(BC)$, however, is not (try it!). It is much easier to prove by relating matrix multiplication to composition of transformations, and using the obvious fact that composition of transformations is associative.

3.5 Matrix Inverses

Objectives

1. Understand what it means for a square matrix to be invertible.
2. Learn about invertible transformations, and understand the relationship between invertible matrices and invertible transformations.
3. *Recipes:* compute the inverse matrix, solve a linear system by taking inverses.
4. *Picture:* the inverse of a transformation.
5. *Vocabulary words:* **inverse matrix**, **inverse transformation**.

In Section 3.1 we learned to multiply matrices together. In this section, we learn to "divide" by a matrix. This allows us to solve the matrix equation $Ax = b$ in an elegant way:

$$Ax = b \quad \iff \quad x = A^{-1}b.$$

One has to take care when "dividing by matrices", however, because not every matrix has an inverse, and the order of matrix multiplication is important.

3.5.1 Invertible Matrices

The *reciprocal* or *inverse* of a nonzero number a is the number b which is characterized by the property that $ab = 1$. For instance, the inverse of 7 is $1/7$. We use this formulation to define the inverse of a matrix.

Definition. Let A be an $n \times n$ (square) matrix. We say that A is **invertible** if there is an $n \times n$ matrix B such that

$$AB = I_n \quad \text{and} \quad BA = I_n.$$

In this case, the matrix B is called the **inverse** of A, and we write $B = A^{-1}$.

We have to require $AB = I_n$ *and* $BA = I_n$ because in general matrix multiplication is not commutative. However, we will show in this corollary in Section 3.6 that if A and B are $n \times n$ matrices such that $AB = I_n$, then automatically $BA = I_n$.

Example. Verify that the matrices

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

are inverses.

Solution. We will check that $AB = I_2$ and that $BA = I_2$.

$$AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$BA = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Therefore, A is invertible, with inverse B.

Remark. There exist non-square matrices whose product is the identity. Indeed, if

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

then $AB = I_2$. However, $BA \neq I_3$, so B does not deserve to be called the inverse of A.

One can show using the ideas later in this section that if A is an $n \times m$ matrix for $n \neq m$, then there is no $m \times n$ matrix B such that $AB = I_m$ and $BA = I_n$. For this reason, we restrict ourselves to *square* matrices when we discuss matrix invertibility.

Facts about invertible matrices. *Let A and B be invertible $n \times n$ matrices.*

1. *A^{-1} is invertible, and its inverse is $(A^{-1})^{-1} = A$.*
2. *AB is invertible, and its inverse is $(AB)^{-1} = B^{-1}A^{-1}$ (note the order).*

Proof.

1. The equations $AA^{-1} = I_n$ and $A^{-1}A = I_n$ at the same time exhibit A^{-1} as the inverse of A and A as the inverse of A^{-1}.

2. We compute

$$(B^{-1}A^{-1})AB = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n.$$

   Here we used the associativity of matrix multiplication and the fact that $I_nB = B$. This shows that $B^{-1}A^{-1}$ is the inverse of AB.

□

Why is the inverse of AB not equal to $A^{-1}B^{-1}$? If it were, then we would have

$$I_n = (AB)(A^{-1}B^{-1}) = ABA^{-1}B^{-1}.$$

But there is no reason for $ABA^{-1}B^{-1}$ to equal the identity matrix: one cannot switch the order of A^{-1} and B, so there is nothing to cancel in this expression. In fact, if $I_n = (AB)(A^{-1}B^{-1})$, then we can multiply both sides on the right by BA to conclude that $AB = BA$. In other words, $(AB)^{-1} = A^{-1}B^{-1}$ if and only if $AB = BA$.

More generally, the inverse of a product of several invertible matrices is the product of the inverses, in the opposite order; the proof is the same. For instance,

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}.$$

3.5.2 Computing the Inverse Matrix

So far we have defined the inverse matrix without giving any strategy for computing it. We do so now, beginning with the special case of 2×2 matrices. Then we will give a recipe for the $n \times n$ case.

Definition. The **determinant** of a 2×2 matrix is the number

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Proposition. *Let* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

1. *If* $\det(A) \neq 0$*, then* A *is invertible, and*

$$A^{-1} = \frac{1}{\det(A)}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

2. *If* $\det(A) = 0$*, then* A *is not invertible.*

Proof.

1. Suppose that $\det(A) \neq 0$. Define $B = \dfrac{1}{\det(A)}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Then

$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = I_2.$$

   The reader can check that $BA = I_2$, so A is invertible and $B = A^{-1}$.

2. Suppose that $\det(A) = ad - bc = 0$. Let $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ be the matrix transformation $T(x) = Ax$. Then

$$T\begin{pmatrix} -b \\ a \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} -b \\ a \end{pmatrix} = \begin{pmatrix} -ab+ab \\ -bc+ad \end{pmatrix} = \begin{pmatrix} 0 \\ \det(A) \end{pmatrix} = 0$$
$$T\begin{pmatrix} d \\ -c \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d \\ -c \end{pmatrix} = \begin{pmatrix} ad-bc \\ cd-cd \end{pmatrix} = \begin{pmatrix} \det(A) \\ 0 \end{pmatrix} = 0.$$

   If A is the zero matrix, then it is obviously not invertible. Otherwise, one of $v = \binom{-b}{a}$ and $v = \binom{d}{-c}$ will be a nonzero vector in the null space of A. Suppose that there were a matrix B such that $BA = I_2$. Then

$$v = I_2 v = BAv = B0 = 0,$$

   which is impossible as $v \neq 0$. Therefore, A is not invertible.

□

There is an analogous formula for the inverse of an $n \times n$ matrix, but it is not as simple, and it is computationally intensive. The interested reader can find it in this subsection in Section 4.2.

Example. Let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Then $\det(A) = 1 \cdot 4 - 2 \cdot 3 = -2$. By the proposition, the matrix A is invertible with inverse

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^{-1} = \frac{1}{\det(A)}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}.$$

We check:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot -\frac{1}{2}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = -\frac{1}{2}\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} = I_2.$$

The following theorem gives a procedure for computing A^{-1} in general.

Theorem. *Let A be an $n \times n$ matrix, and let $(\,A \mid I_n\,)$ be the matrix obtained by augmenting A by the identity matrix. If the reduced row echelon form of $(\,A \mid I_n\,)$ has the form $(\,I_n \mid B\,)$, then A is invertible and $B = A^{-1}$. Otherwise, A is not invertible.*

Proof. First suppose that the reduced row echelon form of $(A \mid I_n)$ does not have the form $(I_n \mid B)$. This means that fewer than n pivots are contained in the first n columns (the non-augmented part), so A has fewer than n pivots. It follows that $\mathrm{Nul}(A) \neq \{0\}$ (the equation $Ax = 0$ has a free variable), so there exists a nonzero vector v in $\mathrm{Nul}(A)$. Suppose that there were a matrix B such that $BA = I_n$. Then

$$v = I_n v = BAv = B0 = 0,$$

which is impossible as $v \neq 0$. Therefore, A is not invertible.

Now suppose that the reduced row echelon form of $(A \mid I_n)$ has the form $(I_n \mid B)$. In this case, all pivots are contained in the non-augmented part of the matrix, so the augmented part plays no role in the row reduction: the entries of the augmented part do not influence the choice of row operations used. Hence, row reducing $(A \mid I_n)$ is equivalent to solving the n systems of linear equations $Ax_1 = e_1, Ax_2 = e_2, \ldots, Ax_n = e_n$, where $e_1, e_2, \ldots, e_n$ are the standard coordinate vectors:

$$Ax_1 = e_1: \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -4 & 0 & 0 & 1 \end{array}\right)$$

$$Ax_2 = e_2: \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -4 & 0 & 0 & 1 \end{array}\right)$$

$$Ax_3 = e_3: \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -4 & 0 & 0 & 1 \end{array}\right).$$

The columns $x_1, x_2, \ldots, x_n$ of the matrix B in the row reduced form are the solutions to these equations:

$$A\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = e_1: \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & 3/2 & 1/2 \end{array}\right)$$

$$A\begin{pmatrix} -6 \\ -2 \\ 3/2 \end{pmatrix} = e_2: \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & 3/2 & 1/2 \end{array}\right)$$

$$A\begin{pmatrix} -2 \\ -1 \\ 1/2 \end{pmatrix} = e_3: \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & 3/2 & 1/2 \end{array}\right).$$

By this fact in Section 3.3, the product Be_i is just the ith column x_i of B, so

$$e_i = Ax_i = ABe_i$$

for all i. By the same fact, the ith column of AB is e_i, which means that AB is the identity matrix. Thus B is the inverse of A. □

Example (An invertible matrix)**.** Find the inverse of the matrix

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

Solution. We augment by the identity and row reduce:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -4 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_3 = R_3 + 3R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 3 & 1 \end{array}\right)$$

$$\xrightarrow[R_2 = R_2 - R_3]{R_1 = R_1 - 2R_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 2 & 0 & 3 & 1 \end{array}\right)$$

$$\xrightarrow{R_3 = R_3 \div 2} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -6 & -2 \\ 0 & 1 & 0 & 0 & -2 & -1 \\ 0 & 0 & 1 & 0 & 3/2 & 1/2 \end{array}\right).$$

By the theorem, the inverse matrix is

$$\begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & -3 & -4 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -6 & -2 \\ 0 & -2 & -1 \\ 0 & 3/2 & 1/2 \end{pmatrix}.$$

We check:

$$\begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & -3 & -4 \end{pmatrix}\begin{pmatrix} 1 & -6 & -2 \\ 0 & -2 & -1 \\ 0 & 3/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Example (A non-invertible matrix)**.** Is the following matrix invertible?

$$A = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & -3 & -6 \end{pmatrix}.$$

Solution. We augment by the identity and row reduce:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & -3 & -6 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_3 = R_3 + 3R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 4 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \end{array}\right).$$

At this point we can stop, because it is clear that the reduced row echelon form will not have I_3 in the non-augmented part: it will have a row of zeros. By the theorem, the matrix is not invertible.

3.5.3 Solving Linear Systems using Inverses

In this subsection, we learn to solve $Ax = b$ by "dividing by A."

Theorem. *Let A be an invertible $n \times n$ matrix, and let b be a vector in $\mathbf{R}^n$. Then the matrix equation $Ax = b$ has exactly one solution:*

$$x = A^{-1}b.$$

Proof. We calculate:

$$\begin{aligned} Ax = b \quad &\implies \quad A^{-1}(Ax) = A^{-1}b \\ &\implies \quad (A^{-1}A)x = A^{-1}b \\ &\implies \quad I_n x = A^{-1}b \\ &\implies \quad x = A^{-1}b. \end{aligned}$$

Here we used associativity of matrix multiplication, and the fact that $I_n x = x$ for any vector b. □

Example (Solving a 2×2 system using inverses)**.** Solve the matrix equation

$$\begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix} x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Solution. By the theorem, the only solution of our linear system is

$$x = \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 2 & -3 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

Here we used

$$\det \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix} = 1 \cdot 2 - (-1) \cdot 3 = 5.$$

Example (Solving a 3×3 system using inverses)**.** Solve the system of equations

$$\left\{ \begin{aligned} 2x_1 + 3x_2 + 2x_3 &= 1 \\ x_1 + 3x_3 &= 1 \\ 2x_1 + 2x_2 + 3x_3 &= 1. \end{aligned} \right.$$

Solution. First we write our system as a matrix equation $Ax = b$, where

$$A = \begin{pmatrix} 2 & 3 & 2 \\ 1 & 0 & 3 \\ 2 & 2 & 3 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Next we find the inverse of A by augmenting and row reducing:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right) \xrightarrow{R_1 \longleftrightarrow R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 2 & 3 & 2 & 1 & 0 & 0 \\ 2 & 2 & 3 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[R_3 = R_3 - 2R_1]{R_2 = R_2 - 2R_1} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 3 & -4 & 1 & -2 & 0 \\ 0 & 2 & -3 & 0 & -2 & 1 \end{array}\right)$$

$$\xrightarrow{R_2 = R_2 - R_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 2 & -3 & 0 & -2 & 1 \end{array}\right)$$

$$\xrightarrow{R_3 = R_3 - 2R_2} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & -1 & -2 & -2 & 3 \end{array}\right)$$

$$\xrightarrow{R_3 = -R_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 & 2 & -3 \end{array}\right)$$

$$\xrightarrow[R_2 = R_2 + R_3]{R_1 = R_1 - 3R_3} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -6 & -5 & 9 \\ 0 & 1 & 0 & 3 & 2 & -4 \\ 0 & 0 & 1 & 2 & 2 & -3 \end{array}\right).$$

By the theorem, the only solution of our linear system is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 2 \\ 1 & 0 & 3 \\ 2 & 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -6 & -5 & 9 \\ 3 & 2 & -4 \\ 2 & 2 & -3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}.$$

The advantage of solving a linear system using inverses is that it becomes much faster to solve the matrix equation $Ax = b$ for other, or even unknown, values of b. For instance, in the above example, the solution of the system of equations

$$\begin{cases} 2x_1 + 3x_2 + 2x_3 = b_1 \\ x_1 + 3x_3 = b_2 \\ 2x_1 + 2x_2 + 3x_3 = b_3, \end{cases}$$

where b_1, b_2, b_3 are unknowns, is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 2 \\ 1 & 0 & 3 \\ 2 & 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} -6 & -5 & 9 \\ 3 & 2 & -4 \\ 2 & 2 & -3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} -6b_1 - 5b_2 + 9b_3 \\ 3b_1 + 2b_2 - 4b_3 \\ 2b_1 + 2b_2 - 3b_3 \end{pmatrix}.$$

3.5.4 Invertible linear transformations

As with matrix multiplication, it is helpful to understand matrix inversion as an operation on linear transformations. Recall that the identity transformation on $\mathbf{R}^n$ is denoted $\mathrm{Id}_{\mathbf{R}^n}$.

Definition. A transformation $T: \mathbf{R}^n \to \mathbf{R}^n$ is **invertible** if there exists a transformation $U: \mathbf{R}^n \to \mathbf{R}^n$ such that $T \circ U = \mathrm{Id}_{\mathbf{R}^n}$ and $U \circ T = \mathrm{Id}_{\mathbf{R}^n}$. In this case, the transformation U is called the **inverse** of T, and we write $U = T^{-1}$.

The inverse U of T "undoes" whatever T did. We have

$$T \circ U(x) = x \quad \text{and} \quad U \circ T(x) = x$$

for all vectors x. This means that if you apply T to x, then you apply U, you get the vector x back, and likewise in the other order.

Example (Functions of one variable)**.** Define $f: \mathbf{R} \to \mathbf{R}$ by $f(x) = 2x$. This is an invertible transformation, with inverse $g(x) = x/2$. Indeed,

$$f \circ g(x) = f(g(x)) = f\left(\frac{x}{2}\right) = 2\left(\frac{x}{2}\right) = x$$

and

$$g \circ f(x) = g(f(x)) = g(2x) = \frac{2x}{2} = x.$$

In other words, dividing by 2 undoes the transformation that multiplies by 2.

Define $f: \mathbf{R} \to \mathbf{R}$ by $f(x) = x^3$. This is an invertible transformation, with inverse $g(x) = \sqrt[3]{x}$. Indeed,

$$f \circ g(x) = f(g(x)) = f(\sqrt[3]{x}) = \left(\sqrt[3]{x}\right)^3 = x$$

and

$$g \circ f(x) = g(f(x)) = g(x^3) = \sqrt[3]{x^3} = x.$$

In other words, taking the cube root undoes the transformation that takes a number to its cube.

Define $f: \mathbf{R} \to \mathbf{R}$ by $f(x) = x^2$. This is *not* an invertible function. Indeed, we have $f(2) = 2 = f(-2)$, so there is no way to undo f: the inverse transformation would not know if it should send 2 to 2 or -2. More formally, if $g: \mathbf{R} \to \mathbf{R}$ satisfies $g(f(x)) = x$, then

$$2 = g(f(2)) = g(2) \quad \text{and} \quad -2 = g(f(-2)) = g(2),$$

which is impossible: $g(2)$ is a number, so it cannot be equal to 2 and -2 at the same time.

Define $f: \mathbf{R} \to \mathbf{R}$ by $f(x) = e^x$. This is *not* an invertible function. Indeed, if there were a function $g: \mathbf{R} \to \mathbf{R}$ such that $f \circ g = \mathrm{Id}_{\mathbf{R}}$, then we would have

$$-1 = f \circ g(-1) = f(g(-1)) = e^{g(-1)}.$$

But e^x is a *positive* number for every x, so this is impossible.

Example (Dilation)**.** Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be dilation by a factor of $3/2$: that is, $T(x) = 3/2x$. Is T invertible? If so, what is T^{-1}?

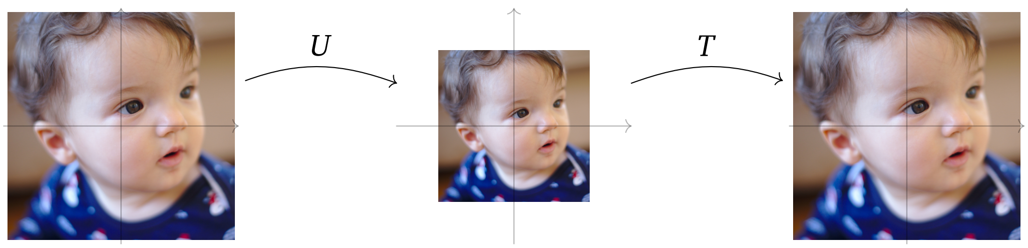
Solution. Let $U: \mathbf{R}^2 \to \mathbf{R}^2$ be dilation by a factor of $2/3$: that is, $U(x) = 2/3x$. Then

$$T \circ U(x) = T\left(\frac{2}{3}x\right) = \frac{3}{2} \cdot \frac{2}{3}x = x$$

and

$$U \circ T(x) = U\left(\frac{3}{2}x\right) = \frac{2}{3} \cdot \frac{3}{2}x = x.$$

Hence $T \circ U = \mathrm{Id}_{\mathbf{R}^2}$ and $U \circ T = \mathrm{Id}_{\mathbf{R}^2}$, so T is invertible, with inverse U. In other words, *shrinking* by a factor of $2/3$ undoes *stretching* by a factor of $3/2$.



Use this link to view the online demo

Shrinking by a factor of $2/3$ *followed by scaling by a factor of* $3/2$ *is the identity transformation.*

Use this link to view the online demo

Scaling by a factor of $3/2$ *followed by shrinking by a factor of* $2/3$ *is the identity transformation.*

Example (Rotation)**.** Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be counterclockwise rotation by $45°$. Is T invertible? If so, what is T^{-1}?

Solution. Let $U: \mathbf{R}^2 \to \mathbf{R}^2$ be *clockwise* rotation by $45°$. Then $T \circ U$ first rotates clockwise by $45°$, then counterclockwise by $45°$, so the composition rotates by zero degrees: it is the identity transformation. Likewise, $U \circ T$ first rotates counterclockwise, then clockwise by the same amount, so it is the identity transformation. In other words, *clockwise* rotation by $45°$ undoes *counterclockwise* rotation by $45°$.

Use this link to view the online demo

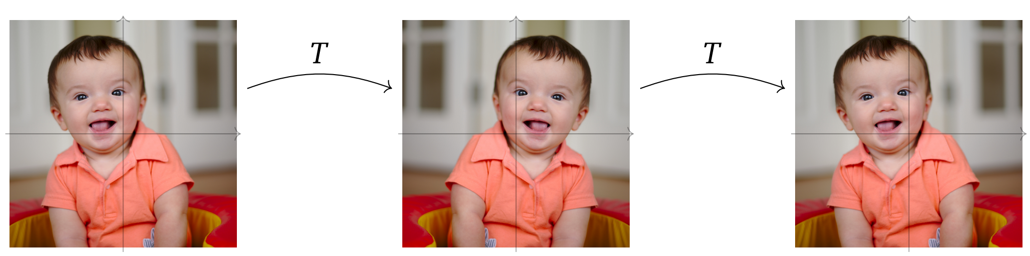
Counterclockwise rotation by 45° followed by clockwise rotation by 45° is the identity transformation.

Use this link to view the online demo

Clockwise rotation by 45° followed by counterclockwise rotation by 45° is the identity transformation.

Example (Reflection). Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be the reflection over the y-axis. Is T invertible? If so, what is T^{-1}?

Solution. The transformation T is invertible; in fact, it is equal to its own inverse. Reflecting a vector x over the y-axis twice brings the vector back to where it started, so $T \circ T = \mathrm{Id}_{\mathbf{R}^2}$.



Use this link to view the online demo

The transformation T is equal to its own inverse: applying T twice takes a vector back to where it started.

Non-Example (Projection). Let $T: \mathbf{R}^3 \to \mathbf{R}^3$ be the projection onto the xy-plane, introduced in this example in Section 3.1. Is T invertible?

Solution. The transformation T is *not* invertible. Every vector on the z-axis projects onto the zero vector, so there is no way to undo what T did: the inverse transformation would not know which vector on the z-axis it should send the zero vector to. More formally, suppose there were a transformation $U\colon \mathbf{R}^3 \to \mathbf{R}^3$ such that $U \circ T = \mathrm{Id}_{\mathbf{R}^3}$. Then

$$0 = U \circ T(0) = U(T(0)) = U(0)$$

and

$$\begin{pmatrix}0\\0\\1\end{pmatrix} = U \circ T\begin{pmatrix}0\\0\\1\end{pmatrix} = U\left(T\begin{pmatrix}0\\0\\1\end{pmatrix}\right) = U(0).$$

But $U(0)$ is as single vector in $\mathbf{R}^3$, so it cannot be equal to 0 and to $(0,0,1)$ at the same time.

Use this link to view the online demo

Projection onto the xy-plane is not an invertible transformation: all points on each vertical line are sent to the same point by T, so there is no way to undo T.

Proposition.

1. *A transformation $T: \mathbf{R}^n \to \mathbf{R}^n$ is invertible if and only if it is both one-to-one and onto.*

2. *If T is already known to be invertible, then $U\colon \mathbf{R}^n \to \mathbf{R}^n$ is the inverse of T provided that* either *$T \circ U = \mathrm{Id}_{\mathbf{R}^n}$ or $U \circ T = \mathrm{Id}_{\mathbf{R}^n}$: it is only necessary to verify one.*

Proof. To say that T is one-to-one and onto means that $T(x) = b$ has *exactly one* solution for every b in $\mathbf{R}^n$.

Suppose that T is invertible. Then $T(x) = b$ always has the unique solution $x = T^{-1}(b)$: indeed, applying T^{-1} to both sides of $T(x) = b$ gives

$$x = T^{-1}(T(x)) = T^{-1}(b),$$

and applying T to both sides of $x = T^{-1}(b)$ gives

$$T(x) = T(T^{-1}(b)) = b.$$

Conversely, suppose that T is one-to-one and onto. Let b be a vector in $\mathbf{R}^n$, and let $x = U(b)$ be the unique solution of $T(x) = b$. Then U defines a transformation from $\mathbf{R}^n$ to $\mathbf{R}^n$. For any x in $\mathbf{R}^n$, we have $U(T(x)) = x$, because x is the unique solution of the equation $T(x) = b$ for $b = T(x)$. For any b in $\mathbf{R}^n$, we have $T(U(b)) = b$, because $x = U(b)$ is the unique solution of $T(x) = b$. Therefore, U is the inverse of T, and T is invertible.

Suppose now that T is an invertible transformation, and that U is another transformation such that $T \circ U = \mathrm{Id}_{\mathbf{R}^n}$. We must show that $U = T^{-1}$, i.e., that

$U \circ T = \mathrm{Id}_{\mathbf{R}^n}$. We compose both sides of the equality $T \circ U = \mathrm{Id}_{\mathbf{R}^n}$ on the left by T^{-1} and on the right by T to obtain

$$T^{-1} \circ T \circ U \circ T = T^{-1} \circ \mathrm{Id}_{\mathbf{R}^n} \circ T.$$

We have $T^{-1} \circ T = \mathrm{Id}_{\mathbf{R}^n}$ and $\mathrm{Id}_{\mathbf{R}^n} \circ U = U$, so the left side of the above equation is $U \circ T$. Likewise, $\mathrm{Id}_{\mathbf{R}^n} \circ T = T$ and $T^{-1} \circ T = \mathrm{Id}_{\mathbf{R}^n}$, so our equality simplifies to $U \circ T = \mathrm{Id}_{\mathbf{R}^n}$, as desired.

If instead we had assumed only that $U \circ T = \mathrm{Id}_{\mathbf{R}^n}$, then the proof that $T \circ U = \mathrm{Id}_{\mathbf{R}^n}$ proceeds similarly. □

Remark. It makes sense in the above definition to define the inverse of a transformation $T: \mathbf{R}^n \to \mathbf{R}^m$, for $m \neq n$, to be a transformation $U: \mathbf{R}^m \to \mathbf{R}^n$ such that $T \circ U = \mathrm{Id}_{\mathbf{R}^m}$ and $U \circ T = \mathrm{Id}_{\mathbf{R}^n}$. In fact, there exist invertible transformations $T: \mathbf{R}^n \to \mathbf{R}^m$ for any m and n, but they are not linear, or even continuous.

If T is a *linear* transformation, then it can only be invertible when $m = n$, i.e., when its domain is equal to its codomain. Indeed, if $T: \mathbf{R}^n \to \mathbf{R}^m$ is one-to-one, then $n \leq m$ by this note in Section 3.2, and if T is onto, then $m \leq n$ by this note in Section 3.2. Therefore, when discussing invertibility we restrict ourselves to the case $m = n$.

Challenge. Find an invertible (non-linear) transformation $T: \mathbf{R}^2 \to \mathbf{R}$.

As you might expect, the matrix for the inverse of a linear transformation is the inverse of the matrix for the transformation, as the following theorem asserts.

Theorem. *Let $T: \mathbf{R}^n \to \mathbf{R}^n$ be a linear transformation with standard matrix A. Then T is invertible if and only if A is invertible, in which case T^{-1} is linear with standard matrix A^{-1}.*

Proof. Suppose that T is invertible. Let $U: \mathbf{R}^n \to \mathbf{R}^n$ be the inverse of T. We claim that U is linear. We need to check the defining properties in Section 3.3. Let u, v be vectors in $\mathbf{R}^n$. Then

$$u + v = T(U(u)) + T(U(v)) = T(U(u) + U(v))$$

by linearity of T. Applying U to both sides gives

$$U(u + v) = U\big(T(U(u) + U(v))\big) = U(u) + U(v).$$

Let c be a scalar. Then

$$cu = cT(U(u)) = T(cU(u))$$

by linearity of T. Applying U to both sides gives

$$U(cu) = U\big(T(cU(u))\big) = cU(u).$$

Since U satisfies the defining properties in Section 3.3, it is a linear transformation.

Now that we know that U is linear, we know that it has a standard matrix B. By the compatibility of matrix multiplication and composition in Section 3.4, the matrix for $T \circ U$ is AB. But $T \circ U$ is the identity transformation $\mathrm{Id}_{\mathbf{R}^n}$, and the standard matrix for $\mathrm{Id}_{\mathbf{R}^n}$ is I_n, so $AB = I_n$. One shows similarly that $BA = I_n$. Hence A is invertible and $B = A^{-1}$.

Conversely, suppose that A is invertible. Let $B = A^{-1}$, and define $U : \mathbf{R}^n \to \mathbf{R}^n$ by $U(x) = Bx$. By the compatibility of matrix multiplication and composition in Section 3.4, the matrix for $T \circ U$ is $AB = I_n$, and the matrix for $U \circ T$ is $BA = I_n$. Therefore,

$$T \circ U(x) = ABx = I_n x = x \quad \text{and} \quad U \circ T(x) = BAx = I_n x = x,$$

which shows that T is invertible with inverse transformation U. □

Example (Dilation)**.** Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be dilation by a factor of $3/2$: that is, $T(x) = 3/2x$. Is T invertible? If so, what is T^{-1}?

Solution. In this example in Section 3.1 we showed that the matrix for T is

$$A = \begin{pmatrix} 3/2 & 0 \\ 0 & 3/2 \end{pmatrix}.$$

The determinant of A is $9/4 \neq 0$, so A is invertible with inverse

$$A^{-1} = \frac{1}{9/4}\begin{pmatrix} 3/2 & 0 \\ 0 & 3/2 \end{pmatrix} = \begin{pmatrix} 2/3 & 0 \\ 0 & 2/3 \end{pmatrix}.$$

By the theorem, T is invertible, and its inverse is the matrix transformation for A^{-1}:

$$T^{-1}(x) = \begin{pmatrix} 2/3 & 0 \\ 0 & 2/3 \end{pmatrix} x.$$

We recognize this as a dilation by a factor of $2/3$.

Example (Rotation)**.** Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be counterclockwise rotation by $45°$. Is T invertible? If so, what is T^{-1}?

Solution. In this example in Section 3.3, we showed that the standard matrix for the counterclockwise rotation of the plane by an angle of θ is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Therefore, the standard matrix A for T is

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

where we have used the trigonometric identities

$$\cos(45°) = \frac{1}{\sqrt{2}} \qquad \sin(45°) = \frac{1}{\sqrt{2}}.$$

The determinant of A is

$$\det(A) = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} \frac{-1}{\sqrt{2}} = \frac{1}{2} + \frac{1}{2} = 1,$$

so the inverse is

$$A^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

By the theorem, T is invertible, and its inverse is the matrix transformation for A^{-1}:

$$T^{-1}(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} x.$$

We recognize this as a clockwise rotation by 45°, using the trigonometric identities

$$\cos(-45°) = \frac{1}{\sqrt{2}} \qquad \sin(-45°) = -\frac{1}{\sqrt{2}}.$$

Example (Reflection)**.** Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be the reflection over the y-axis. Is T invertible? If so, what is T^{-1}?

Solution. In this example in Section 3.1 we showed that the matrix for T is

$$A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

This matrix has determinant -1, so it is invertible, with inverse

$$A^{-1} = -\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = A.$$

By the theorem, T is invertible, and it is equal to its own inverse: $T^{-1} = T$. This is another way of saying that a reflection "undoes" itself.

3.6 The Invertible Matrix Theorem

Objectives

1. *Theorem:* the invertible matrix theorem.

This section consists of a single important theorem containing many equivalent conditions for a matrix to be invertible. This is one of the most important theorems in this textbook. We will append two more criteria in Section 5.1.

Invertible Matrix Theorem. *Let A be an $n \times n$ matrix, and let $T : \mathbf{R}^n \to \mathbf{R}^n$ be the matrix transformation $T(x) = Ax$. The following statements are equivalent:*

1. *A is invertible.*

2. *A has n pivots.*

3. $\mathrm{Nul}(A) = \{0\}$.

4. *The columns of A are linearly independent.*

5. *The columns of A span* $\mathbf{R}^n$.

6. $Ax = b$ *has a unique solution for each* b *in* $\mathbf{R}^n$.

7. T *is invertible.*

8. T *is one-to-one.*

9. T *is onto.*

Proof. ($1 \iff 2$): The matrix A has n pivots if and only if its reduced row echelon form is the identity matrix I_n. This happens exactly when the procedure in Section 3.5 to compute the inverse succeeds.

($2 \iff 3$): The null space of a matrix is $\{0\}$ if and only if the matrix has no free variables, which means that every column is a pivot column, which means A has n pivots. See this recipe in Section 2.6.

($2 \iff 4$, $2 \iff 5$): These follow from this recipe in Section 2.5 and this theorem in Section 2.3, respectively, since A has n pivots if and only if has a pivot in every row/column.

($4{+}5 \iff 6$): We know $Ax = b$ has at least one solution for every b if and only if the columns of A span $\mathbf{R}^n$ by this theorem in Section 3.2, and $Ax = b$ has at most one solution for every b if and only if the columns of A are linearly independent by this theorem in Section 3.2. Hence $Ax = b$ has exactly one solution for every b if and only if its columns are linearly independent and span $\mathbf{R}^n$.

($1 \iff 7$): This is the content of this theorem in Section 3.5.

($7 \implies 8 + 9$): See this proposition in Section 3.5.

($8 \iff 4$, $9 \iff 5$): See this this theorem in Section 3.2 and this theorem in Section 3.2. □

To reiterate, the invertible matrix theorem means:

There are two kinds of *square* matrices:

1. invertible matrices, and

2. non-invertible matrices.

For invertible matrices, all of the statements of the invertible matrix theorem are true.

For non-invertible matrices, all of the statements of the invertible matrix theorem are false.

The reader should be comfortable translating any of the statements in the invertible matrix theorem into a statement about the pivots of a matrix.

Other Conditions for Invertibility. The following conditions are also equivalent to the invertibility of a square matrix A. They are all simple restatements of conditions in the invertible matrix theorem.

1. The reduced row echelon form of A is the identity matrix I_n.
2. $Ax = 0$ has no solutions other than the trivial one.
3. $\text{nullity}(A) = 0$.
4. The columns of A form a basis for $\mathbf{R}^n$.
5. $Ax = b$ is consistent for all b in $\mathbf{R}^n$.
6. $\text{Col}(A) = \mathbf{R}^n$.
7. $\dim \text{Col}(A) = n$.
8. $\text{rank}(A) = n$.

Now we can show that to check $B = A^{-1}$, it's enough to show $AB = I_n$ *or* $BA = I_n$.

Corollary (A Left or Right Inverse Suffices). *Let A be an $n \times n$ matrix, and suppose that there exists an $n \times n$ matrix B such that $AB = I_n$* or *$BA = I_n$. Then A is invertible and $B = A^{-1}$.*

Proof. Suppose that $AB = I_n$. We claim that $T(x) = Ax$ is onto. Indeed, for any b in $\mathbf{R}^n$, we have

$$b = I_n b = (AB)b = A(Bb),$$

so $T(Bb) = b$, and hence b is in the range of T. Therefore, A is invertible by the invertible matrix theorem. Since A is invertible, we have

$$A^{-1} = A^{-1}I_n = A^{-1}(AB) = (A^{-1}A)B = I_n B = B,$$

so $B = A^{-1}$.

Now suppose that $BA = I_n$. We claim that $T(x) = Ax$ is one-to-one. Indeed, suppose that $T(x) = T(y)$. Then $Ax = Ay$, so $BAx = BAy$. But $BA = I_n$, so $I_n x = I_n y$, and hence $x = y$. Therefore, A is invertible by the invertible matrix theorem. One shows that $B = A^{-1}$ as above. □

We conclude with some common situations in which the invertible matrix theorem is useful.

Example. Is this matrix invertible?

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & 7 \\ -2 & -4 & 1 \end{pmatrix}$$

Solution. The second column is a multiple of the first. The columns are linearly dependent, so A does not satisfy condition 4 of the invertible matrix theorem. Therefore, A is not invertible.

Example. Let A be an $n \times n$ matrix and let $T(x) = Ax$. Suppose that the range of T is $\mathbf{R}^n$. Show that the columns of A are linearly independent.

Solution. The range of T is the column space of A, so A satisfies condition 5 of the invertible matrix theorem. Therefore, A also satisfies condition 4, which says that the columns of A are linearly independent.

Example. Let A be a 3×3 matrix such that

$$A\begin{pmatrix} 1 \\ 7 \\ 0 \end{pmatrix} = A\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}.$$

Show that the rank of A is at most 2.

Solution. If we set

$$b = A\begin{pmatrix} 1 \\ 7 \\ 0 \end{pmatrix} = A\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix},$$

then $Ax = b$ has multiple solutions, so it does not satisfy condition 6 of the invertible matrix theorem. Therefore, it does not satisfy condition 5, so the columns of A do not span $\mathbf{R}^3$. Therefore, the column space has dimension strictly less than 3, the rank is at most 2.

Example. Suppose that A is an $n \times n$ matrix such that $Ax = b$ is inconsistent some vector b. Show that $Ax = b$ has infinitely many solutions for some (other) vector b.

Solution. By hypothesis, A does not satisfy condition 6 of the invertible matrix theorem. Therefore, it does not satisfy condition 3, so $\mathrm{Nul}(A)$ is an infinite set. If we take $b = 0$, then the equation $Ax = b$ has infinitely many solutions.

Chapter 4

Determinants

We begin by recalling the overall structure of this book:

1. Solve the matrix equation $Ax = b$.

2. Solve the matrix equation $Ax = \lambda x$, where λ is a number.

3. Approximately solve the matrix equation $Ax = b$.

At this point we have said all that we will say about the first part. This chapter belongs to the second.

Primary Goal. Learn about determinants: their computation and their properties.

The *determinant* of a square matrix A is a number $\det(A)$. This incredible quantity is one of the most important invariants of a matrix; as such, it forms the basis of most advanced computations involving matrices.

In Section 4.1, we will define the determinant in terms of its behavior with respect to row operations. The determinant satisfies many wonderful properties: for instance, $\det(A) \neq 0$ if and only if A is invertible. We will discuss some of these properties in Section 4.1 as well. In Section 4.2, we will give a recursive formula for the determinant of a matrix. This formula is very useful, for instance, when taking the determinant of a matrix with unknown entries; this will be important in Chapter 5. Finally, in Section 4.3, we will relate determinants to volumes. This gives a geometric interpretation for determinants, and explains why the determinant is defined the way it is. This interpretation of determinants is a crucial ingredient in the change-of-variables formula in multivariable calculus.

4.1 Determinants: Definition

Objectives

1. Learn the definition of the determinant.

2. Learn some ways to eyeball a matrix with zero determinant, and how to compute determinants of upper- and lower-triangular matrices.
3. Learn the basic properties of the determinant, and how to apply them.
4. *Recipe:* compute the determinant using row and column operations.
5. *Theorems:* existence theorem, invertibility property, multiplicativity property, transpose property.
6. *Vocabulary words:* **diagonal**, **upper-triangular**, **lower-triangular**, **transpose**.
7. *Essential vocabulary word:* **determinant**.

In this section, we define the determinant, and we present one way to compute it. Then we discuss some of the many wonderful properties the determinant enjoys.

4.1.1 The Definition of the Determinant

The determinant of a square matrix A is a real number $\det(A)$. It is defined via its behavior with respect to row operations; this means we can use row reduction to compute it. We will give a recursive formula for the determinant in Section 4.2. We will also show in this subsection that the determinant is related to invertibility, and in Section 4.3 that it is related to volumes.

Essential Definition. The **determinant** is a function

$$\det\colon \{\text{square matrices}\} \longrightarrow \mathbf{R}$$

satisfying the following properties:

1. Doing a row replacement on A does not change $\det(A)$.
2. Scaling a row of A by a scalar c multiplies the determinant by c.
3. Swapping two rows of a matrix multiplies the determinant by -1.
4. The determinant of the identity matrix I_n is equal to 1.

In other words, to every square matrix A we assign a number $\det(A)$ in a way that satisfies the above properties.

In each of the first three cases, doing a row operation on a matrix scales the determinant by a *nonzero* number. (Multiplying a row by zero is not a row operation.) Therefore, doing row operations on a square matrix A does not change whether or not the determinant is zero.

The main motivation behind using these particular defining properties is geometric: see Section 4.3. Another motivation for this definition is that it tells us how to compute the determinant: we row reduce and keep track of the changes.

Example. Let us compute $\det\begin{pmatrix}2&1\\1&4\end{pmatrix}$. First we row reduce, then we compute the determinant in the opposite order:

$$\begin{aligned}
&\begin{pmatrix}2&1\\1&4\end{pmatrix} && \det = 7\\
\xrightarrow{R_1 \longleftrightarrow R_2}\ &\begin{pmatrix}1&4\\2&1\end{pmatrix} && \det = -7\\
\xrightarrow{R_2 = R_2 - 2R_1}\ &\begin{pmatrix}1&4\\0&-7\end{pmatrix} && \det = -7\\
\xrightarrow{R_2 = R_2 \div -7}\ &\begin{pmatrix}1&4\\0&1\end{pmatrix} && \det = 1\\
\xrightarrow{R_1 = R_1 - 4R_2}\ &\begin{pmatrix}1&0\\0&1\end{pmatrix} && \det = 1
\end{aligned}$$

The reduced row echelon form of the matrix is the identity matrix I_2, so its determinant is 1. The second-last step in the row reduction was a row replacement, so the second-final matrix also has determinant 1. The previous step in the row reduction was a row scaling by $-1/7$; since (the determinant of the second matrix times $-1/7$) is 1, the determinant of the second matrix must be -7. The first step in the row reduction was a row swap, so the determinant of the first matrix is negative the determinant of the second. Thus, the determinant of the original matrix is 7.

Note that our answer agrees with this definition of the determinant.

Example. Compute $\det\begin{pmatrix}1&0\\0&3\end{pmatrix}$.

Solution. Let $A = \begin{pmatrix}1&0\\0&3\end{pmatrix}$. Since A is obtained from I_2 by multiplying the second row by the constant 3, we have

$$\det(A) = 3\det(I_2) = 3 \cdot 1 = 3.$$

Note that our answer agrees with this definition of the determinant.

Example. Compute $\det\begin{pmatrix}1&0&0\\0&0&1\\5&1&0\end{pmatrix}$.

Solution. First we row reduce, then we compute the determinant in the opposite order:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 5 & 1 & 0 \end{pmatrix} \qquad \det = -1$$

$$\xrightarrow{R_2 \longleftrightarrow R_3} \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \det = 1$$

$$\xrightarrow{R_2 = R_2 - 5R_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \det = 1$$

The reduced row echelon form is I_3, which has determinant 1. Working backwards from I_3 and using the four defining properties, we see that the second matrix also has determinant 1 (it differs from I_3 by a row replacement), and the first matrix has determinant -1 (it differs from the second by a row swap).

Here is the general method for computing determinants using row reduction.

Recipe: Computing determinants by row reducing. Let A be a square matrix. Suppose that you do some number of row operations on A to obtain a matrix B in row echelon form. Then

$$\det(A) = (-1)^r \cdot \frac{(\text{product of the diagonal entries of } B)}{(\text{product of scaling factors used})},$$

where r is the number of row swaps performed.

In other words, the determinant of A is the product of diagonal entries of the row echelon form B, times a factor of ± 1 coming from the number of row swaps you made, divided by the product of the scaling factors used in the row reduction.

Remark. This is an efficient way of computing the determinant of a large matrix, either by hand or by computer. The computational complexity of row reduction is $O(n^3)$; by contrast, the cofactor expansion algorithm we will learn in Section 4.2 has complexity $O(n!) \approx O(n^n \sqrt{n})$, which is much larger. (Cofactor expansion has other uses.)

Example. Compute $\det \begin{pmatrix} 0 & -7 & -4 \\ 2 & 4 & 6 \\ 3 & 7 & -1 \end{pmatrix}$.

Solution. We row reduce the matrix, keeping track of the number of row swaps and of the scaling factors used.

$$\begin{pmatrix} 0 & -7 & -4 \\ 2 & 4 & 6 \\ 3 & 7 & -1 \end{pmatrix} \xrightarrow{R_1 \longleftrightarrow R_2} \begin{pmatrix} 2 & 4 & 6 \\ 0 & -7 & -4 \\ 3 & 7 & -1 \end{pmatrix} \qquad r = 1$$

$$\xrightarrow{R_1 = R_1 \div 2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -7 & -4 \\ 3 & 7 & -1 \end{pmatrix} \qquad \text{scaling factors} = \tfrac{1}{2}$$

$$\xrightarrow{R_3 = R_3 - 3R_1} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -7 & -4 \\ 0 & 1 & -10 \end{pmatrix}$$

$$\xrightarrow{R_2 \longleftrightarrow R_3} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -10 \\ 0 & -7 & -4 \end{pmatrix} \qquad r = 2$$

$$\xrightarrow{R_3 = R_3 + 7R_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & -10 \\ 0 & 0 & -74 \end{pmatrix}$$

We made two row swaps and scaled once by a factor of $1/2$, so the recipe says that

$$\det \begin{pmatrix} 0 & -7 & -4 \\ 2 & 4 & 6 \\ 3 & 7 & -1 \end{pmatrix} = (-1)^2 \cdot \frac{1 \cdot 1 \cdot (-74)}{1/2} = -148.$$

Example. Compute $\det \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & 1 \end{pmatrix}$.

Solution. We row reduce the matrix, keeping track of the number of row swaps and of the scaling factors used.

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & 1 \end{pmatrix} \xrightarrow[R_3 = R_3 - 3R_1]{R_2 = R_2 - 2R_1} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -5 & -5 \\ 0 & -6 & -8 \end{pmatrix}$$

$$\xrightarrow{R_2 = R_2 \div -5} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & -6 & -8 \end{pmatrix} \qquad \text{scaling factors} = -\tfrac{1}{5}$$

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

We did not make any row swaps, and we scaled once by a factor of $-1/5$, so the recipe says that

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & 1 \end{pmatrix} = \frac{1 \cdot 1 \cdot (-2)}{-1/5} = 10.$$

Example (The determinant of a 2×2 matrix)**.** Let us use the recipe to compute the determinant of a general 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

- If $a = 0$, then

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det\begin{pmatrix} 0 & b \\ c & d \end{pmatrix} = -\det\begin{pmatrix} c & d \\ 0 & b \end{pmatrix} = -bc.$$

- If $a \neq 0$, then

$$\begin{aligned} \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} &= a \cdot \det\begin{pmatrix} 1 & b/a \\ c & d \end{pmatrix} = a \cdot \det\begin{pmatrix} 1 & b/a \\ 0 & d - c \cdot b/a \end{pmatrix} \\ &= a \cdot 1 \cdot (d - bc/a) = ad - bc. \end{aligned}$$

In either case, we recover the formula in Section 3.5:

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

If a matrix is already in row echelon form, then you can simply read off the determinant as the product of the diagonal entries. It turns out this is true for a slightly larger class of matrices called *triangular*.

Definition.

- The **diagonal** entries of a matrix A are the entries $a_{11}, a_{22}, \ldots$:

diagonal entries

$$\begin{pmatrix} \boxed{a_{11}} & a_{12} & a_{13} & a_{14} \\ a_{21} & \boxed{a_{22}} & a_{23} & a_{24} \\ a_{31} & a_{32} & \boxed{a_{33}} & a_{34} \end{pmatrix} \qquad \begin{pmatrix} \boxed{a_{11}} & a_{12} & a_{13} \\ a_{21} & \boxed{a_{22}} & a_{23} \\ a_{31} & a_{32} & \boxed{a_{33}} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}$$

- A square matrix is called **upper-triangular** if its nonzero entries all lie above the diagonal, and it is called **lower-triangular** if its nonzero entries all lie below the diagonal. It is called **diagonal** if all of its nonzero entries lie on the diagonal, i.e., if it is both upper-triangular and lower-triangular.

upper-triangular $\begin{pmatrix} \star & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & 0 & \star & \star \\ 0 & 0 & 0 & \star \end{pmatrix}$ lower-triangular $\begin{pmatrix} \star & 0 & 0 & 0 \\ \star & \star & 0 & 0 \\ \star & \star & \star & 0 \\ \star & \star & \star & \star \end{pmatrix}$ diagonal $\begin{pmatrix} \star & 0 & 0 & 0 \\ 0 & \star & 0 & 0 \\ 0 & 0 & \star & 0 \\ 0 & 0 & 0 & \star \end{pmatrix}$

Proposition. *Let A be an $n \times n$ matrix.*

1. *If A has a zero row or column, then* $\det(A) = 0$.
2. *If A is upper-triangular or lower-triangular, then* $\det(A)$ *is the product of its diagonal entries.*

Proof.

1. Suppose that A has a zero row. Let B be the matrix obtained by negating the zero row. Then $\det(A) = -\det(B)$ by the second defining property. But $A = B$, so $\det(A) = \det(B)$:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \end{pmatrix} \xrightarrow{R_2 = -R_2} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \end{pmatrix}.$$

   Putting these together yields $\det(A) = -\det(A)$, so $\det(A) = 0$.

   Now suppose that A has a zero column. Then A is not invertible by the invertible matrix theorem in Section 3.6, so its reduced row echelon form has a zero row. Since row operations do not change whether the determinant is zero, we conclude $\det(A) = 0$.

2. First suppose that A is upper-triangular, and that one of the diagonal entries is zero, say $a_{ii} = 0$. We can perform row operations to clear the entries above the nonzero diagonal entries:

$$\begin{pmatrix} a_{11} & \star & \star & \star \\ 0 & a_{22} & \star & \star \\ 0 & 0 & 0 & \star \\ 0 & 0 & 0 & a_{44} \end{pmatrix} \longrightarrow \begin{pmatrix} a_{11} & 0 & \star & 0 \\ 0 & a_{22} & \star & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_{44} \end{pmatrix}$$

   In the resulting matrix, the ith row is zero, so $\det(A) = 0$ by the first part.

   Still assuming that A is upper-triangular, now suppose that all of the diagonal entries of A are nonzero. Then A can be transformed to the identity matrix by scaling the diagonal entries and then doing row replacements:

$$\begin{array}{ccccc} \begin{pmatrix} a & \star & \star \\ 0 & b & \star \\ 0 & 0 & c \end{pmatrix} & \xrightarrow[a^{-1}, b^{-1}, c^{-1}]{\text{scale by}} & \begin{pmatrix} 1 & \star & \star \\ 0 & 1 & \star \\ 0 & 0 & 1 \end{pmatrix} & \xrightarrow{\text{row replacements}} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ \det = abc & \longleftarrow & \det = 1 & \longleftarrow & \det = 1 \end{array}$$

   Since $\det(I_n) = 1$ and we scaled by the reciprocals of the diagonal entries, this implies $\det(A)$ is the product of the diagonal entries.

   The same argument works for lower triangular matrices, except that the the row replacements go down instead of up.

□

Example. Compute the determinants of these matrices:

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \qquad \begin{pmatrix} -20 & 0 & 0 \\ \pi & 0 & 0 \\ 100 & 3 & -7 \end{pmatrix} \qquad \begin{pmatrix} 17 & -3 & 4 \\ 0 & 0 & 0 \\ 11/2 & 1 & e \end{pmatrix}.$$

Solution. The first matrix is upper-triangular, the second is lower-triangular, and the third has a zero row:

$$\det\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} = 1\cdot 4\cdot 6 = 24$$

$$\det\begin{pmatrix} -20 & 0 & 0 \\ \pi & 0 & 0 \\ 100 & 3 & -7 \end{pmatrix} = -20\cdot 0\cdot -7 = 0$$

$$\det\begin{pmatrix} 17 & -3 & 4 \\ 0 & 0 & 0 \\ 11/2 & 1 & e \end{pmatrix} = 0.$$

A matrix can always be transformed into row echelon form by a series of row operations, and a matrix in row echelon form is upper-triangular. Therefore, we have completely justified the recipe for computing the determinant.

The determinant is characterized by its defining properties, since we can compute the determinant of any matrix using row reduction, as in the above recipe. However, we have not yet proved the existence of a function satisfying the defining properties! Row reducing will compute the determinant *if it exists*, but we cannot use row reduction to prove existence, because we do not yet know that you compute the same number by row reducing in two different ways.

Theorem (Existence of the determinant)**.** *There exists one and only one function from the set of square matrices to the real numbers, that satisfies the four defining properties.*

We will prove the existence theorem in Section 4.2, by exhibiting a recursive formula for the determinant. Again, the real content of the existence theorem is:

> No matter which row operations you do, you will always compute the same value for the determinant.

4.1.2 Magical Properties of the Determinant

In this subsection, we will discuss a number of the amazing properties enjoyed by the determinant: the invertibility property, the multiplicativity property, and the transpose property.

Invertibility Property. *A square matrix is invertible if and only if* $\det(A) \neq 0$.

Proof. If A is invertible, then it has a pivot in every row and column by the invertible matrix theorem in Section 3.6, so its reduced row echelon form is the identity matrix. Since row operations do not change whether the determinant is zero, and since $\det(I_n) = 1$, this implies $\det(A) \neq 0$. Conversely, if A is not invertible, then it is row equivalent to a matrix with a zero row. Again, row operations do not change whether the determinant is nonzero, so in this case $\det(A) = 0$. □

By the invertibility property, a matrix that does not satisfy any of the properties of the invertible matrix theorem in Section 3.6 has zero determinant.

Corollary. *Let A be a square matrix. If the rows or columns of A are linearly dependent, then* $\det(A) = 0$.

Proof. If the columns of A are linearly dependent, then A is not invertible by condition 4 of the invertible matrix theorem in Section 3.6. Suppose now that the rows of A are linearly dependent. If $r_1, r_2, \ldots, r_n$ are the rows of A, then one of the rows is in the span of the others, so we have an equation like

$$r_2 = 3r_1 - r_3 + 2r_4.$$

If we perform the following row operations on A:

$$R_2 = R_2 - 3R_1; \quad R_2 = R_2 + R_3; \quad R_2 = R_2 - 2R_4$$

then the second row of the resulting matrix is zero. Hence A is not invertible in this case either.

Alternatively, if the rows of A are linearly dependent, then one can combine condition 4 of the invertible matrix theorem in Section 3.6 and the transpose property below to conclude that $\det(A) = 0$. □

In particular, if two rows/columns of A are multiples of each other, then $\det(A) = 0$. We also recover the fact that a matrix with a row or column of zeros has determinant zero.

Example. The following matrices all have zero determinant:

$$\begin{pmatrix} 0 & 2 & -1 \\ 0 & 5 & 10 \\ 0 & -7 & 3 \end{pmatrix}, \quad \begin{pmatrix} 5 & -15 & 11 \\ 3 & -9 & 2 \\ 2 & -6 & 16 \end{pmatrix}, \quad \begin{pmatrix} 3 & 1 & 2 & 4 \\ 0 & 0 & 0 & 0 \\ 4 & 2 & 5 & 12 \\ -1 & 3 & 4 & 8 \end{pmatrix}, \quad \begin{pmatrix} \pi & e & 11 \\ 3\pi & 3e & 33 \\ 12 & -7 & 2 \end{pmatrix}.$$

The proofs of the multiplicativity property and the transpose property below, as well as the cofactor expansion theorem in Section 4.2 and the determinants and volumes theorem in Section 4.3, use the following strategy: define another function $d: \{n \times n \text{ matrices}\} \to \mathbf{R}$, and prove that d satisfies the same four defining properties as the determinant. By the existence theorem, *the function d is equal to the determinant*. This is an advantage of defining a function via its properties: in order to prove it is equal to another function, one only has to check the defining properties.

Multiplicativity Property. *If A and B are $n \times n$ matrices, then*

$$\det(AB) = \det(A)\det(B).$$

Proof. In this proof, we need to use the notion of an **elementary matrix**. This is a matrix obtained by doing one row operation to the identity matrix. There are three kinds of elementary matrices: those arising from row replacement, row scaling, and row swaps:

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

The important property of elementary matrices is the following claim.

Claim: If E is the elementary matrix for a row operation, then EA is the matrix obtained by performing the same row operation on A.

In other words, left-multiplication by an elementary matrix applies a row operation. For example,

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} - 2a_{11} & a_{22} - 2a_{12} & a_{23} - 2a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$
$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 3a_{11} & 3a_{12} & 3a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$
$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

The proof of the Claim is by direct calculation; we leave it to the reader to generalize the above equalities to $n \times n$ matrices.

As a consequence of the Claim and the four defining properties, we have the following observation. Let C be any square matrix.

1. If E is the elementary matrix for a row replacement, then $\det(EC) = \det(C)$. In other words, *left-multiplication by E does not change the determinant.*

2. If E is the elementary matrix for a row scale by a factor of c, then $\det(EC) = c\det(C)$. In other words, *left-multiplication by E scales the determinant by a factor of c.*

3. If E is the elementary matrix for a row swap, then $\det(EC) = -\det(C)$. In other words, *left-multiplication by E negates the determinant.*

Now we turn to the proof of the multiplicativity property. Suppose to begin that B is not invertible. Then AB is also not invertible: otherwise, $(AB)^{-1}AB = I_n$ implies $B^{-1} = (AB)^{-1}A$. By the invertibility property, both sides of the equation $\det(AB) = \det(A)\det(B)$ are zero.

Now assume that B is invertible, so $\det(B) \neq 0$. Define a function

$$d\colon \{n \times n \text{ matrices}\} \longrightarrow \mathbf{R} \quad \text{by} \quad d(C) = \frac{\det(CB)}{\det(B)}.$$

We claim that d satisfies the four defining properties of the determinant.

1. Let C' be the matrix obtained by doing a row replacement on C, and let E be the elementary matrix for this row replacement, so $C' = EC$. Since left-multiplication by E does not change the determinant, we have $\det(ECB) = \det(CB)$, so

$$d(C') = \frac{\det(C'B)}{\det(B)} = \frac{\det(ECB)}{\det(B)} = \frac{\det(CB)}{\det(B)} = d(C).$$

2. Let C' be the matrix obtained by scaling a row of C by a factor of c, and let E be the elementary matrix for this row replacement, so $C' = EC$. Since left-multiplication by E scales the determinant by a factor of c, we have $\det(ECB) = c\det(CB)$, so

$$d(C') = \frac{\det(C'B)}{\det(B)} = \frac{\det(ECB)}{\det(B)} = \frac{c\det(CB)}{\det(B)} = c \cdot d(C).$$

3. Let C' be the matrix obtained by swapping two rows of C, and let E be the elementary matrix for this row replacement, so $C' = EC$. Since left-multiplication by E negates the determinant, we have $\det(ECB) = -\det(CB)$, so

$$d(C') = \frac{\det(C'B)}{\det(B)} = \frac{\det(ECB)}{\det(B)} = \frac{-\det(CB)}{\det(B)} = -d(C).$$

4. We have

$$d(I_n) = \frac{\det(I_nB)}{\det(B)} = \frac{\det(B)}{\det(B)} = 1.$$

Since d satisfies the four defining properties of the determinant, *it is equal to the determinant* by the existence theorem. In other words, for all matrices A, we have

$$\det(A) = d(A) = \frac{\det(AB)}{\det(B)}.$$

Multiplying through by $\det(B)$ gives $\det(A)\det(B) = \det(AB)$. □

Recall that taking a power of a square matrix A means taking products of A with itself:

$$A^2 = AA \qquad A^3 = AAA \qquad \text{etc.}$$

If A is invertible, then we define

$$A^{-2} = A^{-1}A^{-1} \qquad A^{-3} = A^{-1}A^{-1}A^{-1} \qquad \text{etc.}$$

For completeness, we set $A^0 = I_n$ if $A \neq 0$.

Corollary. *If A is a square matrix, then*

$$\det(A^n) = \det(A)^n$$

for all $n \geq 1$. If A is invertible, then the equation holds for all $n \leq 0$ as well; in particular,

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Proof. Using the multiplicativity property, we compute

$$\det(A^2) = \det(AA) = \det(A)\det(A) = \det(A)^2$$

and

$$\det(A^3) = \det(AAA) = \det(A)\det(AA) = \det(A)\det(A)\det(A) = \det(A)^3;$$

the pattern is clear.

We have

$$1 = \det(I_n) = \det(AA^{-1}) = \det(A)\det(A^{-1})$$

by the multiplicativity property and the fourth defining property, which shows that $\det(A^{-1}) = \det(A)^{-1}$. Thus

$$\det(A^{-2}) = \det(A^{-1}A^{-1}) = \det(A^{-1})\det(A^{-1}) = \det(A^{-1})^2 = \det(A)^{-2},$$

and so on. □

Example. Compute $\det(A^{100})$, where

$$A = \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix}.$$

Solution. We have $\det(A) = 4 - 2 = 2$, so

$$\det(A^{100}) = \det(A)^{100} = 2^{100}.$$

Nowhere did we have to compute the 100th power of A! (We will learn an efficient way to do that in Section 5.4.)

Here is another application of the multiplicativity property.

Corollary. *Let $A_1, A_2, \ldots, A_k$ be $n \times n$ matrices. Then the product $A_1 A_2 \cdots A_k$ is invertible if and only if each A_i is invertible.*

Proof. The determinant of the product is the product of the determinants by the multiplicativity property:

$$\det(A_1 A_2 \cdots A_k) = \det(A_1)\det(A_2)\cdots\det(A_k).$$

By the invertibility property, this is nonzero if and only if $A_1 A_2 \cdots A_k$ is invertible. On the other hand, $\det(A_1)\det(A_2)\cdots\det(A_k)$ is nonzero if and only if each $\det(A_i) \neq 0$, which means each A_i is invertible. □

Example. For any number n we define

$$A_n = \begin{pmatrix} 1 & n \\ 1 & 2 \end{pmatrix}.$$

Show that the product

$$A_1 A_2 A_3 A_4 A_5$$

is not invertible.

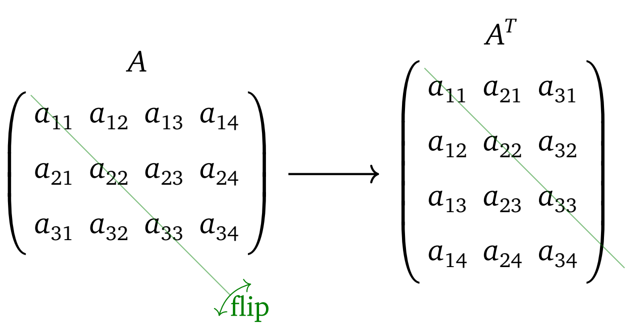
Solution. When $n = 2$, the matrix A_2 is not invertible, because its rows are identical:

$$A_2 = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}.$$

Hence any product involving A_2 is not invertible.

In order to state the transpose property, we need to define the transpose of a matrix.

Definition. The **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose rows are the columns of A. In other words, the ij entry of A^T is a_{ji}.



Like inversion, transposition reverses the order of matrix multiplication.

Fact. *Let A be an $m \times n$ matrix, and let B be an $n \times p$ matrix. Then*

$$(AB)^T = B^T A^T.$$

Proof. First suppose that A is a row vector an B is a column vector, i.e., $m = p = 1$. Then

$$\begin{aligned} AB &= \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \\ &= \begin{pmatrix} b_1 & b_2 & \cdots & b_n \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = B^T A^T. \end{aligned}$$

Now we use the row-column rule for matrix multiplication. Let $r_1, r_2, \ldots, r_m$ be the rows of A, and let $c_1, c_2, \ldots, c_p$ be the columns of B, so

$$AB = \begin{pmatrix} — r_1 — \\ — r_2 — \\ \vdots \\ — r_m — \end{pmatrix} \begin{pmatrix} | & | & & | \\ c_1 & c_2 & \cdots & c_p \\ | & | & & | \end{pmatrix} = \begin{pmatrix} r_1 c_1 & r_1 c_2 & \cdots & r_1 c_p \\ r_2 c_1 & r_2 c_2 & \cdots & r_2 c_p \\ \vdots & \vdots & & \vdots \\ r_m c_1 & r_m c_2 & \cdots & r_m c_p \end{pmatrix}.$$

By the case we handled above, we have $r_i c_j = c_j^T r_i^T$. Then

$$\begin{aligned} (AB)^T &= \begin{pmatrix} r_1 c_1 & r_2 c_1 & \cdots & r_m c_1 \\ r_1 c_2 & r_2 c_2 & \cdots & r_m c_2 \\ \vdots & \vdots & & \vdots \\ r_1 c_p & r_2 c_p & \cdots & r_m c_p \end{pmatrix} \\ &= \begin{pmatrix} c_1^T r_1^T & c_1^T r_2^T & \cdots & c_1^T r_m^T \\ c_2^T r_1^T & c_2^T r_2^T & \cdots & c_2^T r_m^T \\ \vdots & \vdots & & \vdots \\ c_p^T r_1^T & c_p^T r_2^T & \cdots & c_p^T r_m^T \end{pmatrix} \\ &= \begin{pmatrix} — c_1^T — \\ — c_2^T — \\ \vdots \\ — c_p^T — \end{pmatrix} \begin{pmatrix} | & | & & | \\ r_1^T & r_2^T & \cdots & r_m^T \\ | & | & & | \end{pmatrix} = B^T A^T. \end{aligned}$$

□

Transpose Property. *For any square matrix A, we have*

$$\det(A) = \det(A^T).$$

Proof. We follow the same strategy as in the proof of the multiplicativity property: namely, we define

$$d(A) = \det(A^T),$$

and we show that d satisfies the four defining properties of the determinant. Again we use elementary matrices, also introduced in the proof of the multiplicativity property.

1. Let C' be the matrix obtained by doing a row replacement on C, and let E be the elementary matrix for this row replacement, so $C' = EC$. The elementary matrix for a row replacement is either upper-triangular or lower-triangular, with ones on the diagonal:

$$R_1 = R_1 + 3R_3 : \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad R_3 = R_3 + 3R_1 : \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

   It follows that E^T is also either upper-triangular or lower-triangular, with ones on the diagonal, so $\det(E^T) = 1$ by this proposition. By the fact and the multiplicativity property,

$$\begin{aligned} d(C') = \det((C')^T) &= \det((EC)^T) = \det(C^T E^T) \\ &= \det(C^T)\det(E^T) = \det(C^T) = d(C). \end{aligned}$$

2. Let C' be the matrix obtained by scaling a row of C by a factor of c, and let E be the elementary matrix for this row replacement, so $C' = EC$. Then E is a diagonal matrix:

$$R_2 = cR_2 : \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

   Thus $\det(E^T) = c$. By the fact and the multiplicativity property,

$$\begin{aligned} d(C') = \det((C')^T) &= \det((EC)^T) = \det(C^T E^T) \\ &= \det(C^T)\det(E^T) = c\det(C^T) = c \cdot d(C). \end{aligned}$$

3. Let C' be the matrix obtained by swapping two rows of C, and let E be the elementary matrix for this row replacement, so $C' = EC$. The E is equal to its own transpose:

$$R_1 \longleftrightarrow R_2 : \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}^T.$$

   Since E (hence E^T) is obtained by performing one row swap on the identity matrix, we have $\det(E^T) = -1$. By the fact and the multiplicativity property,

$$\begin{aligned} d(C') = \det((C')^T) &= \det((EC)^T) = \det(C^T E^T) \\ &= \det(C^T)\det(E^T) = -\det(C^T) = -d(C). \end{aligned}$$

4. Since $I_n^T = I_n$, we have

$$d(I_n) = \det(I_n^T) = det(I_n) = 1.$$

Since d satisfies the four defining properties of the determinant, *it is equal to the determinant* by the existence theorem. In other words, for all matrices A, we have

$$\det(A) = d(A) = \det(A^T). \qquad \square$$

The transpose property is very useful. For concreteness, we note that $\det(A) = \det(A^T)$ means, for instance, that

$$\det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \det\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

This implies that the determinant has the curious feature that it also behaves well with respect to *column* operations. Indeed, a column operation on A is the same as a row operation on A^T, and $\det(A) = \det(A^T)$.

Corollary. *The determinant satisfies the following properties with respect to column operations:*

1. *Doing a column replacement on A does not change* $\det(A)$.
2. *Scaling a column of A by a scalar c multiplies the determinant by c.*
3. *Swapping two columns of a matrix multiplies the determinant by* -1.

The previous corollary makes it easier to compute the determinant: one is allowed to do row *and* column operations when simplifying the matrix. (Of course, one still has to keep track of how the row and column operations change the determinant.)

Example. Compute $\det\begin{pmatrix} 2 & 7 & 4 \\ 3 & 1 & 3 \\ 4 & 0 & 1 \end{pmatrix}$.

Solution. It takes fewer column operations than row operations to make this matrix upper-triangular:

$$\begin{pmatrix} 2 & 7 & 4 \\ 3 & 1 & 3 \\ 4 & 0 & 1 \end{pmatrix} \xrightarrow{C_1=C_1-4C_3} \begin{pmatrix} -14 & 7 & 4 \\ -9 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow{C_1=C_1+9C_2} \begin{pmatrix} 49 & 7 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

We performed two column replacements, which does not change the determinant; therefore,

$$\det\begin{pmatrix} 2 & 7 & 4 \\ 3 & 1 & 3 \\ 4 & 0 & 1 \end{pmatrix} = \det\begin{pmatrix} 49 & 7 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} = 49.$$

Multilinearity The following observation is useful for theoretical purposes.

We can think of det as a function of the rows of a matrix:

$$\det(v_1, v_2, \ldots, v_n) = \det\begin{pmatrix} - v_1 - \\ - v_2 - \\ \vdots \\ - v_n - \end{pmatrix}.$$

Multilinearity Property. *Let i be a whole number between 1 and n, and fix $n-1$ vectors $v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ in $\mathbf{R}^n$. Then the transformation $T: \mathbf{R}^n \to \mathbf{R}$ defined by*

$$T(x) = \det(v_1, v_2, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n)$$

is linear.

Proof. First assume that $i = 1$, so

$$T(x) = \det(x, v_2, \ldots, v_n).$$

We have to show that T satisfies the defining properties in Section 3.3.

- By the first defining property, scaling any row of a matrix by a number c scales the determinant by a factor of c. This implies that T satisfies the second property, i.e., that

$$T(cx) = \det(cx, v_2, \ldots, v_n) = c\det(x, v_2, \ldots, v_n) = cT(x).$$

- We claim that $T(v + w) = T(v) + T(w)$. If w is in $\mathrm{Span}\{v, v_2, \ldots, v_n\}$, then

$$w = cv + c_2v_2 + \cdots + c_nv_n$$

for some scalars $c, c_2, \ldots, c_n$. Let A be the matrix with rows $v + w, v_2, \ldots, v_n$, so $T(v + w) = \det(A)$. By performing the row operations

$$R_1 = R_1 - c_2R_2; \quad R_1 = R_1 - c_3R_3; \quad \ldots \quad R_1 = R_1 - c_nR_n,$$

the first row of the matrix A becomes

$$v + w - (c_2v_2 + \cdots + c_nv_n) = v + cv = (1 + c)v.$$

Therefore,

$$\begin{aligned} T(v+w) = \det(A) &= \det((1+c)v, v_2, \ldots, v_n) \\ &= (1+c)\det(v, v_2, \ldots, v_n) \\ &= T(v) + cT(v) = T(v) + T(cv). \end{aligned}$$

Doing the opposite row operations

$$R_1 = R_1 + c_2R_2; \quad R_1 = R_1 + c_3R_3; \quad \ldots \quad R_1 = R_1 + c_nR_n$$

to the matrix with rows $cv, v_2, \ldots, v_n$ shows that

$$\begin{aligned} T(cv) &= \det(cv, v_2, \ldots, v_n) \\ &= \det(cv + c_2v_2 + \cdots + c_nv_n, v_2, \ldots, v_n) \\ &= \det(w, v_2, \ldots, v_n) = T(w), \end{aligned}$$

which finishes the proof of the first property in this case.

Now suppose that w is not in $\mathrm{Span}\{v, v_2, \ldots, v_n\}$. This implies that $\{v, v_2, \ldots, v_n\}$ is linearly *dependent* (otherwise it would form a basis for $\mathbf{R}^n$), so $T(v) = 0$. If v is not in $\mathrm{Span}\{v_2, \ldots, v_n\}$, then $\{v_2, \ldots, v_n\}$ is linearly dependent by the increasing span criterion in Section 2.5, so $T(x) = 0$ for all x, as the matrix with rows $x, v_2, \ldots, v_n$ is not invertible. Hence we may assume v is in $\mathrm{Span}\{v_2, \ldots, v_n\}$. By the above argument with the roles of v and w reversed, we have $T(v+w) = T(v) + T(w)$.

For $i \neq 1$, we note that

$$\begin{aligned} T(x) &= \det(v_1, v_2, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_n) \\ &= -\det(x, v_2, \ldots, v_{i-1}, v_1, v_{i+1}, \ldots, v_n). \end{aligned}$$

By the previously handled case, we know that $-T$ is linear:

$$-T(cx) = -cT(x) \qquad -T(v+w) = -T(v) - T(w).$$

Multiplying both sides by -1, we see that T is linear. □

For example, we have

$$\det\begin{pmatrix} — & v_1 & — \\ — & av+bw & — \\ — & v_3 & — \end{pmatrix} = a\det\begin{pmatrix} — v_1 — \\ — v — \\ — v_3 — \end{pmatrix} + b\det\begin{pmatrix} — v_1 — \\ — w — \\ — v_3 — \end{pmatrix}$$

By the transpose property, the determinant is also multilinear in the *columns* of a matrix:

$$\det\begin{pmatrix} | & | & | \\ v_1 & av+bw & v_3 \\ | & | & | \end{pmatrix} = a\det\begin{pmatrix} | & | & | \\ v_1 & v & v_3 \\ | & | & | \end{pmatrix} + b\det\begin{pmatrix} | & | & | \\ v_1 & w & v_3 \\ | & | & | \end{pmatrix}.$$

Remark (Alternative defining properties)**.** In more theoretical treatments of the topic, where row reduction plays a secondary role, the defining properties of the determinant are often taken to be:

1. The determinant $\det(A)$ is multilinear in the rows of A.
2. If A has two identical rows, then $\det(A) = 0$.
3. The determinant of the identity matrix is equal to one.

We have already shown that our four defining properties imply these three. Conversely, we will prove that these three alternative properties imply our four, so that both sets of properties are equivalent.

Defining property 2 is just the second defining property in Section 3.3. Suppose that the rows of A are $v_1, v_2, \ldots, v_n$. If we perform the row replacement $R_i = R_i + cR_j$ on A, then the rows of our new matrix are $v_1, v_2, \ldots, v_{i-1}, v_i + cv_j, v_{i+1}, \ldots, v_n$, so by linearity in the ith row,

$$\begin{aligned}&\det(v_1, v_2, \ldots, v_{i-1}, v_i + cv_j, v_{i+1}, \ldots, v_n)\\ &\qquad = \det(v_1, v_2, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) + c\det(v_1, v_2, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_n)\\ &\qquad = \det(v_1, v_2, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) = \det(A),\end{aligned}$$

where $\det(v_1, v_2, \ldots, v_{i-1}, v_j, v_{i+1}, \ldots, v_n) = 0$ because v_j is repeated. Thus, the alternative defining properties imply our first two defining properties. For the third, suppose that we want to swap row i with row j. Using the second alternative defining property and multilinearity in the ith and jth rows, we have

$$\begin{aligned}0 &= \det(v_1, \ldots, v_i + v_j, \ldots, v_i + v_j, \ldots, v_n)\\ &= \det(v_1, \ldots, v_i, \ldots, v_i + v_j, \ldots, v_n) + \det(v_1, \ldots, v_j, \ldots, v_i + v_j, \ldots, v_n)\\ &= \det(v_1, \ldots, v_i, \ldots, v_i, \ldots, v_n) + \det(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n)\\ &\qquad + \det(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n) + \det(v_1, \ldots, v_j, \ldots, v_j, \ldots, v_n)\\ &= \det(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) + \det(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_n),\end{aligned}$$

as desired.

Example. We have

$$\begin{pmatrix}-1\\2\\3\end{pmatrix} = -\begin{pmatrix}1\\0\\0\end{pmatrix} + 2\begin{pmatrix}0\\1\\0\end{pmatrix} + 3\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

Therefore,

$$\begin{aligned}\det\begin{pmatrix}-1 & 7 & 2\\ 2 & -3 & 2\\ 3 & 1 & 1\end{pmatrix} &= -\det\begin{pmatrix}1 & 7 & 2\\ 0 & -3 & 2\\ 0 & 1 & 1\end{pmatrix}\\ &\quad + 2\det\begin{pmatrix}0 & 7 & 2\\ 1 & -3 & 2\\ 0 & 1 & 1\end{pmatrix} + 3\det\begin{pmatrix}0 & 7 & 2\\ 0 & -3 & 2\\ 1 & 1 & 1\end{pmatrix}.\end{aligned}$$

This is the basic idea behind cofactor expansions in Section 4.2.

Summary: Magical Properties of the Determinant.

1. There is one and only one function $\det\colon \{n \times n \text{ matrices}\} \to \mathbf{R}$ satisfying the four defining properties.
2. The determinant of an upper-triangular or lower-triangular matrix is the product of the diagonal entries.
3. A square matrix is invertible if and only if $\det(A) \neq 0$; in this case,
$$\det(A^{-1}) = \frac{1}{\det(A)}.$$
4. If A and B are $n \times n$ matrices, then
$$\det(AB) = \det(A)\det(B).$$
5. For any square matrix A, we have
$$\det(A^T) = \det(A).$$
6. The determinant can be computed by performing row and/or column operations.

4.2 Cofactor Expansions

Objectives

1. Learn to recognize which methods are best suited to compute the determinant of a given matrix.
2. *Recipes:* the determinant of a 3×3 matrix, compute the determinant using cofactor expansions.
3. *Vocabulary words:* **minor**, **cofactor**.

In this section, we give a recursive formula for the determinant of a matrix, called a *cofactor expansion*. The formula is recursive in that we will compute the determinant of an $n \times n$ matrix *assuming* we already know how to compute the determinant of an $(n-1) \times (n-1)$ matrix.

At the end is a supplementary subsection on Cramer's rule and a cofactor formula for the inverse of a matrix.

4.2.1 Cofactor Expansions

A recursive formula must have a starting point. For cofactor expansions, the starting point is the case of 1×1 matrices. The definition of determinant directly implies that

$$\det\begin{pmatrix} a \end{pmatrix} = a.$$

To describe cofactor expansions, we need to introduce some notation.

Definition. Let A be an $n \times n$ matrix.

1. The (i,j) **minor**, denoted A_{ij}, is the $(n-1)\times(n-1)$ matrix obtained from A by deleting the ith row and the jth column.

2. The (i,j) **cofactor** C_{ij} is defined in terms of the minor by

$$C_{ij} = (-1)^{i+j}\det(A_{ij}).$$

Note that the signs of the cofactors follow a "checkerboard pattern." Namely, $(-1)^{i+j}$ is pictured in this matrix:

$$\begin{pmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{pmatrix}.$$

Example. For

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$

compute A_{23} and C_{23}.

Solution.

$$A_{23} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} \qquad C_{23} = (-1)^{2+3}\det\begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix} = (-1)(-6) = 6$$

The cofactors C_{ij} of an $n \times n$ matrix are determinants of $(n-1)\times(n-1)$ submatrices. Hence the following theorem is in fact a recursive procedure for computing the determinant.

Theorem (Cofactor expansion)**.** *Let A be an $n \times n$ matrix with entries a_{ij}.*

1. *For any $i = 1, 2, \ldots, n$, we have*

$$\det(A) = \sum_{j=1}^{n} a_{ij}C_{ij} = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.$$

   This is called ***cofactor expansion along the ith row.***

2. *For any $j = 1, 2, \ldots, n$, we have*

$$\det(A) = \sum_{i=1}^{n} a_{ij}C_{ij} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}.$$

   This is called ***cofactor expansion along the jth column.***

Proof. First we will prove that cofactor expansion along the first column computes the determinant. Define a function $d\colon \{n \times n \text{ matrices}\} \to \mathbf{R}$ by

$$d(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det(A_{i1}).$$

We want to show that $d(A) = \det(A)$. Instead of showing that d satisfies the four defining properties of the determinant in Section 4.1, we will prove that it satsifies the three alternative defining properties in Section 4.1, which were shown to be equivalent.

1. We claim that d is multilinear in the rows of A. Let A be the matrix with rows $v_1, v_2, \ldots, v_{i-1}, v + w, v_{i+1}, \ldots, v_n$:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ b_1 + c_1 & b_2 + c_2 & b_3 + c_3 \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

   Here we let b_i and c_i be the entries of v and w, respectively. Let B and C be the matrices with rows $v_1, v_2, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_n$ and $v_1, v_2, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n$, respectively:

$$B = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ b_1 & b_2 & b_3 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad C = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ c_1 & c_2 & c_3 \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

   We wish to show $d(A) = d(B) + d(C)$. For $i' \neq i$, the $(i', 1)$-cofactor of A is the sum of the $(i', 1)$-cofactors of B and C, by multilinearity of the determinants of $(n-1) \times (n-1)$ matrices:

$$\begin{aligned} (-1)^{3+1}\det(A_{31}) &= (-1)^{3+1}\det\begin{pmatrix} a_{12} & a_{13} \\ b_2 + c_2 & b_3 + c_3 \end{pmatrix} \\ &= (-1)^{3+1}\det\begin{pmatrix} a_{12} & a_{13} \\ b_2 & b_3 \end{pmatrix} + (-1)^{3+1}\det\begin{pmatrix} a_{12} & a_{13} \\ c_2 & c_3 \end{pmatrix} \\ &= (-1)^{3+1}\det(B_{31}) + (-1)^{3+1}\det(C_{31}). \end{aligned}$$

On the other hand, the $(i,1)$-cofactors of A, B, and C are all the same:

$$\begin{aligned}(-1)^{2+1}\det(A_{21}) &= (-1)^{2+1}\det\begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33}\end{pmatrix} \\ &= (-1)^{2+1}\det(B_{21}) = (-1)^{2+1}\det(C_{21}).\end{aligned}$$

Now we compute

$$\begin{aligned} d(A) &= (-1)^{i+1}(b_i + c_i)\det(A_{i1}) + \sum_{i'\neq i}(-1)^{i'+1}a_{i1}\det(A_{i'1}) \\ &= (-1)^{i+1}b_i\det(B_{i1}) + (-1)^{i+1}c_i\det(C_{i1}) \\ &\qquad + \sum_{i'\neq i}(-1)^{i'+1}a_{i1}\big(\det(B_{i'1}) + \det(C_{i'1})\big) \\ &= \left[(-1)^{i+1}b_i\det(B_{i1}) + \sum_{i'\neq i}(-1)^{i'+1}a_{i1}\det(B_{i'1})\right] \\ &\qquad + \left[(-1)^{i+1}c_i\det(C_{i1}) + \sum_{i'\neq i}(-1)^{i'+1}a_{i1}\det(C_{i'1})\right] \\ &= d(B) + d(C), \end{aligned}$$

as desired. This shows that $d(A)$ satisfies the first defining property in the rows of A.

We still have to show that $d(A)$ satisfies the second defining property in the rows of A. Let B be the matrix obtained by scaling the ith row of A by a factor of c:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}\end{pmatrix} \qquad B = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ ca_{21} & ca_{22} & ca_{23} \\ a_{31} & a_{32} & a_{33}\end{pmatrix}.$$

We wish to show that $d(B) = c\,d(A)$. For $i' \neq i$, the $(i',1)$-cofactor of B is c times the $(i',1)$-cofactor of A, by multilinearity of the determinants of $(n-1)\times(n-1)$-matrices:

$$\begin{aligned}(-1)^{3+1}\det(B_{31}) &= (-1)^{3+1}\det\begin{pmatrix} a_{12} & a_{13} \\ ca_{22} & ca_{23}\end{pmatrix} \\ &= (-1)^{3+1}\cdot c\det\begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23}\end{pmatrix} = (-1)^{3+1}\cdot c\det(A_{31}).\end{aligned}$$

On the other hand, the $(i,1)$-cofactors of A and B are the same:

$$(-1)^{2+1}\det(B_{21}) = (-1)^{2+1}\det\begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33}\end{pmatrix} = (-1)^{2+1}\det(A_{21}).$$

Now we compute

$$\begin{aligned} d(B) &= (-1)^{i+1} c a_{i1} \det(B_{i1}) + \sum_{i' \neq i} (-1)^{i'+1} a_{i'1} \det(B_{i'1}) \\ &= (-1)^{i+1} c a_{i1} \det(A_{i1}) + \sum_{i' \neq i} (-1)^{i'+1} a_{i'1} \cdot c \det(A_{i'1}) \\ &= c \left[(-1)^{i+1} c a_{i1} \det(A_{i1}) + \sum_{i' \neq i} (-1)^{i'+1} a_{i'1} \det(A_{i'1}) \right] \\ &= c\, d(A), \end{aligned}$$

as desired. This completes the proof that $d(A)$ is multilinear in the rows of A.

2. Now we show that $d(A) = 0$ if A has two identical rows. Suppose that rows i_1, i_2 of A are identical, with $i_1 < i_2$:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{11} & a_{12} & a_{13} & a_{14} \end{pmatrix}.$$

If $i \neq i_1, i_2$ then the $(i,1)$-cofactor of A is equal to zero, since A_{i1} is an $(n-1) \times (n-1)$ matrix with identical rows:

$$(-1)^{2+1} \det(A_{21}) = (-1)^{2+1} \det \begin{pmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \\ a_{12} & a_{13} & a_{14} \end{pmatrix} = 0.$$

The $(i_1, 1)$-minor can be transformed into the $(i_2, 1)$-minor using $i_2 - i_1 - 1$ row swaps:

$$A_{11} = \begin{pmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{12} & a_{13} & a_{14} \end{pmatrix} \quad \begin{pmatrix} a_{22} & a_{23} & a_{24} \\ a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \end{pmatrix} \quad \begin{pmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{pmatrix} = A_{41}$$

Therefore,

$$(-1)^{i_1+1} \det(A_{i_1 1}) = (-1)^{i_1+1} \cdot (-1)^{i_2 - i_1 - 1} \det(A_{i_2 1}) = -(-1)^{i_2+1} \det(A_{i_2 1}).$$

The two remaining cofactors cancel out, so $d(A) = 0$, as desired.

3. It remains to show that $d(I_n) = 1$. The first is the only one nonzero term in the cofactor expansion of the identity:

$$d(I_n) = 1 \cdot (-1)^{1+1} \det(I_{n-1}) = 1.$$

This proves that $\det(A) = d(A)$, i.e., that cofactor expansion along the first column computes the determinant.

Now we show that cofactor expansion along the jth column also computes the determinant. By performing $j-1$ column swaps, one can move the jth column of a matrix to the first column, keeping the other columns in order. For example, here we move the third column to the first, using two column swaps:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \qquad \begin{pmatrix} a_{11} & a_{13} & a_{12} & a_{14} \\ a_{21} & a_{23} & a_{22} & a_{24} \\ a_{31} & a_{33} & a_{32} & a_{34} \\ a_{41} & a_{43} & a_{42} & a_{44} \end{pmatrix} \qquad \begin{pmatrix} a_{13} & a_{12} & a_{11} & a_{14} \\ a_{23} & a_{22} & a_{21} & a_{24} \\ a_{33} & a_{32} & a_{31} & a_{34} \\ a_{43} & a_{42} & a_{41} & a_{44} \end{pmatrix}$$

Let B be the matrix obtained by moving the jth column of A to the first column in this way. Then the (i,j) minor A_{ij} is equal to the $(i,1)$ minor B_{i1}, since deleting the ith column of A is the same as deleting the first column of B. By construction, the (i,j)-entry a_{ij} of A is equal to the $(i,1)$-entry b_{i1} of B. Since we know that we can compute determinants by expanding along the first column, we have

$$\det(B) = \sum_{i=1}^{n} (-1)^{i+1} b_{i1} \det(B_{i1}) = \sum_{i=1}^{n} (-1)^{i+1} a_{ij} \det(A_{ij}).$$

Since B was obtained from A by performing $j-1$ column swaps, we have

$$\begin{aligned} \det(A) = (-1)^{j-1} \det(B) &= (-1)^{j-1} \sum_{i=1}^{n} (-1)^{i+1} a_{ij} \det(A_{ij}) \\ &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}). \end{aligned}$$

This proves that cofactor expansion along the ith column computes the determinant of A.

By the transpose property in Section 4.1, the cofactor expansion along the ith row of A is the same as the cofactor expansion along the ith column of A^T. Again by the transpose property, we have $\det(A) = \det(A^T)$, so expanding cofactors along a row also computes the determinant. □

Note that the theorem actually gives $2n$ different formulas for the determinant: one for each row and one for each column. For instance, the formula for cofactor expansion along the first column is

$$\begin{aligned} \det(A) = \sum_{i=1}^{n} a_{i1} C_{i1} &= a_{11} C_{11} + a_{21} C_{21} + \cdots + a_{n1} C_{n1} \\ &= a_{11} \det(A_{11}) - a_{21} \det(A_{21}) + a_{31} \det(A_{31}) - \cdots \pm a_{n1} \det(A_{n1}). \end{aligned}$$

Remember, the determinant of a matrix is just a number, defined by the four defining properties in Section 4.1, so to be clear:

You obtain the same number by expanding cofactors along *any* row or column.

Now that we have a recursive formula for the determinant, we can finally prove the existence theorem in Section 4.1.

Proof. Let us review what we actually proved in Section 4.1. We showed that if $\det\colon \{n\times n \text{ matrices}\} \to \mathbf{R}$ is any function satisfying the four defining properties of the determinant (or the three alternative defining properties), then it also satisfies all of the wonderful properties proved in that section. In particular, since det can be computed using row reduction by this recipe in Section 4.1, it is uniquely characterized by the defining properties. What we did not prove was the existence of such a function, since we did not know that two different row reduction procedures would always compute the same answer.

Consider the function d defined by cofactor expansion along the first row:

$$d(A) = \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det(A_{i1}).$$

If we assume that the determinant exists for $(n-1)\times(n-1)$ matrices, then there is no question that the function d exists, since we gave a formula for it. Moreover, we showed in the proof of the theorem above that d satisfies the three alternative defining properties of the determinant, again only assuming that the determinant exists for $(n-1)\times(n-1)$ matrices. This proves the existence of the determinant for $n\times n$ matrices!

This is an example of a proof by *mathematical induction*. We start by noticing that $\det\begin{pmatrix} a \end{pmatrix} = a$ satisfies the four defining properties of the determinant of a 1×1 matrix. Then we showed that the determinant of $n\times n$ matrices exists, assuming the determinant of $(n-1)\times(n-1)$ matrices exists. This implies that all determinants exist, by the following chain of logic:

$$1\times 1 \text{ exists} \implies 2\times 2 \text{ exists} \implies 3\times 3 \text{ exists} \implies \cdots. \qquad \square$$

Example. Find the determinant of

$$A = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{pmatrix}.$$

Solution. We make the somewhat arbitrary choice to expand along the first row. The minors and cofactors are

$$A_{11} = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix} \qquad C_{11} = +\det\begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix} = 4$$

$$A_{12} = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} \qquad C_{12} = -\det\begin{pmatrix} -1 & 1 \\ -2 & 3 \end{pmatrix} = 1$$

$$A_{13} = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 2 & 1 \\ -2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ -2 & 2 \end{pmatrix} \qquad C_{13} = +\det\begin{pmatrix} -1 & 2 \\ -2 & 2 \end{pmatrix} = 2.$$

Thus,

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = (2)(4) + (1)(1) + (3)(2) = 15.$$

The determinant of a 2×2 matrix. Let us compute (again) the determinant of a general 2×2 matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

The minors are

$$A_{11} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d \end{pmatrix} \qquad A_{12} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c \end{pmatrix}$$

$$A_{21} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} b \end{pmatrix} \qquad A_{22} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \end{pmatrix}.$$

The minors are all 1×1 matrices. As we have seen that the determinant of a 1×1 matrix is just the number inside of it, the cofactors are therefore

$$C_{11} = +\det(A_{11}) = d \qquad C_{12} = -\det(A_{12}) = -c$$
$$C_{21} = -\det(A_{21}) = -b \qquad C_{22} = +\det(A_{22}) = a$$

Expanding cofactors along the first column, we find that

$$\det(A) = aC_{11} + cC_{21} = ad - bc,$$

which agrees with the formulas in this definition in Section 3.5 and this example in Section 4.1.

The determinant of a 3×3 matrix. We can also use cofactor expansions to find a formula for the determinant of a 3×3 matrix. Let is compute the determinant of

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

by expanding along the first row. The minors and cofactors are:

$$A_{11} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \qquad C_{11} = +\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$$

$$A_{12} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \qquad C_{12} = -\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}$$

$$A_{13} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \qquad C_{13} = +\det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

The determinant is:

$$\begin{aligned} \det(A) &= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \\ &= a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12}\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13}\det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

The formula for the determinant of a 3×3 matrix looks too complicated to memorize outright. Fortunately, there is the following mnemonic device.

Recipe: Computing the Determinant of a 3×3 Matrix. To compute the determinant of a 3×3 matrix, first draw a larger matrix with the first two columns repeated on the right. Then add the products of the downward diagonals together, and subtract the products of the upward diagonals:

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{array}{l} a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ \quad - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \end{array}$$

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix} \quad - \quad \begin{matrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{matrix}$$

Alternatively, it is not necessary to repeat the first two columns if you allow your diagonals to "wrap around" the sides of a matrix, like in Pac-Man or Asteroids.

Example. Find the determinant of $A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 0 & -1 \\ 4 & -3 & 1 \end{pmatrix}$.

Solution. We repeat the first two columns on the right, then add the products of the downward diagonals and subtract the products of the upward diagonals:

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

$$\det\begin{pmatrix} 1 & 3 & 5 \\ 2 & 0 & -1 \\ 4 & -3 & 1 \end{pmatrix} = \begin{matrix} (1)(0)(1) + (3)(-1)(4) + (5)(2)(-3) \\ -(5)(0)(4) - (1)(-1)(-3) - (3)(2)(1) \end{matrix} = -51.$$

Cofactor expansions are most useful when computing the determinant of a matrix that has a row or column with several zero entries. Indeed, if the (i,j) entry of A is zero, then there is no reason to compute the (i,j) cofactor. In the following example we compute the determinant of a matrix with two zeros in the fourth column by expanding cofactors along the fourth column.

Example. Find the determinant of

$$A = \begin{pmatrix} 2 & 5 & -3 & -2 \\ -2 & -3 & 2 & -5 \\ 1 & 3 & -2 & 0 \\ -1 & 6 & 4 & 0 \end{pmatrix}.$$

Solution. The fourth column has two zero entries. We expand along the fourth column to find

$$\begin{aligned} \det(A) = 2\det\begin{pmatrix} -2 & -3 & 2 \\ 1 & 3 & -2 \\ -1 & 6 & 4 \end{pmatrix} - 5\det\begin{pmatrix} 2 & 5 & -3 \\ 1 & 3 & -2 \\ -1 & 6 & 4 \end{pmatrix} \\ - 0\det(\text{don't care}) + 0\det(\text{don't care}). \end{aligned}$$

We only have to compute two cofactors. We can find these determinants using any method we wish; for the sake of illustration, we will expand cofactors on one and use the formula for the 3×3 determinant on the other.

Expanding along the first column, we compute

$$\begin{aligned} &\det\begin{pmatrix} -2 & -3 & 2 \\ 1 & 3 & -2 \\ -1 & 6 & 4 \end{pmatrix} \\ &= -2\det\begin{pmatrix} 3 & -2 \\ 6 & 4 \end{pmatrix} - \det\begin{pmatrix} -3 & 2 \\ 6 & 4 \end{pmatrix} - \det\begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix} \\ &= -2(24) - (-24) - 0 = -48 + 24 + 0 = -24. \end{aligned}$$

Using the formula for the 3×3 determinant, we have

$$\det\begin{pmatrix} 2 & 5 & -3 \\ 1 & 3 & -2 \\ -1 & 6 & 4 \end{pmatrix} = \begin{array}{l} (2)(3)(4)+(5)(-2)(-1)+(-3)(1)(6) \\ \quad -(2)(-2)(6)-(5)(1)(4)-(-3)(3)(-1) \end{array} = 11.$$

Thus, we find that

$$\det(A) = 2(-24) - 5(11) = -103.$$

Cofactor expansions are also very useful when computing the determinant of a matrix with unknown entries. Indeed, it is inconvenient to row reduce in this case, because one cannot be sure whether an entry containing an unknown is a pivot or not.

Example. Compute the determinant of this matrix containing the unknown λ:

$$A = \begin{pmatrix} -\lambda & 2 & 7 & 12 \\ 3 & 1-\lambda & 2 & -4 \\ 0 & 1 & -\lambda & 7 \\ 0 & 0 & 0 & 2-\lambda \end{pmatrix}.$$

Solution. First we expand cofactors along the fourth row:

$$\begin{aligned} \det(A) &= 0\det\big(\cdots\big) + 0\det\big(\cdots\big) + 0\det\big(\cdots\big) \\ &\quad + (2-\lambda)\det\begin{pmatrix} -\lambda & 2 & 7 \\ 3 & 1-\lambda & 2 \\ 0 & 1 & -\lambda \end{pmatrix}. \end{aligned}$$

We only have to compute one cofactor. To do so, first we clear the $(3,3)$-entry by performing the column replacement $C_3 = C_3 + \lambda C_2$, which does not change the determinant:

$$\det\begin{pmatrix} -\lambda & 2 & 7 \\ 3 & 1-\lambda & 2 \\ 0 & 1 & -\lambda \end{pmatrix} = \det\begin{pmatrix} -\lambda & 2 & 7+2\lambda \\ 3 & 1-\lambda & 2+\lambda(1-\lambda) \\ 0 & 1 & 0 \end{pmatrix}.$$

Now we expand cofactors along the third row to find

$$\begin{aligned} \det\begin{pmatrix} -\lambda & 2 & 7+2\lambda \\ 3 & 1-\lambda & 2+\lambda(1-\lambda) \\ 0 & 1 & 0 \end{pmatrix} &= (-1)^{2+3}\det\begin{pmatrix} -\lambda & 7+2\lambda \\ 3 & 2+\lambda(1-\lambda) \end{pmatrix} \\ &= -\Big(-\lambda\big(2+\lambda(1-\lambda)\big) - 3(7+2\lambda)\Big) \\ &= -\lambda^3 + \lambda^2 + 8\lambda + 21. \end{aligned}$$

Therefore, we have

$$\det(A) = (2-\lambda)(-\lambda^3 + \lambda^2 + 8\lambda + 21) = \lambda^4 - 3\lambda^3 - 6\lambda^2 - 5\lambda + 42.$$

It is often most efficient to use a combination of several techniques when computing the determinant of a matrix. Indeed, when expanding cofactors on a matrix, one can compute the determinants of the cofactors in whatever way is most convenient. Or, one can perform row and column operations to clear some entries of a matrix before expanding cofactors, as in the previous example.

Summary: methods for computing determinants. We have several ways of computing determinants:

1. *Special formulas for* 2×2 *and* 3×3 *matrices.*

   This is usually the best way to compute the determinant of a small matrix, except for a 3×3 matrix with several zero entries.

2. *Cofactor expansion.*

   This is usually most efficient when there is a row or column with several zero entries, or if the matrix has unknown entries.

3. *Row and column operations.*

   This is generally the fastest when presented with a large matrix which does not have a row or column with a lot of zeros in it.

4. *Any combination of the above.*

   Cofactor expansion is recursive, but one can compute the determinants of the minors using whatever method is most convenient. Or, you can perform row and column operations to clear some entries of a matrix before expanding cofactors.

Remember, *all methods for computing the determinant yield the same number.*

4.2.2 Cramer's Rule and Matrix Inverses

Recall from this proposition in Section 3.5 that one can compute the determinant of a 2×2 matrix using the rule

$$A = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \implies A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

We computed the cofactors of a 2×2 matrix in this example; using $C_{11} = d$, $C_{12} = -c$, $C_{21} = -b$, $C_{22} = a$, we can rewrite the above formula as

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{pmatrix}.$$

It turns out that this formula generalizes to $n \times n$ matrices.

Theorem. *Let A be an invertible $n \times n$ matrix, with cofactors C_{ij}. Then*

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n-1,1} & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n-1,2} & C_{n2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{1,n-1} & C_{2,n-1} & \cdots & C_{n-1,n-1} & C_{n,n-1} \\ C_{1n} & C_{2n} & \cdots & C_{n-1,n} & C_{nn} \end{pmatrix}. \tag{4.2.1}$$

The matrix of cofactors is sometimes called the **adjugate matrix** of A, and is denoted $\mathrm{adj}(A)$:

$$\mathrm{adj}(A) = \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n-1,1} & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n-1,2} & C_{n2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{1,n-1} & C_{2,n-1} & \cdots & C_{n-1,n-1} & C_{n,n-1} \\ C_{1n} & C_{2n} & \cdots & C_{n-1,n} & C_{nn} \end{pmatrix}.$$

Note that the (i,j) cofactor C_{ij} goes in the (j,i) entry the adjugate matrix, not the (i,j) entry: the adjugate matrix is the *transpose* of the cofactor matrix.

Remark. In fact, one always has $A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) I_n$, whether or not A is invertible.

Example. Use the theorem to compute A^{-1}, where

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Solution. The minors are:

$$A_{11} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad A_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad A_{13} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

$$A_{21} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad A_{22} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \qquad A_{23} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

$$A_{31} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \qquad A_{32} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad A_{33} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The cofactors are:

$$\begin{array}{lll} C_{11} = -1 & C_{12} = 1 & C_{13} = -1 \\ C_{21} = 1 & C_{22} = -1 & C_{23} = -1 \\ C_{31} = -1 & C_{32} = -1 & C_{33} = 1 \end{array}$$

Expanding along the first row, we compute the determinant to be

$$\det(A) = 1 \cdot C_{11} + 0 \cdot C_{12} + 1 \cdot C_{13} = -2.$$

Therefore, the inverse is

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} & C_{31} \\ C_{12} & C_{22} & C_{32} \\ C_{13} & C_{23} & C_{33} \end{pmatrix} = -\frac{1}{2} \begin{pmatrix} -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{pmatrix}.$$

It is clear from the previous example that (4.2.1) is a very inefficient way of computing the inverse of a matrix, compared to augmenting by the identity matrix and row reducing, as in this subsection in Section 3.5. However, it has its uses.

- If a matrix has unknown entries, then it is difficult to compute its inverse using row reduction, for the same reason it is difficult to compute the determinant that way: one cannot be sure whether an entry containing an unknown is a pivot or not.

- This formula is useful for theoretical purposes. Notice that the only denominators in (4.2.1) occur when dividing by the determinant: computing cofactors only involves multiplication and addition, never division. This means, for instance, that if the determinant is very small, then any measurement error in the entries of the matrix is greatly magnified when computing the inverse. In this way, (4.2.1) is useful in error analysis.

The proof of the theorem uses an interesting trick called *Cramer's Rule*, which gives a formula for the entries of the solution of an invertible matrix equation.

Cramer's Rule. *Let $x = (x_1, x_2, \ldots, x_n)$ be the solution of $Ax = b$, where A is an invertible $n \times n$ matrix and b is a vector in $\mathbf{R}^n$. Let A_i be the matrix obtained from A by replacing the ith column by b. Then*

$$x_i = \frac{\det(A_i)}{\det(A)}.$$

Proof. First suppose that A is the identity matrix, so that $x = b$. Then the matrix A_i looks like this:

$$\begin{pmatrix} 1 & 0 & b_1 & 0 \\ 0 & 1 & b_2 & 0 \\ 0 & 0 & b_3 & 0 \\ 0 & 0 & b_4 & 1 \end{pmatrix}.$$

Expanding cofactors along the ith row, we see that $\det(A_i) = b_i$, so in this case,

$$x_i = b_i = \det(A_i) = \frac{\det(A_i)}{\det(A)}.$$

Now let A be a general $n \times n$ matrix. One way to solve $Ax = b$ is to row reduce the augmented matrix $(\, A \mid b \,)$; the result is $(\, I_n \mid x \,)$. By the case we handled above, it is enough to check that the quantity $\det(A_i)/\det(A)$ does not change when we do a row operation to $(\, A \mid b \,)$, since $\det(A_i)/\det(A) = x_i$ when $A = I_n$.

1. Doing a row replacement on $(\, A \mid b \,)$ does the same row replacement on A and on A_i:

$$\left(\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right) \xrightarrow{R_2 = R_2 - 2R_3} \left(\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} - 2a_{31} & a_{22} - 2a_{32} & a_{23} - 2a_{33} & b_2 - 2b_3 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right)$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \xrightarrow{R_2 = R_2 - 2R_3} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} - 2a_{31} & a_{22} - 2a_{32} & a_{23} - 2a_{33} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix} \xrightarrow{R_2 = R_2 - 2R_3} \begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} - 2a_{31} & b_2 - 2b_3 & a_{23} - 2a_{33} \\ a_{31} & b_3 & a_{33} \end{pmatrix}.$$

   In particular, $\det(A)$ and $\det(A_i)$ are unchanged, so $\det(A)/\det(A_i)$ is unchanged.

2. Scaling a row of $(\, A \mid b \,)$ by a factor of c scales the same row of A and of A_i by the same factor:

$$\left(\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right) \xrightarrow{R_2 = cR_2} \left(\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ ca_{21} & ca_{22} & ca_{23} & cb_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right)$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \xrightarrow{R_2 = cR_2} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ ca_{21} & ca_{22} & ca_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$\begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix} \xrightarrow{R_2 = cR_2} \begin{pmatrix} a_{11} & b_1 & a_{13} \\ ca_{21} & cb_2 & ca_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix}.$$

   In particular, $\det(A)$ and $\det(A_i)$ are both scaled by a factor of c, so $\det(A_i)/\det(A)$ is unchanged.

3. Swapping two rows of $(A \mid b)$ swaps the same rows of A and of A_i:

$$\left(\begin{array}{ccc|c} a_{11} & a_{12} & a_{13} & b_1 \\ a_{21} & a_{22} & a_{23} & b_2 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right) \xrightarrow{R_1 \longleftrightarrow R_2} \left(\begin{array}{ccc|c} a_{21} & a_{22} & a_{23} & b_2 \\ a_{11} & a_{12} & a_{13} & b_1 \\ a_{31} & a_{32} & a_{33} & b_3 \end{array}\right)$$
$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \xrightarrow{R_1 \longleftrightarrow R_2} \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$
$$\begin{pmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{pmatrix} \xrightarrow{R_1 \longleftrightarrow R_2} \begin{pmatrix} a_{21} & b_2 & a_{23} \\ a_{11} & b_1 & a_{13} \\ a_{31} & b_3 & a_{33} \end{pmatrix}.$$

In particular, $\det(A)$ and $\det(A_i)$ are both negated, so $\det(A_i)/\det(A)$ is unchanged.

□

Example. Compute the solution of $Ax = b$ using Cramer's rule, where

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Here the coefficients of A are unknown, but A may be assumed invertible.

Solution. First we compute the determinants of the matrices obtained by replacing the columns of A with b:

$$A_1 = \begin{pmatrix} 1 & b \\ 2 & d \end{pmatrix} \qquad \det(A_1) = d - 2b$$
$$A_2 = \begin{pmatrix} a & 1 \\ c & 2 \end{pmatrix} \qquad \det(A_2) = 2a - c.$$

Now we compute

$$\frac{\det(A_1)}{\det(A)} = \frac{d-2b}{ad-bc} \qquad \frac{\det(A_2)}{\det(A)} = \frac{2a-c}{ad-bc}.$$

It follows that

$$x = \frac{1}{ad-bc}\begin{pmatrix} d-2b \\ 2a-c \end{pmatrix}.$$

Now we use Cramer's rule to prove the first theorem of this subsection.

Proof. The jth column of A^{-1} is $x_j = A^{-1}e_j$. This vector is the solution of the matrix equation

$$Ax = A\big(A^{-1}e_j\big) = I_n e_j = e_j.$$

By Cramer's rule, the ith entry of x_j is $\det(A_i)/\det(A)$, where A_i is the matrix obtained from A by replacing the ith column of A by e_j:

$$A_i = \begin{pmatrix} a_{11} & a_{12} & 0 & a_{14} \\ a_{21} & a_{22} & 1 & a_{24} \\ a_{31} & a_{32} & 0 & a_{34} \\ a_{41} & a_{42} & 0 & a_{44} \end{pmatrix} \qquad (i=3,\ j=2).$$

Expanding cofactors along the ith column, we see the determinant of A_i is exactly the (j,i)-cofactor C_{ji} of A. Therefore, the jth column of A^{-1} is

$$x_j = \frac{1}{\det(A)} \begin{pmatrix} C_{j1} \\ C_{j2} \\ \vdots \\ C_{jn} \end{pmatrix},$$

and thus

$$A^{-1} = \begin{pmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & & | \end{pmatrix} = \frac{1}{\det(A)} \begin{pmatrix} C_{11} & C_{21} & \cdots & C_{n-1,1} & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n-1,2} & C_{n2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_{1,n-1} & C_{2,n-1} & \cdots & C_{n-1,n-1} & C_{n,n-1} \\ C_{1n} & C_{2n} & \cdots & C_{n-1,n} & C_{nn} \end{pmatrix}. \qquad \square$$

4.3 Determinants and Volumes

Objectives

1. Understand the relationship between the determinant of a matrix and the volume of a parallelepiped.
2. Learn to use determinants to compute volumes of parallelograms and triangles.
3. Learn to use determinants to compute the volume of some curvy shapes like ellipses.
4. *Pictures:* parallelepiped, the image of a curvy shape under a linear transformation.
5. *Theorem:* determinants and volumes.
6. *Vocabulary word:* **parallelepiped**.

In this section we give a geometric interpretation of determinants, in terms of *volumes*. This will shed light on the reason behind three of the four defining properties of the determinant. It is also a crucial ingredient in the change-of-variables formula in multivariable calculus.

4.3.1 Parallelograms and Paralellepipeds

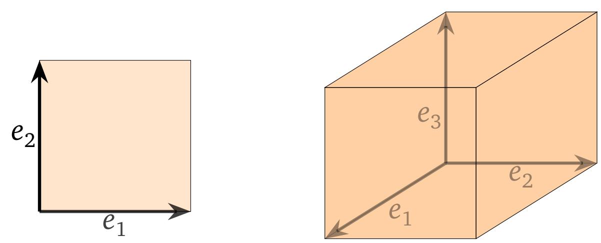
The determinant computes the volume of the following kind of geometric object.

Definition. The **paralellepiped** determined by n vectors $v_1, v_2, \ldots, v_n$ in $\mathbf{R}^n$ is the subset

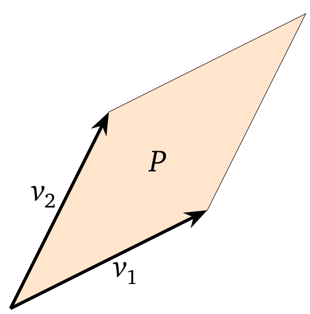
$$P = \left\{ a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \;\middle|\; 0 \le a_1, a_2, \ldots, a_n \le 1 \right\}.$$

In other words, a parallelepiped is the set of all linear combinations of n vectors with coefficients in $[0,1]$. We can draw parallelepipeds using the parallelogram law for vector addition.

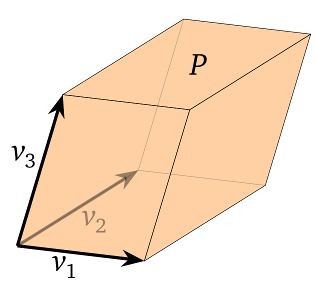
Example (The unit cube)**.** The parallelepiped determined by the standard coordinate vectors $e_1, e_2, \ldots, e_n$ is the unit n-dimensional cube.



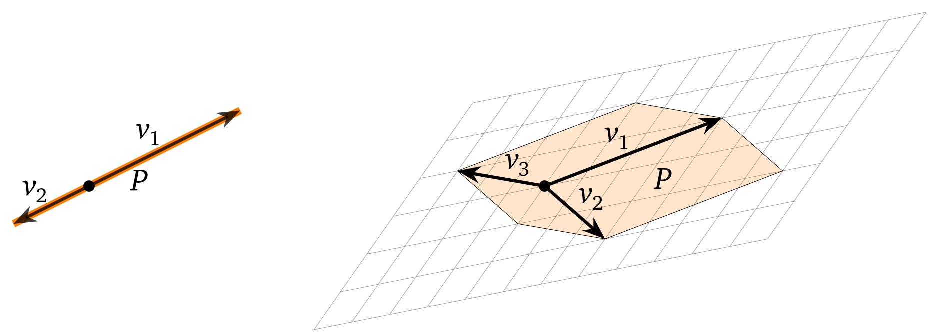
Example (Parallelograms)**.** When $n = 2$, a paralellepiped is just a paralellogram in $\mathbf{R}^2$. Note that the edges come in parallel pairs.



Example. When $n = 3$, a parallelepiped is a kind of a skewed cube. Note that the faces come in parallel pairs.



When does a parallelepiped have zero volume? This can happen only if the parallelepiped is flat, i.e., it is squashed into a lower dimension.



This means exactly that $\{v_1, v_2, \ldots, v_n\}$ is *linearly dependent,* which by this corollary in Section 4.1 means that the matrix with rows $v_1, v_2, \ldots, v_n$ has determinant zero. To summarize:

Key Observation. The parallelepiped defined by $v_1, v_2, \ldots, v_n$ has zero volume if and only if the matrix with rows $v_1, v_2, \ldots, v_n$ has zero determinant.

4.3.2 Determinants and Volumes

The key observation above is only the beginning of the story: the volume of a parallelepiped is *always* a determinant.

Theorem (Determinants and volumes)**.** *Let $v_1, v_2, \ldots, v_n$ be vectors in $\mathbf{R}^n$, let P be the parallelepiped determined by these vectors, and let A be the matrix with rows $v_1, v_2, \ldots, v_n$. Then the absolute value of the determinant of A is the volume of P:*

$$|\det(A)| = \mathrm{vol}(P).$$

Proof. Since the four defining properties characterize the determinant, they also characterize the absolute value of the determinant. Explicitly, $|\det|$ is a function on square matrices which satisfies these properties:

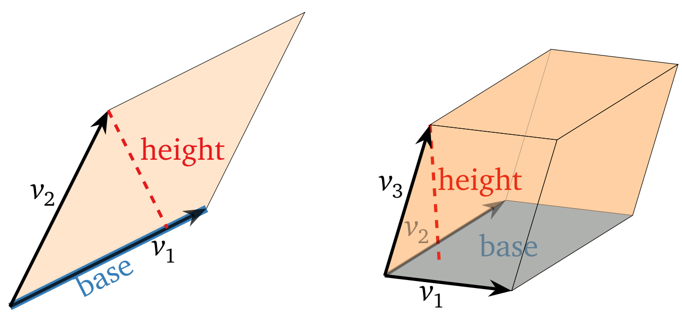
1. Doing a row replacement on A does not change $|\det(A)|$.
2. Scaling a row of A by a scalar c multiplies $|\det(A)|$ by $|c|$.
3. Swapping two rows of a matrix does not change $|\det(A)|$.
4. The determinant of the identity matrix I_n is equal to 1.

The absolute value of the determinant is the *only* such function: indeed, by this recipe in Section 4.1, if you do some number of row operations on A to obtain a matrix B in row echelon form, then

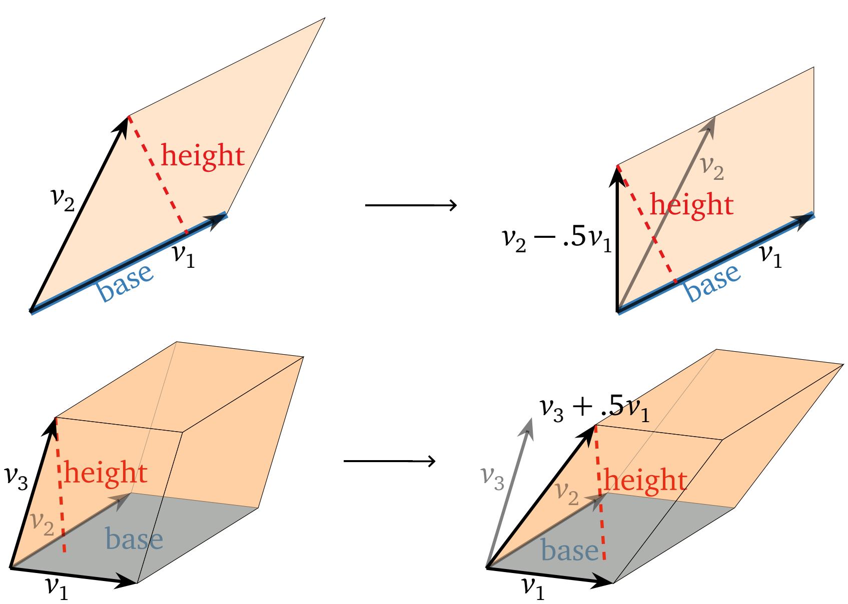
$$|\det(A)| = \left|\frac{(\text{product of the diagonal entries of } B)}{(\text{product of scaling factors used})}\right|.$$

For a square matrix A, we abuse notation and let $\text{vol}(A)$ denote the volume of the paralellepiped determined by the rows of A. Then we can regard vol as a function from the set of square matrices to the real numbers. We will show that vol also satisfies the above four properties.

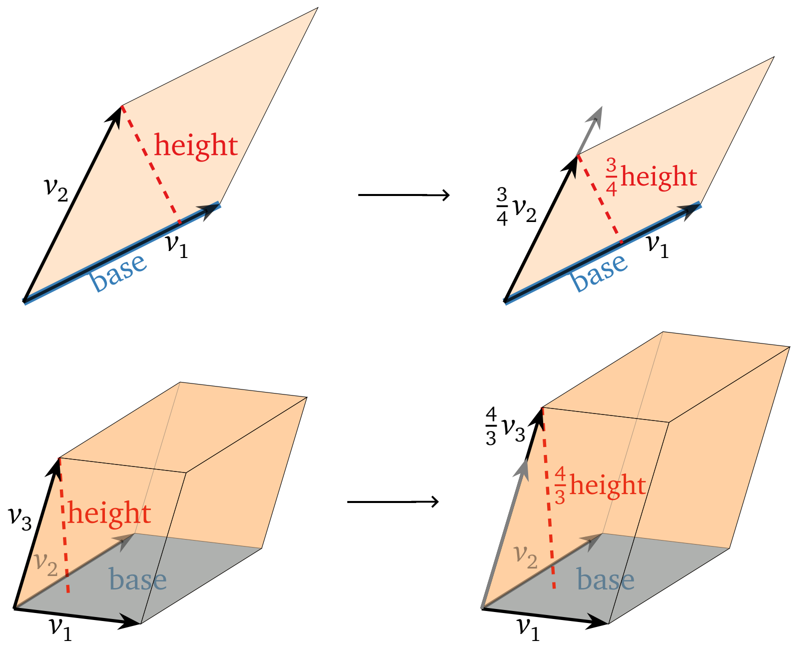
1. For simplicity, we consider a row replacement of the form $R_n = R_n + cR_i$. The volume of a paralellepiped is the volume of its base, times its height: here the "base" is the paralellepiped determined by $v_1, v_2, \ldots, v_{n-1}$, and the "height" is the perpendicular distance of v_n from the base.



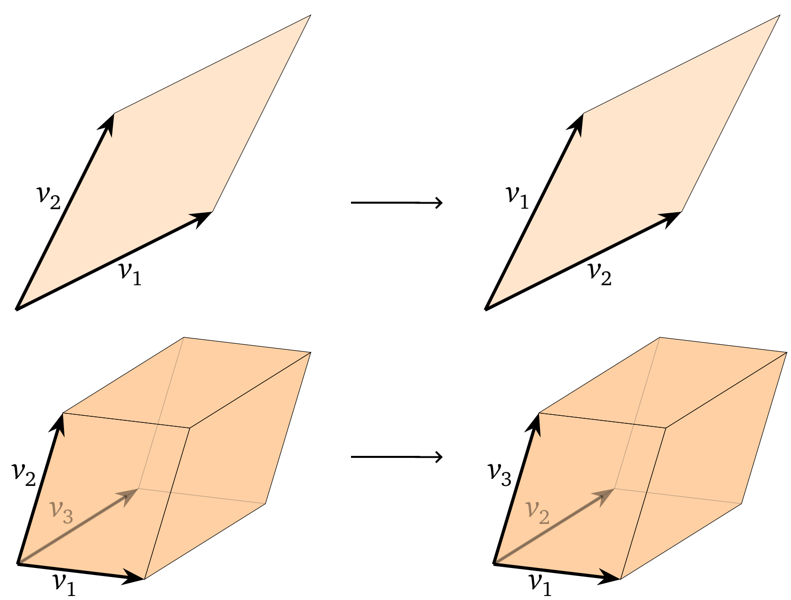
Translating v_n by a multiple of v_i moves v_n in a direction parallel to the base. This changes neither the base nor the height! Thus, $\text{vol}(A)$ is unchanged by row replacements.



2. For simplicity, we consider a row scale of the form $R_n = cR_n$. This scales the length of v_n by a factor of $|c|$, which also scales the perpendicular distance of v_n from the base by a factor of $|c|$. Thus, $\mathrm{vol}(A)$ is scaled by $|c|$.



3. Swapping two rows of A just reorders the vectors $v_1, v_2, \ldots, v_n$, hence has no effect on the parallelepiped determined by those vectors. Thus, $\mathrm{vol}(A)$ is unchanged by row swaps.



4. The rows of the identity matrix I_n are the standard coordinate vectors $e_1, e_2, \ldots, e_n$. The associated paralellepiped is the unit cube, which has volume 1. Thus, $\text{vol}(I_n) = 1$.

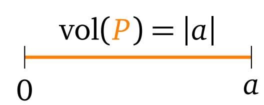
Since $|\det|$ is the only function satisfying these properties, we have

$$\text{vol}(P) = \text{vol}(A) = |\det(A)|.$$

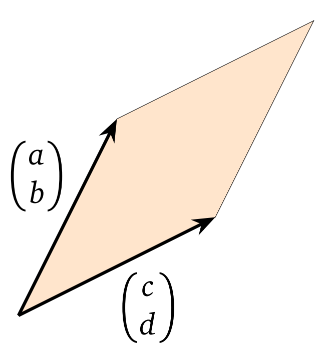
This completes the proof. □

Since $\det(A) = \det(A^T)$ by the transpose property, the absolute value of $\det(A)$ is also equal to the volume of the paralellepiped determined by the *columns* of A as well.

Example (Length)**.** A 1×1 matrix A is just a number $\begin{pmatrix} a \end{pmatrix}$. In this case, the parallelepiped P determined by its one row is just the interval $[0, a]$ (or $[a, 0]$ if $a < 0$). The "volume" of a region in $\mathbf{R}^1 = \mathbf{R}$ is just its length, so it is clear in this case that $\text{vol}(P) = |a|$.



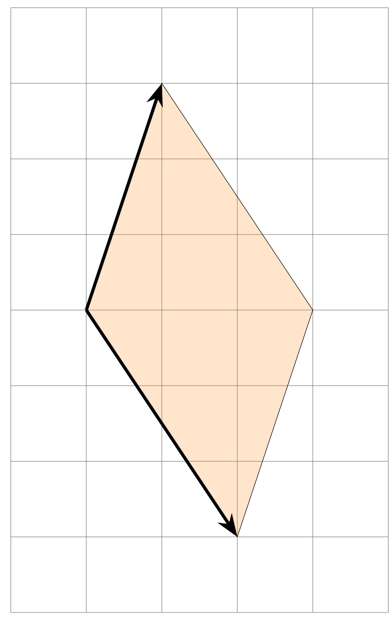
Example (Area)**.** When A is a 2×2 matrix, its rows determine a parallelogram in $\mathbf{R}^2$. The "volume" of a region in $\mathbf{R}^2$ is its area, so we obtain a formula for the area of a parallelogram: it is the determinant of the matrix whose rows are the vectors forming two adjacent sides of the parallelogram.



$$\text{area} = \left|\det\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right| = |ad - bc|$$

It is perhaps surprising that it is possible to compute the area of a parallelogram without trigonometry. It is a fun geometry problem to prove this formula by hand. [Hint: first think about the case when the first row of A lies on the x-axis.]

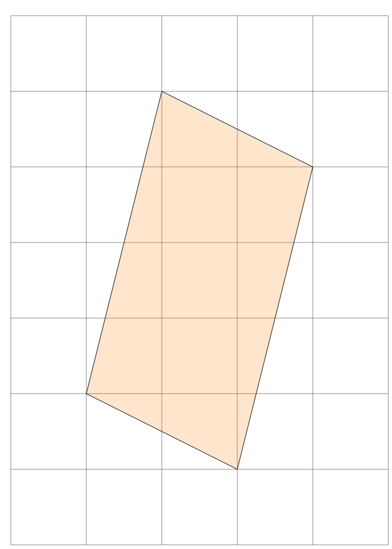
Example. Find the area of the parallelogram with sides $(1,3)$ and $(2,-3)$.



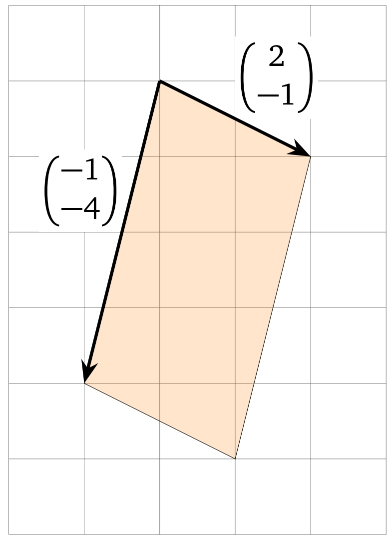
Solution. The area is

$$\left|\det\begin{pmatrix} 1 & 3 \\ 2 & -3 \end{pmatrix}\right| = |-3-6| = 9.$$

Example. Find the area of the parallelogram in the picture.



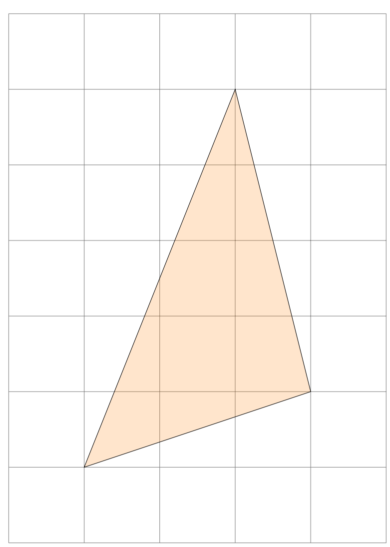
Solution. We choose two adjacent sides to be the rows of a matrix. We choose the top two:



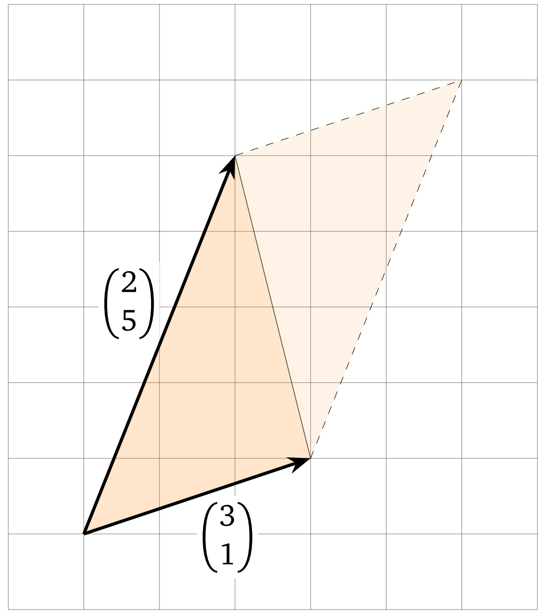
Note that we do not need to know where the origin is in the picture: vectors are determined by their length and direction, not where they start. The area is

$$\left|\det\begin{pmatrix}-1 & -4\\ 2 & -1\end{pmatrix}\right| = |1+8| = 9.$$

Example (Area of a triangle)**.** Find the area of the triangle with vertices $(-1,-2)$, $(2,-1)$, $(1,3)$.



Solution. Doubling a triangle makes a paralellogram. We choose two of its sides to be the rows of a matrix.



The area of the parallelogram is

$$\left|\det\begin{pmatrix}2 & 5\\3 & 1\end{pmatrix}\right| = |2-15| = 13,$$

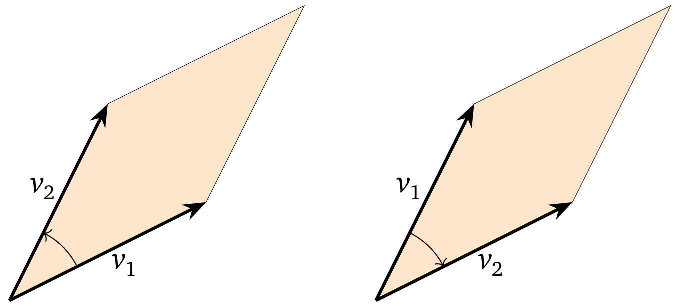
so the area of the triangle is $13/2$.

You might be wondering: if the absolute value of the determinant is a volume, what is the geometric meaning of the determinant without the absolute value? The next remark explains that we can think of the determinant as a *signed* volume. If you have taken an integral calculus course, you probably computed negative areas under curves; the idea here is similar.

Remark (Signed volumes)**.** The theorem on determinants and volumes tells us that the *absolute value* of the determinant is the volume of a paralellepiped. This raises the question of whether the sign of the determinant has any geometric meaning.

A 1×1 matrix A is just a number $\begin{pmatrix} a \end{pmatrix}$. In this case, the parallelepiped P determined by its one row is just the interval $[0,a]$ if $a\geq 0$, and it is $[a,0]$ if $a<0$. In this case, the sign of the determinant determines whether the interval is to the left or the right of the origin.

For a 2×2 matrix with rows v_1, v_2, the sign of the determinant determines whether v_2 is counterclockwise or clockwise from v_1. That is, if the counterclockwise angle from v_1 to v_2 is less than $180°$, then the determinant is positive; otherwise it is negative (or zero).



$$\det\begin{pmatrix} — v_1 — \\ — v_2 — \end{pmatrix} > 0 \qquad\qquad \det\begin{pmatrix} — v_1 — \\ — v_2 — \end{pmatrix} < 0$$

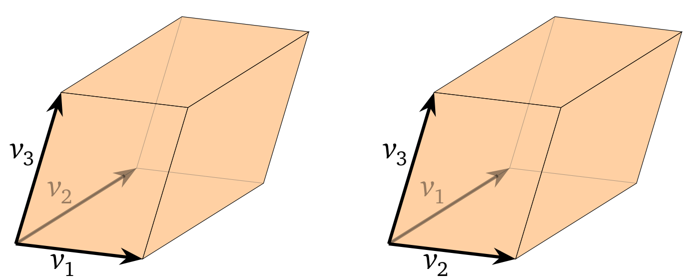
For example, if $v_1 = \binom{a}{b}$, then the counterclockwise rotation of v_1 by 90° is $v_2 = \binom{-b}{a}$ by this example in Section 3.3, and

$$\det\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = a^2 + b^2 > 0.$$

On the other hand, the *clockwise* rotation of v_1 by 90° is $\binom{b}{-a}$, and

$$\det\begin{pmatrix} a & b \\ b & -a \end{pmatrix} = -a^2 - b^2 < 0.$$

For a 3×3 matrix with rows v_1, v_2, v_3, the *right-hand rule* determines the sign of the determinant. If you point the index finger of your right hand in the direction of v_1 and your middle finger in the direction of v_2, then the determinant is positive if your thumb points roughly in the direction of v_3, and it is negative otherwise.

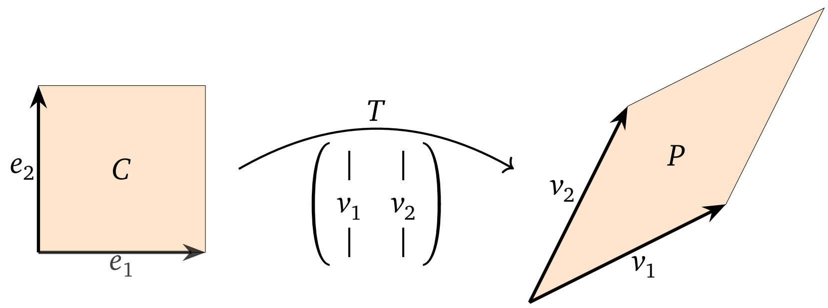


$$\det\begin{pmatrix} — v_1 — \\ — v_2 — \\ — v_3 — \end{pmatrix} > 0 \qquad\qquad \det\begin{pmatrix} — v_1 — \\ — v_2 — \\ — v_3 — \end{pmatrix} < 0$$

In higher dimensions, the notion of signed volume is still important, but it is usually *defined* in terms of the sign of a determinant.

4.3.3 Volumes of Regions

Let A be an $n\times n$ matrix with columns $v_1, v_2, \ldots, v_n$, and let $T\colon \mathbf{R}^n \to \mathbf{R}^n$ be the associated matrix transformation $T(x) = Ax$. Then $T(e_1) = v_1$ and $T(e_2) = v_2$, so T takes the unit cube C to the parallelepiped P determined by $v_1, v_2, \ldots, v_n$:



Since the unit cube has volume 1 and its image has volume $|\det(A)|$, the transformation T scaled the volume of the cube by a factor of $|\det(A)|$. To rephrase:

If A is an $n\times n$ matrix with corresponding matrix transformation $T\colon \mathbf{R}^n \to \mathbf{R}^n$, and if C is the unit cube in $\mathbf{R}^n$, then the volume of $T(C)$ is $|\det(A)|$.

The notation $T(S)$ means the image of the region S under the transformation T. In set builder notation, this is the subset

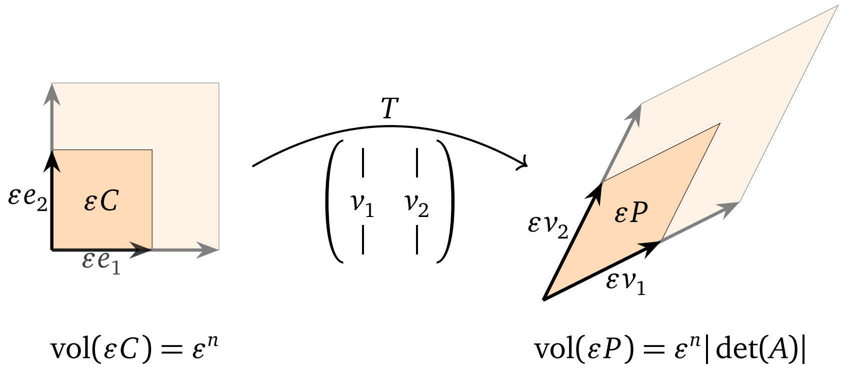
$$T(S) = \bigl\{T(x) \mid x \text{ in } S\bigr\}.$$

In fact, T scales the volume of *any* region in $\mathbf{R}^n$ by the same factor, even for curvy regions.

Theorem. *Let A be an $n\times n$ matrix, and let $T\colon \mathbf{R}^n \to \mathbf{R}^n$ be the associated matrix transformation $T(x) = Ax$. If S is any region in $\mathbf{R}^n$, then*

$$\mathrm{vol}(T(S)) = |\det(A)|\cdot \mathrm{vol}(S).$$

Proof. Let C be the unit cube, let $v_1, v_2, \ldots, v_n$ be the columns of A, and let P be the paralellepiped determined by these vectors, so $T(C) = P$ and $\mathrm{vol}(P) = |\det(A)|$. For $\varepsilon > 0$ we let εC be the cube with side lengths ε, i.e., the paralellepiped determined by the vectors $\varepsilon e_1, \varepsilon e_2, \ldots, \varepsilon e_n$, and we define εP similarly. By the second defining property, T takes εC to εP. The volume of εC is ε^n (we scaled each of the n standard vectors by a factor of ε) and the volume of εP is $\varepsilon^n|\det(A)|$ (for the same reason), so we have shown that T scales the volume of εC by $|\det(A)|$.

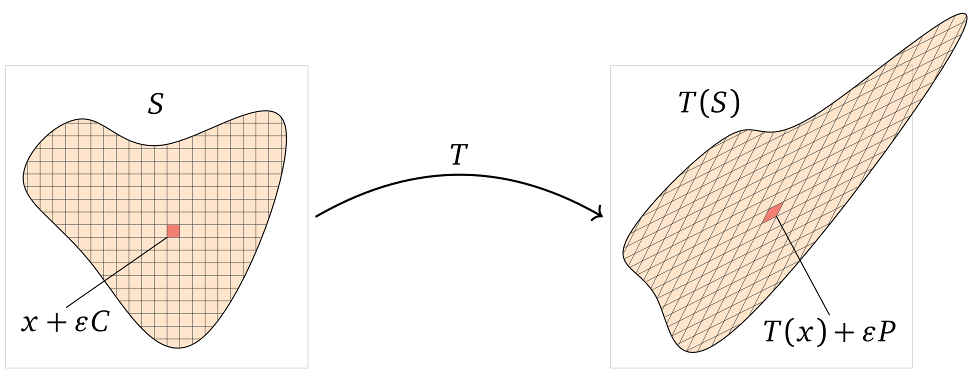


By the first defining property, the image of a translate of εC is a translate of εP:

$$T(x + \varepsilon C) = T(x) + \varepsilon T(C) = T(x) + \varepsilon P.$$

Since a translation does not change volumes, this proves that T scales the volume of a translate of εC by $|\det(A)|$.

At this point, we need to use techniques from multivariable calculus, so we only give an idea of the rest of the proof. Any region S can be approximated by a collection of very small cubes of the form $x + \varepsilon C$. The image $T(S)$ is then approximated by the image of this collection of cubes, which is a collection of very small paralellepipeds of the form $T(x) + \varepsilon P$.

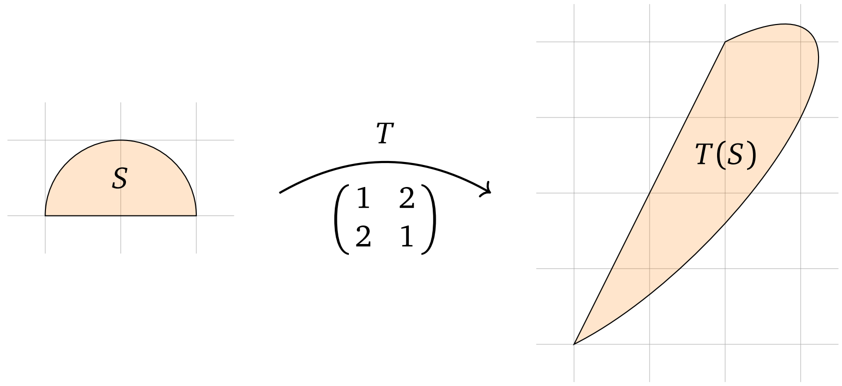


The volume of S is closely approximated by the sum of the volumes of the cubes; in fact, as ε goes to zero, the limit of this sum is precisely $\mathrm{vol}(S)$. Likewise, the volume of $T(S)$ is equal to the sum of the volumes of the paralellepipeds, take in the limit as $\varepsilon \to 0$. The key point is that *the volume of each cube is scaled by* $|\det(A)|$. Therefore, the sum of the volumes of the paralellepipeds is $|\det(A)|$ times the sum of the volumes of the cubes. This proves that $\mathrm{vol}(T(S)) = |\det(A)|\mathrm{vol}(S)$. □

Example. Let S be a half-circle of radius 1, let

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

and define $T: \mathbf{R}^2 \to \mathbf{R}^2$ by $T(x) = Ax$. What is the area of $T(S)$?



Solution. The area of the unit circle is π, so the area of S is $\pi/2$. The transformation T scales areas by a factor of $|\det(A)| = |1-4| = 3$, so

$$\mathrm{vol}(T(S)) = 3\,\mathrm{vol}(S) = \frac{3\pi}{2}.$$

Example (Area of an ellipse)**.** Find the area of the interior E of the ellipse defined by the equation

$$\left(\frac{2x-y}{2}\right)^2 + \left(\frac{y+3x}{3}\right)^2 = 1.$$

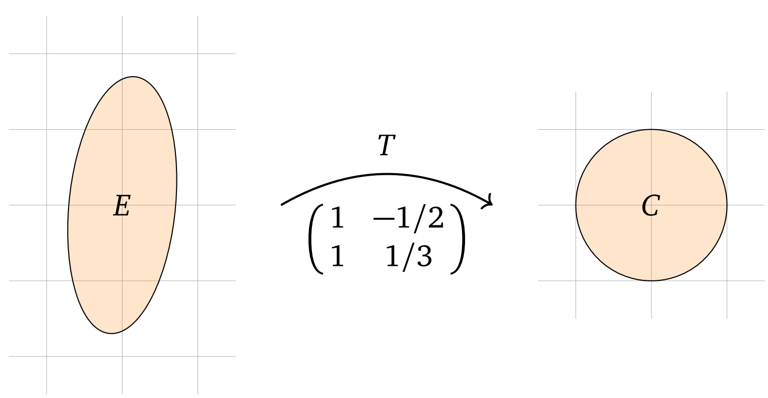
Solution. This ellipse is obtained from the unit circle $X^2 + Y^2 = 1$ by the linear change of coordinates

$$X = \frac{2x-y}{2}$$
$$Y = \frac{y+3x}{3}.$$

In other words, if we define a linear transformation $T: \mathbf{R}^2 \to \mathbf{R}^2$ by

$$T\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}(2x-y)/2\\(y+3x)/3\end{pmatrix},$$

then $T\binom{x}{y}$ lies on the unit circle C whenever $\binom{x}{y}$ lies on E.



We compute the standard matrix A for T by evaluating on the standard coordinate vectors:

$$T\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix} \qquad T\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}-1/2\\1/3\end{pmatrix} \quad\implies\quad A=\begin{pmatrix}1 & -1/2\\1 & 1/3\end{pmatrix}.$$

Therefore, T scales areas by a factor of $|\det(A)| = |\frac{1}{3}+\frac{1}{2}| = \frac{5}{6}$. The area of the unit circle is π, so

$$\pi = \text{vol}(C) = \text{vol}(T(E)) = |\det(A)| \cdot \text{vol}(E) = \frac{5}{6}\text{vol}(E),$$

and thus the area of E is $6\pi/5$.

Remark (Multiplicativity of $|\det|$)**.** The above theorem also gives a geometric reason for multiplicativity of the (absolute value of the) determinant. Indeed, let A and B be $n\times n$ matrices, and let $T,U\colon \mathbf{R}^n \to \mathbf{R}^n$ be the corresponding matrix transformations. If C is the unit cube, then

$$\begin{aligned}\text{vol}\big(T\circ U(C)\big) = \text{vol}\big(T(U(C))\big) &= |\det(A)|\,\text{vol}(U(C))\\ &= |\det(A)|\cdot|\det(B)|\,\text{vol}(C)\\ &= |\det(A)|\cdot|\det(B)|.\end{aligned}$$

On the other hand, the matrix for the composition $T\circ U$ is the product AB, so

$$\text{vol}\big(T\circ U(C)\big) = |\det(AB)|\,\text{vol}(C) = |\det(AB)|.$$

Thus $|\det(AB)| = |\det(A)|\cdot|\det(B)|$.

Chapter 5

Eigenvalues and Eigenvectors

Primary Goal. Solve the matrix equation $Ax = \lambda x$.

This chapter constitutes the core of any first course on linear algebra: eigenvalues and eigenvectors play a crucial role in most real-world applications of the subject.

Example. 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 baby rabbits in their first, second, and third years, respectively.

What is the *asymptotic* behavior of this system? What will the rabbit population look like in 100 years?

Use this link to view the online demo

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. What do you notice about the long-term behavior of the ratios? This phenomenon turns out to be due to eigenvectors.

In Section 5.1, we will define eigenvalues and eigenvectors, and show how to compute the latter; in Section 5.2 we will learn to compute the former. In Section 5.3 we introduce the notion of *similar* matrices, and demonstrate that similar matrices do indeed behave similarly. In Section 5.4 we study matrices that are similar to diagonal matrices and in Section 5.5 we study matrices that are similar to rotation-scaling matrices, thus gaining a solid geometric understanding of large classes of matrices. Finally, we spend Section 5.6 presenting a common kind of application of eigenvalues and eigenvectors to real-world problems, including searching the Internet using Google's PageRank algorithm.

5.1 Eigenvalues and Eigenvectors

Objectives

1. Learn the definition of eigenvector and eigenvalue.
2. Learn to find eigenvectors and eigenvalues geometrically.
3. Learn to decide if a number is an eigenvalue of a matrix, and if so, how to find an associated eigenvector.
4. *Recipe:* find a basis for the λ-eigenspace.
5. *Pictures:* whether or not a vector is an eigenvector, eigenvectors of standard matrix transformations.
6. *Theorem:* the expanded invertible matrix theorem.
7. *Vocabulary word:* **eigenspace**.
8. *Essential vocabulary words:* **eigenvector**, **eigenvalue**.

In this section, we define eigenvalues and eigenvectors. These form the most important facet of the structure theory of square matrices. As such, eigenvalues and eigenvectors tend to play a key role in the real-life applications of linear algebra.

5.1.1 Eigenvalues and Eigenvectors

Here is the most important definition in this text.

Essential Definition. Let A be an $n \times n$ matrix.

1. An **eigenvector** of A is a *nonzero* vector v in $\mathbf{R}^n$ such that $Av = \lambda v$, for some scalar λ.
2. An **eigenvalue** of A is a scalar λ such that the equation $Av = \lambda v$ has a *nontrivial* solution.

If $Av = \lambda v$ for $v \neq 0$, we say that λ is the **eigenvalue for** v, and that v is an **eigenvector for** λ.

The German prefix "eigen" roughly translates to "self" or "own". An eigenvector of A is a vector that is taken to a multiple of itself by the matrix transformation $T(x) = Ax$, which perhaps explains the terminology. On the other hand, "eigen" is often translated as "characteristic"; we may think of an eigenvector as describing an intrinsic, or characteristic, property of A.

Note. Eigenvalues and eigenvectors are only for square matrices.

Eigenvectors are *by definition nonzero*. Eigenvalues may be equal to zero.

We do not consider the zero vector to be an eigenvector: since $A0 = 0 = \lambda 0$ for *every* scalar λ, the associated eigenvalue would be undefined.

If someone hands you a matrix A and a vector v, it is easy to check if v is an eigenvector of A: simply multiply v by A and see if Av is a scalar multiple of v. On the other hand, given just the matrix A, it is not obvious at all how to find the eigenvectors. We will learn how to do this in Section 5.2.

Example (Verifying eigenvectors)**.** Consider the matrix

$$A = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix} \quad \text{and vectors} \quad v = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad w = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Which are eigenvectors? What are their eigenvalues?

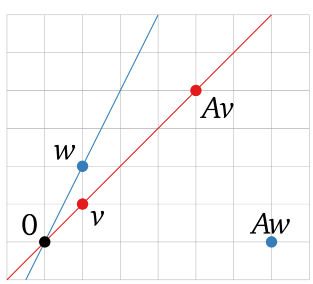
Solution. We have

$$Av = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4v.$$

Hence, v is an eigenvector of A, with eigenvalue $\lambda = 4$. On the other hand,

$$Aw = \begin{pmatrix} 2 & 2 \\ -4 & 8 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix}.$$

which is not a scalar multiple of w. Hence, w is not an eigenvector of A.



v is an eigenvector

w is not an eigenvector

Example (Verifying eigenvectors)**.** Consider the matrix

$$A = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \quad \text{and vectors} \quad v = \begin{pmatrix} 16 \\ 4 \\ 1 \end{pmatrix} \qquad w = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}.$$

Which are eigenvectors? What are their eigenvalues?

Solution. We have

$$Av = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 16 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 32 \\ 8 \\ 2 \end{pmatrix} = 2v.$$

Hence, v is an eigenvector of A, with eigenvalue $\lambda = 2$. On the other hand,

$$Aw = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 28 \\ 1 \\ 1 \end{pmatrix},$$

which is not a scalar multiple of w. Hence, w is not an eigenvector of A.

Example (An eigenvector with eigenvalue 0). Let

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \qquad v = \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

Is v an eigenvector of A? If so, what is its eigenvalue?

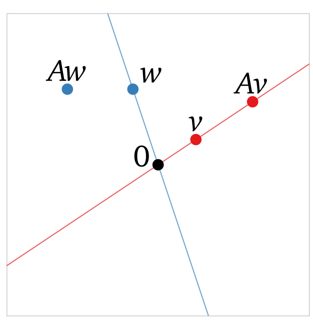
Solution. The product is

$$Av = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0v.$$

Hence, v is an eigenvector with eigenvalue *zero*.

As noted above, an eigen*value* is allowed to be zero, but an eigen*vector* is not.

To say that $Av = \lambda v$ means that Av and λv are *collinear with the origin*. So, an eigenvector of A is a nonzero vector v such that Av and v lie on the same line through the origin. In this case, Av is a scalar multiple of v; the eigenvalue is the scaling factor.



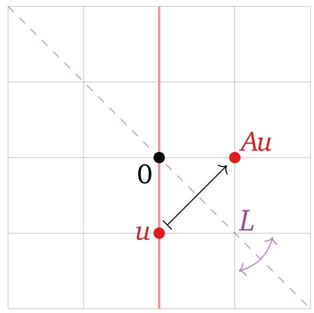
v is an eigenvector

w is not an eigenvector

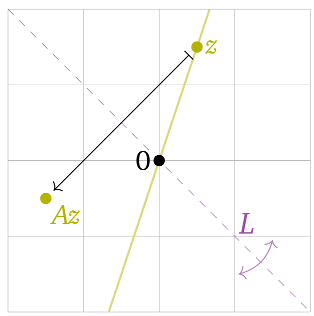
For matrices that arise as the standard matrix of a linear transformation, it is often best to draw a picture, then find the eigenvectors and eigenvalues geometrically by studying which vectors are not moved off of their line. For a transformation that is defined geometrically, it is not necessary even to compute its matrix to find the eigenvectors and eigenvalues.

Example (Reflection). Here is an example of this. Let $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ be the linear transformation that reflects over the line L defined by $y = -x$, and let A be the matrix for T. We will find the eigenvalues and eigenvectors of A without doing any computations.

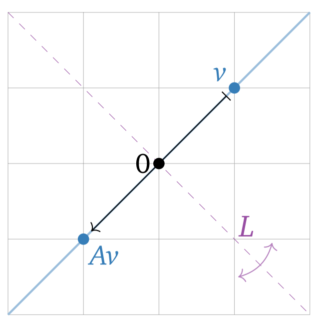
This transformation is defined geometrically, so we draw a picture.



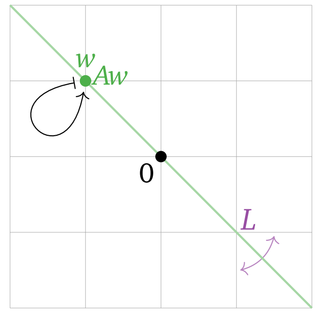
The vector u is not an eigenvector, because Au is not collinear with u and the origin.



The vector z is not an eigenvector either.



The vector v is an eigenvector because Av is collinear with v and the origin. The vector Av has the same length as v, but the opposite direction, so the associated eigenvalue is -1.



The vector w is an eigenvector because Aw is collinear with w and the origin: indeed, Aw is equal to w! This means that w is an eigenvector with eigenvalue 1.

It appears that all eigenvectors lie either on L, or on the line perpendicular to L. The vectors on L have eigenvalue 1, and the vectors perpendicular to L have eigenvalue -1.

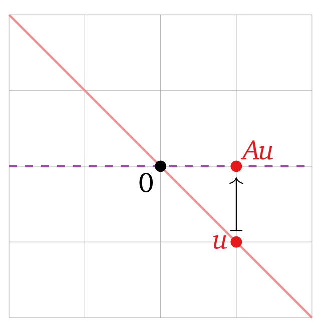
Use this link to view the online demo

An eigenvector of A is a vector x such that Ax is collinear with x and the origin. Click and drag the head of x to convince yourself that all such vectors lie either on L, or on the line perpendicular to L.

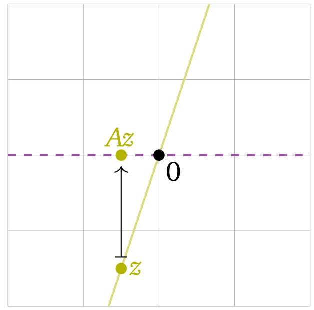
We will now give five more examples of this nature

Example (Projection)**.** Let $T: \mathbf{R}^2 \to \mathbf{R}^2$ be the linear transformation that projects a vector vertically onto the x-axis, and let A be the matrix for T. Find the eigenvalues and eigenvectors of A without doing any computations.

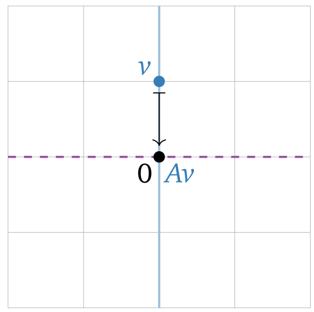
Solution. This transformation is defined geometrically, so we draw a picture.



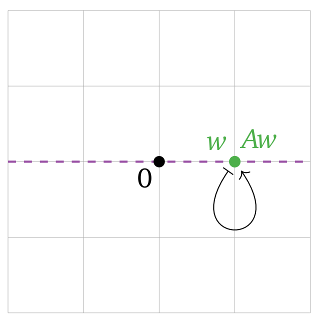
The vector u is not an eigenvector, because Au is not collinear with u and the origin.



The vector z is not an eigenvector either.



The vector v is an eigenvector. Indeed, Av is the zero vector, which is collinear with v and the origin; since $Av = 0v$, the associated eigenvalue is 0.



The vector w is an eigenvector because Aw is collinear with w and the origin: indeed, Aw is equal to w! This means that w is an eigenvector with eigenvalue 1.

It appears that all eigenvectors lie on the x-axis or the y-axis. The vectors on the x-axis have eigenvalue 1, and the vectors on the y-axis have eigenvalue 0.

Use this link to view the online demo

An eigenvector of A is a vector x such that Ax is collinear with x and the origin. Click and drag the head of x to convince yourself that all such vectors lie on the coordinate axes.

Example (Identity)**.** Find all eigenvalues and eigenvectors of the identity matrix I_n.

Solution. The identity matrix has the property that $I_n v = v$ for all vectors v in $\mathbf{R}^n$. We can write this as $I_n v = 1 \cdot v$, so *every* nonzero vector is an eigenvector with eigenvalue 1.

Use this link to view the online demo

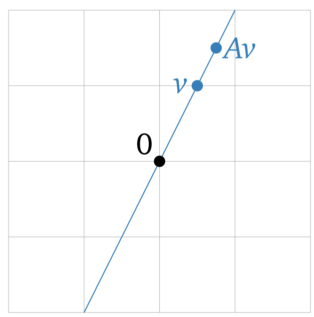
Every nonzero vector is an eigenvector of the identity matrix.

Example (Dilation)**.** Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be the linear transformation that dilates by a factor of 1.5, and let A be the matrix for T. Find the eigenvalues and eigenvectors of A without doing any computations.

Solution. We have

$$Av = T(v) = 1.5v$$

for every vector v in $\mathbf{R}^2$. Therefore, by definition every nonzero vector is an eigenvector with eigenvalue 1.5.



Use this link to view the online demo

Every nonzero vector is an eigenvector of a dilation matrix.

Example (Shear)**.** Let

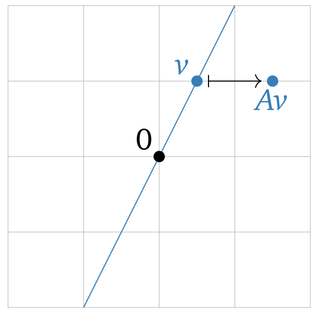
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and let $T(x) = Ax$, so T is a shear in the x-direction. Find the eigenvalues and eigenvectors of A without doing any computations.

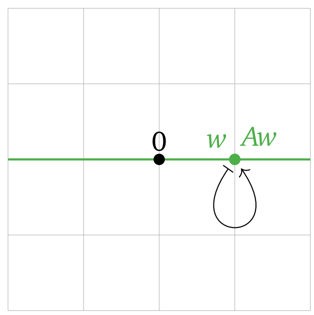
Solution. In equations, we have

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ y \end{pmatrix}.$$

This tells us that a shear takes a vector and adds its y-coordinate to its x-coordinate. Since the x-coordinate changes but not the y-coordinate, this tells us that any vector v with nonzero y-coordinate cannot be collinear with Av and the origin.



On the other hand, any vector v on the x-axis has zero y-coordinate, so it is not moved by A. Hence v is an eigenvector with eigenvalue 1.



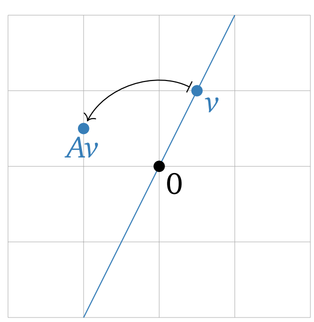
Accordingly, all eigenvectors of A lie on the x-axis, and have eigenvalue 1.

Use this link to view the online demo

All eigenvectors of a shear lie on the x-axis. Click and drag the head of x to find the eigenvectors.

Example (Rotation). Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be the linear transformation that rotates counterclockwise by 90°, and let A be the matrix for T. Find the eigenvalues and eigenvectors of A without doing any computations.

Solution. If v is any nonzero vector, then Av is rotated by an angle of 90° from v. Therefore, Av is not on the same line as v, so v is not an eigenvector. And of course, the zero vector is never an eigenvector.



Therefore, this matrix *has no eigenvectors and eigenvalues.*

Use this link to view the online demo

This rotation matrix has no eigenvectors. Click and drag the head of x to find one.

Here we mention one basic fact about eigenvectors.

Fact (Eigenvectors with distinct eigenvalues are linearly independent)**.** *Let $v_1, v_2, \ldots, v_k$ be eigenvectors of a matrix A, and suppose that the corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ are* distinct *(all different from each other). Then $\{v_1, v_2, \ldots, v_k\}$ is linearly independent.*

Proof. Suppose that $\{v_1, v_2, \ldots, v_k\}$ were linearly *de*pendent. According to the increasing span criterion in Section 2.5, this means that for some j, the vector v_j is in $\text{Span}\{v_1, v_2, \ldots, v_{j-1}\}$. If we choose the *first* such j, then $\{v_1, v_2, \ldots, v_{j-1}\}$ is linearly independent. Note that $j > 1$ since $v_1 \neq 0$.

Since v_j is in $\text{Span}\{v_1, v_2, \ldots, v_{j-1}\}$,, we can write

$$v_j = c_1 v_1 + c_2 v_2 + \cdots + c_{j-1} v_{j-1}$$

for some scalars $c_1, c_2, \ldots, c_{j-1}$. Multiplying both sides of the above equation by A gives

$$\begin{aligned} \lambda_j v_j = A v_j &= A\big(c_1 v_1 + c_2 v_2 + \cdots + c_{j-1} v_{j-1}\big) \\ &= c_1 A v_1 + c_2 A v_2 + \cdots + c_{j-1} A v_{j-1} \\ &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_{j-1} \lambda_{j-1} v_{j-1}. \end{aligned}$$

Subtracting λ_j times the first equation from the second gives

$$0 = \lambda_j v_j - \lambda_j v_j = c_1(\lambda_1 - \lambda_j) v_1 + c_2(\lambda_2 - \lambda_j) v_2 + \cdots + c_{j-1}(\lambda_{j-1} - \lambda_j) v_{j-1}.$$

Since $\lambda_i \neq \lambda_j$ for $i < j$, this is an equation of linear dependence among $v_1, v_2, \ldots, v_{j-1}$, which is impossible because those vectors are linearly independent. Therefore, $\{v_1, v_2, \ldots, v_k\}$ must have been linearly independent after all. □

When $k = 2$, this says that if v_1, v_2 are eigenvectors with eigenvalues $\lambda_1 \neq \lambda_2$, then v_2 is not a multiple of v_1. In fact, any nonzero multiple $c v_1$ of v_1 is also an eigenvector with eigenvalue λ_1:

$$A(c v_1) = c A v_1 = c(\lambda_1 v_1) = \lambda_1 (c v_1).$$

As a consequence of the above fact, we have the following.

An $n \times n$ matrix A has at most n eigenvalues.

5.1.2 Eigenspaces

Suppose that A is a square matrix. We already know how to check if a given vector is an eigenvector of A and in that case to find the eigenvalue. Our next goal is to check if a given real number is an eigenvalue of A and in that case to find all of the corresponding eigenvectors. Again this will be straightforward, but more involved. The only missing piece, then, will be to find the eigenvalues of A; this is the main content of Section 5.2.

Let A be an $n\times n$ matrix, and let λ be a scalar. The eigenvectors with eigenvalue λ, if any, are the nonzero solutions of the equation $Av=\lambda v$. We can rewrite this equation as follows:

$$\begin{aligned} & Av = \lambda v \\ \iff\quad & Av - \lambda v = 0 \\ \iff\quad & Av - \lambda I_n v = 0 \\ \iff\quad & (A-\lambda I_n)v = 0. \end{aligned}$$

Therefore, the eigenvectors of A with eigenvalue λ, if any, are the nontrivial solutions of the matrix equation $(A-\lambda I_n)v=0$, i.e., the nonzero vectors in $\mathrm{Nul}(A-\lambda I_n)$. If this equation has no nontrivial solutions, then λ is not an eigenvector of A.

The above observation is important because it says that *finding the eigenvectors for a given eigenvalue means solving a homogeneous system of equations.* For instance, if

$$A=\begin{pmatrix} 7 & 1 & 3 \\ -3 & 2 & -3 \\ -3 & -2 & -1 \end{pmatrix},$$

then an eigenvector with eigenvalue λ is a nontrivial solution of the matrix equation

$$\begin{pmatrix} 7 & 1 & 3 \\ -3 & 2 & -3 \\ -3 & -2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

This translates to the system of equations

$$\left\{\begin{aligned} 7x + y + 3z &= \lambda x \\ -3x + 2y - 3z &= \lambda y \\ -3x - 2y - z &= \lambda z \end{aligned}\right. \quad\longrightarrow\quad \left\{\begin{aligned} (7-\lambda)x + y + 3z &= 0 \\ -3x + (2-\lambda)y - 3z &= 0 \\ -3x - 2y + (-1-\lambda)z &= 0. \end{aligned}\right.$$

This is the same as the homogeneous matrix equation

$$\begin{pmatrix} 7-\lambda & 1 & 3 \\ -3 & 2-\lambda & -3 \\ -3 & -2 & -1-\lambda \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,$$

i.e., $(A-\lambda I_3)v=0$.

Definition. Let A be an $n \times n$ matrix, and let λ be an eigenvalue of A. The **λ-eigenspace** of A is the solution set of $(A-\lambda I_n)v = 0$, i.e., the subspace $\mathrm{Nul}(A-\lambda I_n)$.

The λ-eigenspace is a subspace because it is the null space of a matrix, namely, the matrix $A-\lambda I_n$. This subspace consists of the zero vector and all eigenvectors of A with eigenvalue λ.

Note. Since a nonzero subspace is infinite, *every eigenvalue has infinitely many eigenvectors*. (For example, multiplying an eigenvector by a nonzero scalar gives another eigenvector.) On the other hand, there can be at most n linearly independent eigenvectors of an $n \times n$ matrix, since $\mathbf{R}^n$ has dimension n.

Example (Computing eigenspaces)**.** For each of the numbers $\lambda = -2, 1, 3$, decide if λ is an eigenvalue of the matrix

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

and if so, compute a basis for the λ-eigenspace.

Solution. The number 3 is an eigenvalue of A if and only if $\mathrm{Nul}(A-3I_2)$ is nonzero. Hence, we have to solve the matrix equation $(A-3I_2)v = 0$. We have

$$A - 3I_2 = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix} - 3\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -4 \\ -1 & -4 \end{pmatrix}.$$

The reduced row echelon form of this matrix is

$$\begin{pmatrix} 1 & 4 \\ 0 & 0 \end{pmatrix} \xrightarrow[\text{form}]{\text{parametric}} \begin{cases} x = -4y \\ y = y \end{cases} \xrightarrow[\text{vector form}]{\text{parametric}} \begin{pmatrix} x \\ y \end{pmatrix} = y\begin{pmatrix} -4 \\ 1 \end{pmatrix}.$$

Since y is a free variable, the null space of $A-3I_2$ is nonzero, so 3 is an eigenvector. A basis for the 3-eigenspace is $\left\{\binom{-4}{1}\right\}$.

Concretely, we have shown that the eigenvectors of A with eigenvalue 3 are exactly the nonzero multiples of $\binom{-4}{1}$. In particular, $\binom{-4}{1}$ is an eigenvector, which we can verify:

$$\begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} -4 \\ 1 \end{pmatrix} = \begin{pmatrix} -12 \\ 3 \end{pmatrix} = 3\begin{pmatrix} -4 \\ 1 \end{pmatrix}.$$

The number 1 is an eigenvalue of A if and only if $\mathrm{Nul}(A-I_2)$ is nonzero. Hence, we have to solve the matrix equation $(A-I_2)v = 0$. We have

$$A - I_2 = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -4 \\ -1 & -2 \end{pmatrix}.$$

This matrix has determinant -6, so it is invertible. By the invertible matrix theorem in Section 3.6, we have $\mathrm{Nul}(A-I_2) = \{0\}$, so 1 is not an eigenvalue.

The eigenvectors of A with eigenvalue -2, if any, are the nonzero solutions of the matrix equation $(A+2I_2)v = 0$. We have

$$A + 2I_2 = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix} + 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -4 \\ -1 & 1 \end{pmatrix}.$$

The reduced row echelon form of this matrix is

$$\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \xrightarrow[\text{form}]{\text{parametric}} \begin{cases} x = y \\ y = y \end{cases} \xrightarrow[\text{vector form}]{\text{parametric}} \begin{pmatrix} x \\ y \end{pmatrix} = y \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Hence there exist eigenvectors with eigenvalue -2, namely, any nonzero multiple of $\binom{1}{1}$. A basis for the -2-eigenspace is $\{\binom{1}{1}\}$.

Use this link to view the online demo

The 3*-eigenspace is the line spanned by* $\binom{-4}{1}$*. This means that* A *scales every vector in that line by a factor of* 3*. Likewise, the* -2*-eigenspace is the line spanned by* $\binom{1}{1}$*. Click and drag the vector* x *around to see how* A *acts on that vector.*

Example (Computing eigenspaces)**.** For each of the numbers $\lambda = 0, \frac{1}{2}, 2$, decide if λ is an eigenvector of the matrix

$$A = \begin{pmatrix} 7/2 & 0 & 3 \\ -3/2 & 2 & -3 \\ -3/2 & 0 & -1 \end{pmatrix},$$

and if so, compute a basis for the λ-eigenspace.

Solution. The number 2 is an eigenvalue of A if and only if $\mathrm{Nul}(A - 2I_3)$ is nonzero. Hence, we have to solve the matrix equation $(A - 2I_3)v = 0$. We have

$$A - 2I_3 = \begin{pmatrix} 7/2 & 0 & 3 \\ -3/2 & 2 & -3 \\ -3/2 & 0 & 1 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3/2 & 0 & 3 \\ -3/2 & 0 & -3 \\ -3/2 & 0 & -3 \end{pmatrix}.$$

The reduced row echelon form of this matrix is

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{form}]{\text{parametric}} \begin{cases} x = -2z \\ y = y \\ z = z \end{cases} \xrightarrow[\text{vector form}]{\text{parametric}} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = y\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

The matrix $A - 2I_3$ has two free variables, so the null space of $A - 2I_3$ is nonzero, and thus 2 is an eigenvector. A basis for the 2-eigenspace is

$$\left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

This is a *plane* in $\mathbf{R}^3$.

The eigenvectors of A with eigenvalue $\frac{1}{2}$, if any, are the nonzero solutions of the matrix equation $(A - \frac{1}{2}I_3)v = 0$. We have

$$A - \frac{1}{2}I_3 = \begin{pmatrix} 7/2 & 0 & 3 \\ -3/2 & 2 & -3 \\ -3/2 & 0 & 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 3 \\ -3/2 & 3/2 & -3 \\ -3/2 & 0 & -3/2 \end{pmatrix}.$$

The reduced row echelon form of this matrix is

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{form}]{\text{parametric}} \begin{cases} x = -z \\ y = z \\ z = z \end{cases} \xrightarrow[\text{vector form}]{\text{parametric}} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Hence there exist eigenvectors with eigenvalue $\frac{1}{2}$, so $\frac{1}{2}$ is an eigenvalue. A basis for the $\frac{1}{2}$-eigenspace is

$$\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

This is a *line* in $\mathbf{R}^3$.

The number 0 is an eigenvalue of A if and only if $\mathrm{Nul}(A - 0I_3) = \mathrm{Nul}(A)$ is nonzero. This is the same as asking whether A is noninvertible, by the invertible matrix theorem in Section 3.6. The determinant of A is $\det(A) = 2 \neq 0$, so A is invertible by the invertibility property in Section 4.1. It follows that 0 is not an eigenvalue of A.

Use this link to view the online demo

The 2-eigenspace is the violet plane. This means that A scales every vector in that plane by a factor of 2. The $\frac{1}{2}$-eigenspace is the green line. Click and drag the vector x around to see how A acts on that vector.

Example (Reflection)**.** Let $T : \mathbf{R}^2 \to \mathbf{R}^2$ be the linear transformation that reflects over the line L defined by $y = -x$, and let A be the matrix for T. Find all eigenspaces of A.

Solution. We showed in this example that all eigenvectors with eigenvalue 1 lie on L, and all eigenvectors with eigenvalue -1 lie on the line $L^\perp$ that is perpendicular to L. Hence, L is the 1-eigenspace, and $L^\perp$ is the -1-eigenspace.

None of this required any computations, but we can verify our conclusions using algebra. First we compute the matrix A:

$$T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \qquad T\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \qquad \Longrightarrow \qquad A = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

Computing the 1-eigenspace means solving the matrix equation $(A - I_2)v = 0$. We have

$$A - I_2 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

The parametric form of the solution set is $x = -y$, or equivalently, $y = -x$, which is exactly the equation for L. Computing the -1-eigenspace means solving the matrix equation $(A + I_2)v = 0$; we have

$$A + I_2 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}.$$

The parametric form of the solution set is $x = y$, or equivalently, $y = x$, which is exactly the equation for $L^\perp$.

Use this link to view the online demo

The violet line L is the 1-eigenspace, and the green line $L^\perp$ is the -1-eigenspace.

Recipes: Eigenspaces. Let A be an $n \times n$ matrix and let λ be a number.

1. λ is an eigenvalue of A if and only if $(A - \lambda I_n)v = 0$ has a nontrivial solution, if and only if $\mathrm{Nul}(A - \lambda I_n) \neq \{0\}$.

2. In this case, finding a basis for the λ-eigenspace of A means finding a basis for $\mathrm{Nul}(A - \lambda I_n)$, which can be done by finding the parametric vector form of the solutions of the homogeneous system of equations $(A - \lambda I_n)v = 0$.

3. The dimension of the λ-eigenspace of A is equal to the number of free variables in the system of equations $(A - \lambda I_n)v = 0$, which is the number of columns of $A - \lambda I_n$ without pivots.

4. The eigenvectors with eigenvalue λ are the nonzero vectors in $\mathrm{Nul}(A - \lambda I_n)$, or equivalently, the nontrivial solutions of $(A - \lambda I_n)v = 0$.

We conclude with an observation about the 0-eigenspace of a matrix.

Fact. *Let A be an $n \times n$ matrix.*

1. *The number 0 is an eigenvalue of A if and only if A is not invertible.*

2. *In this case, the 0-eigenspace of A is $\mathrm{Nul}(A)$.*

Proof. We know that 0 is an eigenvalue of A if and only if $\mathrm{Nul}(A - 0I_n) = \mathrm{Nul}(A)$ is nonzero, which is equivalent to the noninvertibility of A by the invertible matrix theorem in Section 3.6. In this case, the 0-eigenspace is by definition $\mathrm{Nul}(A–0I_n) = \mathrm{Nul}(A)$. □

Concretely, an eigenvector with eigenvalue 0 is a nonzero vector v such that $Av = 0v$, i.e., such that $Av = 0$. These are exactly the nonzero vectors in the null space of A.

5.1.3 The Invertible Matrix Theorem: Addenda

We now have two new ways of saying that a matrix is invertible, so we add them to the invertible matrix theorem.

Invertible Matrix Theorem. *Let A be an $n \times n$ matrix, and let $T : \mathbf{R}^n \to \mathbf{R}^n$ be the matrix transformation $T(x) = Ax$. The following statements are equivalent:*

1. *A is invertible.*
2. *A has n pivots.*
3. $\mathrm{Nul}(A) = \{0\}$.
4. *The columns of A are linearly independent.*
5. *The columns of A span $\mathbf{R}^n$.*
6. *$Ax = b$ has a unique solution for each b in $\mathbf{R}^n$.*
7. *T is invertible.*
8. *T is one-to-one.*
9. *T is onto.*
10. $\det(A) \neq 0$.
11. 0 *is not an eigenvalue of A.*

5.2 The Characteristic Polynomial

Objectives

1. Learn that the eigenvalues of a triangular matrix are the diagonal entries.
2. Find all eigenvalues of a matrix using the characteristic polynomial.
3. Learn some strategies for finding the zeros of a polynomial.
4. *Recipe:* the characteristic polynomial of a 2×2 matrix.
5. *Vocabulary words:* **characteristic polynomial**, **trace**.

In Section 5.1 we discussed how to decide whether a given number λ is an eigenvalue of a matrix, and if so, how to find all of the associated eigenvectors. In this section, we will give a method for computing all of the *eigenvalues* of a matrix. This does not reduce to solving a system of linear equations: indeed, it requires solving a *nonlinear* equation in one variable, namely, finding the roots of the characteristic polynomial.

Definition. Let A be an $n \times n$ matrix. The **characteristic polynomial** of A is the function $f(\lambda)$ given by

$$f(\lambda) = \det(A - \lambda I_n).$$

We will see below that the characteristic polynomial is in fact a polynomial. Finding the characterestic polynomial means computing the determinant of the matrix $A - \lambda I_n$, whose entries contain the unknown λ.

Example. Find the characteristic polynomial of the matrix

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}.$$

Solution. We have

$$\begin{aligned} f(\lambda) = \det(A - \lambda I_2) &= \det\left(\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) \\ &= \det\begin{pmatrix} 5-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} \\ &= (5-\lambda)(1-\lambda) - 2\cdot 2 = \lambda^2 - 6\lambda + 1. \end{aligned}$$

Example. Find the characteristic polynomial of the matrix

$$A = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Solution. We compute the determinant by expanding cofactors along the third column:

$$\begin{aligned} f(\lambda) = \det(A - \lambda I_3) &= \det\begin{pmatrix} -\lambda & 6 & 8 \\ \frac{1}{2} & -\lambda & 0 \\ 0 & \frac{1}{2} & -\lambda \end{pmatrix} \\ &= 8\left(\frac{1}{4} - 0\cdot -\lambda\right) - \lambda\left(\lambda^2 - 6\cdot\frac{1}{2}\right) \\ &= -\lambda^3 + 3\lambda + 2. \end{aligned}$$

The point of the characteristic polynomial is that we can use it to compute eigenvalues.

Theorem (Eigenvalues are roots of the characteristic polynomial)**.** *Let A be an $n\times n$ matrix, and let $f(\lambda) = \det(A - \lambda I_n)$ be its characteristic polynomial. Then a number λ_0 is an eigenvalue of A if and only if $f(\lambda_0) = 0$.*

Proof. By the invertible matrix theorem in Section 5.1, the matrix equation $(A - \lambda_0 I_n)x = 0$ has a nontrivial solution if and only if $\det(A - \lambda_0 I_n) = 0$. Therefore,

$$\begin{aligned} \lambda_0 \text{ is an eigenvalue of } A &\iff Ax = \lambda_0 x \text{ has a nontrivial solution} \\ &\iff (A - \lambda_0 I_n)x = 0 \text{ has a nontrivial solution} \\ &\iff A - \lambda_0 I_n \text{ is not invertible} \\ &\iff \det(A - \lambda_0 I_n) = 0 \\ &\iff f(\lambda_0) = 0. \end{aligned}$$

□

Example (Finding eigenvalues)**.** Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}.$$

Solution. In the above example we computed the characteristic polynomial of A to be $f(\lambda) = \lambda^2 - 6\lambda + 1$. We can solve the equation $\lambda^2 - 6\lambda + 1 = 0$ using the quadratic formula:

$$\lambda = \frac{6 \pm \sqrt{36-4}}{2} = 3 \pm 2\sqrt{2}.$$

Therefore, the eigenvalues are $3 + 2\sqrt{2}$ and $3 - 2\sqrt{2}$.

To compute the eigenvectors, we solve the homogeneous system of equations $(A - \lambda I_2)x = 0$ for each eigenvalue λ. When $\lambda = 3 + 2\sqrt{2}$, we have

$$\begin{aligned} A - (3+\sqrt{2})I_2 &= \begin{pmatrix} 2-2\sqrt{2} & 2 \\ 2 & -2-2\sqrt{2} \end{pmatrix} \\ &\xrightarrow{R_1 = R_1 \times (2+2\sqrt{2})} \begin{pmatrix} -4 & 4+4\sqrt{2} \\ 2 & -2-2\sqrt{2} \end{pmatrix} \\ &\xrightarrow{R_2 = R_2 + R_1/2} \begin{pmatrix} -4 & 4+4\sqrt{2} \\ 0 & 0 \end{pmatrix} \\ &\xrightarrow{R_1 = R_1 \div -4} \begin{pmatrix} 1 & -1-\sqrt{2} \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

The parametric form of the general solution is $x = (1 + \sqrt{2})y$, so the $(3 + 2\sqrt{2})$-eigenspace is the line spanned by $\binom{1+\sqrt{2}}{1}$. We compute in the same way that the $(3 - 2\sqrt{2})$-eigenspace is the line spanned by $\binom{1-\sqrt{2}}{1}$.

Use this link to view the online demo

The green line is the $(3 - 2\sqrt{2})$*-eigenspace, and the violet line is the* $(3 + 2\sqrt{2})$*-eigenspace.*

Example (Finding eigenvalues)**.** Find the eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 0 & 6 & 8 \\ \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

Solution. In the above example we computed the characteristic polynomial of A to be $f(\lambda) = -\lambda^3 + 3\lambda + 2$. We eyeball that $f(2) = -8 + 3 \cdot 2 + 2 = 0$. Thus $\lambda - 2$ divides $f(\lambda)$; to find the other roots, we perform polynomial long division:

$$\frac{-\lambda^3 + 3\lambda + 2}{\lambda - 2} = -\lambda^2 - 2\lambda - 1 = -(\lambda + 1)^2.$$

Therefore,

$$f(\lambda) = -(\lambda - 2)(\lambda + 1)^2,$$

so the only eigenvalues are $\lambda = 2, -1$.

We compute the 2-eigenspace by solving the homogeneous system $(A - 2I_3)x = 0$. We have

$$A - 2I_3 = \begin{pmatrix} -2 & 6 & 8 \\ \frac{1}{2} & -2 & 0 \\ 0 & \frac{1}{2} & -2 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -16 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{pmatrix}.$$

The parametric form and parametric vector form of the solutions are:

$$\begin{cases} x = 16z \\ y = 4z \\ z = z \end{cases} \longrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} 16 \\ 4 \\ 1 \end{pmatrix}.$$

Therefore, the 2-eigenspace is the line

$$\text{Span}\left\{ \begin{pmatrix} 16 \\ 4 \\ 1 \end{pmatrix} \right\}.$$

We compute the -1-eigenspace by solving the homogeneous system $(A + I_3)x = 0$. We have

$$A + I_3 = \begin{pmatrix} 1 & 6 & 8 \\ \frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

The parametric form and parametric vector form of the solutions are:

$$\begin{cases} x = 4z \\ y = -2z \\ z = z \end{cases} \longrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} 4 \\ -2 \\ 1 \end{pmatrix}.$$

Therefore, the -1-eigenspace is the line

$$\text{Span}\left\{ \begin{pmatrix} 4 \\ -2 \\ 1 \end{pmatrix} \right\}.$$

Use this link to view the online demo

The green line is the -1-eigenspace, and the violet line is the 2-eigenspace.

Form of the characteristic polynomial It is time that we justified the use of the term "polynomial." First we need a vocabulary word.

Definition. The **trace** of a square matrix A is the number $\mathrm{Tr}(A)$ obtained by summing the diagonal entries of A:

$$\mathrm{Tr}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & a_{nn} \end{pmatrix} = a_{11} + a_{22} + \cdots + a_{nn}.$$

Theorem. *Let A be an $n \times n$ matrix, and let $f(\lambda) = \det(A - \lambda I_n)$ be its characteristic polynomial. Then $f(\lambda)$ is a polynomial of degree n. Moreover, $f(\lambda)$ has the form*

$$f(\lambda) = (-1)^n \lambda^n + (-1)^{n-1}\mathrm{Tr}(A)\lambda^{n-1} + \cdots + \det(A).$$

In other words, the coefficient of λ^{n-1} is $\pm\mathrm{Tr}(A)$, and the constant term is $\det(A)$ (the other coefficients are just numbers without names).

Proof. First we notice that

$$f(0) = \det(A - 0I_n) = \det(A),$$

so that the constant term is always $\det(A)$.

We will prove the rest of the theorem only for 2×2 matrices; the reader is encouraged to complete the proof in general using cofactor expansions. We can write a 2×2 matrix as $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$; then

$$\begin{aligned} f(\lambda) = \det(A - \lambda I_2) = \det\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} &= (a-\lambda)(d-\lambda) - bc \\ &= \lambda^2 - (a+d)\lambda + (ad - bc) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A). \end{aligned}$$

□

Recipe: The characteristic polynomial of a 2×2 matrix. When $n = 2$, the previous theorem tells us all of the coefficients of the characteristic polynomial:

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A).$$

This is generally the fastest way to compute the characteristic polynomial of a 2×2 matrix.

Example. Find the characteristic polynomial of the matrix

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}.$$

Solution. We have

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) = \lambda^2 - (5+1)\lambda + (5\cdot 1 - 2\cdot 2) = \lambda^2 - 6\lambda + 1,$$

as in the above example.

Remark. By the above theorem, the characteristic polynomial of an $n \times n$ matrix is a polynomial of degree n. Since a polynomial of degree n has at most n roots, this gives another proof of the fact that an $n \times n$ matrix has at most n eigenvalues. See this important note in Section 5.1.

Eigenvalues of a triangular matrix It is easy to compute the determinant of an upper- or lower-triangular matrix; this makes it easy to find its eigenvalues as well.

Corollary. *If A is an upper- or lower-triangular matrix, then the eigenvalues of A are its diagonal entries.*

Proof. Suppose for simplicity that A is a 3×3 upper-triangular matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix}.$$

Its characteristic polynomial is

$$f(\lambda) = \det(A - \lambda I_3) = \det\begin{pmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ 0 & a_{22}-\lambda & a_{23} \\ 0 & 0 & a_{33}-\lambda \end{pmatrix}.$$

This is also an upper-triangular matrix, so the determinant is the product of the diagonal entries:

$$f(\lambda) = (a_{11}-\lambda)(a_{22}-\lambda)(a_{33}-\lambda).$$

The zeros of this polynomial are exactly a_{11}, a_{22}, a_{33}. □

Example. Find the eigenvalues of the matrix

$$A = \begin{pmatrix} 1 & 7 & 2 & 4 \\ 0 & 1 & 3 & 11 \\ 0 & 0 & \pi & 101 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Solution. The eigenvalues are the diagonal entries $1, \pi, 0$. (The eigenvalue 1 occurs twice, but it counts as one eigenvalue; in Section 5.4 we will define the notion of *algebraic multiplicity* of an eigenvalue.)

Factoring the characteristic polynomial If A is an $n \times n$ matrix, then the characteristic polynomial $f(\lambda)$ has degree n by the above theorem. When $n = 2$, one can use the quadratic formula to find the roots of $f(\lambda)$. There exist algebraic formulas for the roots of cubic and quartic polynomials, but these are generally too cumbersome to apply by hand. Even worse, it is known that there is no algebraic formula for the roots of a general polynomial of degree at least 5.

In practice, the roots of the characteristic polynomial are found numerically by computer. That said, there do exist methods for finding roots by hand. For instance, we have the following consequence of the rational root theorem (which we also call the rational root theorem):

Rational Root Theorem. *Suppose that A is an $n \times n$ matrix whose characteristic polynomial $f(\lambda)$ has integer (whole-number) entries. Then all rational roots of its characteristic polynomial are integer divisors of* $\det(A)$.

For example, if A has integer entries, then its characteristic polynomial has integer coefficients. This gives us one way to find a root by hand, if A has an eigenvalue that is a rational number. Once we have found one root, then we can reduce the degree by polynomial long division.

Example. Find the eigenvalues of the matrix

$$A = \begin{pmatrix} 7 & 0 & 3 \\ -3 & 2 & -3 \\ -3 & 0 & -1 \end{pmatrix}.$$

Hint: one eigenvalue is an integer.

Solution. We compute the characteristic polynomial by expanding cofactors along the first row:

$$\begin{aligned} f(\lambda) = \det(A - \lambda I_3) &= \det \begin{pmatrix} 7-\lambda & 0 & 3 \\ -3 & 2-\lambda & -3 \\ -3 & 0 & -1-\lambda \end{pmatrix} \\ &= (7-\lambda)(2-\lambda)(-1-\lambda) + 3 \cdot 3(2-\lambda) \\ &= -\lambda^3 + 8\lambda^2 - 14\lambda + 4. \end{aligned}$$

The determinant of A is the constant term $f(0) = 4$; its integer divisors are $\pm 1, \pm 2, \pm 4$. We check which are roots:

$$f(1) = -3 \quad f(-1) = 27 \quad f(2) = 0 \quad f(-2) = 72 \quad f(4) = 12 \quad f(-4) = 252.$$

The only rational root of $f(\lambda)$ is $\lambda = 2$. We divide by $\lambda - 2$ using polynomial long division:

$$\frac{-\lambda^3 + 8\lambda^2 - 14\lambda + 4}{\lambda - 2} = -\lambda^2 + 6\lambda - 2.$$

We can use the quadratic formula to find the roots of the quotient:

$$\lambda = \frac{-6 \pm \sqrt{36 - 4 \cdot 2}}{-2} = 3 \pm \sqrt{7}.$$

We have factored f completely:

$$f(\lambda) = -(\lambda - 2)\big(\lambda - (3 + \sqrt{7})\big)\big(\lambda - (3 - \sqrt{7})\big).$$

Therefore, the eigenvalues of A are $2,\ 3 + \sqrt{7},\ 3 - \sqrt{7}$.

In the above example, we could have expanded cofactors along the second column to obtain

$$f(\lambda) = (2 - \lambda)\det\begin{pmatrix} 7-\lambda & 3 \\ -3 & -1-\lambda \end{pmatrix}.$$

Since $2 - \lambda$ was the only nonzero entry in its column, this expression already has the $2 - \lambda$ term factored out: the rational root theorem was not needed. The determinant in the above expression is the characteristic polynomial of the matrix $\left(\begin{smallmatrix} 7 & 3 \\ -3 & -1 \end{smallmatrix}\right)$, so we can compute it using the trace and determinant:

$$f(\lambda) = (2-\lambda)\big(\lambda^2 - (7-1)\lambda + (-7+9)\big) = (2-\lambda)(\lambda^2 - 6\lambda + 2).$$

Example. Find the eigenvalues of the matrix

$$A = \begin{pmatrix} 7 & 0 & 3 \\ -3 & 2 & -3 \\ 4 & 2 & 0 \end{pmatrix}.$$

Solution. We compute the characteristic polynomial by expanding cofactors along the first row:

$$\begin{aligned} f(\lambda) = \det(A - \lambda I_3) = \det\begin{pmatrix} 7-\lambda & 0 & 3 \\ -3 & 2-\lambda & -3 \\ 4 & 2 & -\lambda \end{pmatrix} & \\ &= (7-\lambda)\big(-\lambda(2-\lambda) + 6\big) + 3\big(-6 - 4(2-\lambda)\big) \\ &= -\lambda^3 + 9\lambda^2 - 8\lambda. \end{aligned}$$

The constant term is zero, so A has determinant zero. We factor out λ, then eyeball the roots of the quadratic factor:

$$f(\lambda) = -\lambda(\lambda^2 - 9\lambda + 8) = -\lambda(\lambda - 1)(\lambda - 8).$$

Therefore, the eigenvalues of A are $0, 1$, and 8.

Finding Eigenvalues of a Matrix Larger than 2×2. Let A be an $n \times n$ matrix. Here are some strategies for factoring its characteristic polynomial $f(\lambda)$. First, you must find one eigenvalue:

1. Do not multiply out the characteristic polynomial if it is already partially factored! This happens if you expand cofactors along the second column in this example.
2. If there is no constant term, you can factor out λ, as in this example.
3. If the matrix is triangular, the roots are the diagonal entries.
4. Guess one eigenvalue using the rational root theorem: if $\det(A)$ is an integer, substitute all (positive and negative) divisors of $\det(A)$ into $f(\lambda)$.
5. Find an eigenvalue using the geometry of the matrix. For instance, a reflection has eigenvalues ± 1.

After obtaining an eigenvalue λ_1, use polynomial long division to compute $f(\lambda)/(\lambda - \lambda_1)$. This polynomial has lower degree. If $n = 3$ then this is a quadratic polynomial, to which you can apply the quadratic formula to find the remaining roots.

5.3 Similarity

Objectives

1. Learn to interpret similar matrices geoemetrically.
2. Understand the relationship between the eigenvalues, eigenvectors, and characteristic polynomials of similar matrices.
3. *Recipe:* compute Ax in terms of B, C for $A = CBC^{-1}$.
4. *Picture:* the geometry of similar matrices.
5. *Vocabulary word:* **similarity**.

Some matrices are easy to understand. For instance, a diagonal matrix

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$$

just scales the coordinates of a vector: $D\binom{x}{y} = \binom{2x}{y/2}$. The purpose of most of the rest of this chapter is to understand complicated-looking matrices by analyzing to what extent they "behave like" simple matrices. For instance, the matrix

$$A = \frac{1}{10}\begin{pmatrix} 11 & 6 \\ 9 & 14 \end{pmatrix}$$

has eigenvalues 2 and 1/2, with corresponding eigenvectors $v_1 = \binom{2/3}{1}$ and $v_2 = \binom{-1}{1}$. Notice that

$$\begin{aligned} D(xe_1 + ye_2) &= xDe_1 + yDe_2 = 2xe_1 - \tfrac{1}{2}ye_2 \\ A(xv_1 + yv_2) &= xAv_1 + yAv_2 = 2xv_1 - \tfrac{1}{2}yv_2. \end{aligned}$$

Using v_1, v_2 instead of the usual coordinates makes A "behave" like a diagonal matrix.

Use this link to view the online demo

The matrices A and D behave similarly. Click "multiply" to multiply the colored points by D on the left and A on the right. (We will see in Section 5.4 why the points follow hyperbolic paths.)

The other case of particular importance will be matrices that "behave" like a rotation matrix: indeed, this will be crucial for understanding Section 5.5 geometrically. See this important note.

In this section, we study in detail the situation when two matrices behave similarly with respect to different coordinate systems. In Section 5.4 and Section 5.5, we will show how to use eigenvalues and eigenvectors to find a simpler matrix that behaves like a given matrix.

5.3.1 Similar Matrices

We begin with the algebraic definition of similarity.

Definition. Two $n \times n$ matrices A and B are **similar** if there exists an invertible $n \times n$ matrix C such that $A = CBC^{-1}$.

Example. The matrices

$$\begin{pmatrix} -12 & 15 \\ -10 & 13 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$$

are similar because

$$\begin{pmatrix} -12 & 15 \\ -10 & 13 \end{pmatrix} = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}^{-1},$$

as the reader can verify.

Example. The matrices

$$\begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

are not similar. Indeed, the second matrix is the identity matrix I_2, so if C is any invertible 2×2 matrix, then

$$CI_2C^{-1} = CC^{-1} = I_2 \neq \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}.$$

As in the above example, one can show that I_n is the only matrix that is similar to I_n, and likewise for any scalar multiple of I_n.

> Similarity is unrelated to row equivalence. Any invertible matrix is row equivalent to I_n, but I_n is the only matrix similar to I_n. For instance,
>
> $$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
>
> are row equivalent but not similar.

As suggested by its name, similarity is what is called an *equivalence relation*. This means that it satisfies the following properties.

Proposition. *Let A, B, and C be $n \times n$ matrices.*

1. ***Reflexivity:*** *A is similar to itself.*
2. ***Symmetry:*** *if A is similar to B, then B is similar to A.*
3. ***Transitivity:*** *if A is similar to B and B is similar to C, then A is similar to C.*

Proof.

1. Taking $C = I_n = I_n^{-1}$, we have $A = I_n A I_n^{-1}$.
2. Suppose that $A = CBC^{-1}$. Multiplying both sides on the left by C^{-1} and on the right by C gives

   $$C^{-1}AC = C^{-1}(CBC^{-1})C = B.$$

   Since $(C^{-1})^{-1} = C$, we have $B = C^{-1}A(C^{-1})^{-1}$, so that B is similar to A.
3. Suppose that $A = DBD^{-1}$ and $B = ECE^{-1}$. Subsituting for B and remembering that $(DE)^{-1} = E^{-1}D^{-1}$, we have

   $$A = D(ECE^{-1})D^{-1} = (DE)C(DE)^{-1},$$

   which shows that A is similar to C.

□

Example. The matrices

$$\begin{pmatrix} -12 & 15 \\ -10 & 13 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$$

are similar, as we saw in this example. Likewise, the matrices

$$\begin{pmatrix} -12 & 15 \\ -10 & 13 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -12 & 5 \\ -30 & 13 \end{pmatrix}$$

are similar because

$$\begin{pmatrix} -12 & 5 \\ -30 & 13 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} -12 & 15 \\ -10 & 13 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}^{-1}.$$

It follows that

$$\begin{pmatrix} -12 & 5 \\ -30 & 13 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}$$

are similar to each other.

We conclude with an observation about similarity and powers of matrices.

Fact. *Let* $A = CBC^{-1}$. *Then for any* $n \geq 1$, *we have*

$$A^n = CB^nC^{-1}.$$

Proof. First note that

$$A^2 = AA = (CBC^{-1})(CBC^{-1}) = CB(C^{-1}C)BC^{-1} = CBI_nBC^{-1} = CB^2C^{-1}.$$

Next we have

$$A^3 = A^2A = (CB^2C^{-1})(CBC^{-1}) = CB^2(C^{-1}C)BC^{-1} = CB^3C^{-1}.$$

The pattern is clear. □

Example. Compute A^{100}, where

$$A = \begin{pmatrix} 5 & 13 \\ -2 & -5 \end{pmatrix} = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}^{-1}.$$

Solution. By the fact, we have

$$A^{100} = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^{100}\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}^{-1}.$$

The matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ is a counterclockwise rotation by 90°. If we rotate by 90° four times, then we end up where we started. Hence rotating by 90° one hundred times is the identity transformation, so

$$A^{100} = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

5.3.2 Geometry of Similar Matrices

Similarity is a very interesting construction when viewed geometrically. We will see that, roughly, *similar matrices do the same thing in different coordinate systems.* The reader might want to review $\mathcal{B}$-coordinates and nonstandard coordinate grids in Section 2.8 before reading this subsection.

By conditions 4 and 5 of the invertible matrix theorem in Section 5.1, an $n\times n$ matrix C is invertible if and only if its columns $v_1, v_2, \ldots, v_n$ form a basis for $\mathbf{R}^n$. This means we can speak of the $\mathcal{B}$-coordinates of a vector in $\mathbf{R}^n$, where $\mathcal{B}$ is the basis of columns of C. Recall that

$$[x]_{\mathcal{B}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \quad \text{means} \quad x = c_1v_1 + c_2v_2 + \cdots + c_nv_n = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

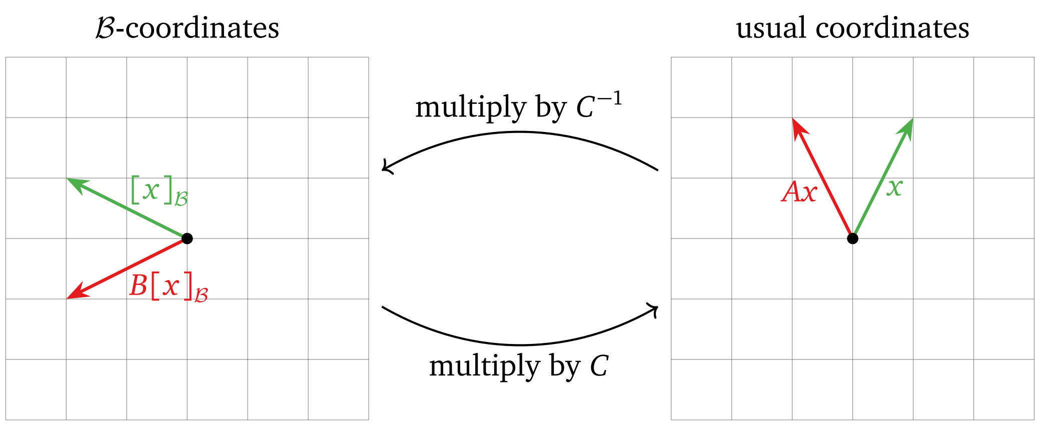
Since C is the matrix with columns $v_1, v_2, \ldots, v_n$, this says that $x = C[x]_{\mathcal{B}}$. Multiplying both sides by C^{-1} gives $[x]_{\mathcal{B}} = C^{-1}x$. To summarize:

> Let C be an invertible $n \times n$ matrix with columns $v_1, v_2, \ldots, v_n$, and let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$, a basis for $\mathbf{R}^n$. Then for any x in $\mathbf{R}^n$, we have
>
> $$C[x]_{\mathcal{B}} = x \quad \text{and} \quad C^{-1}x = [x]_{\mathcal{B}}.$$
>
> This says that C *changes from the $\mathcal{B}$-coordinates to the usual coordinates*, and C^{-1} *changes from the usual coordinates to the $\mathcal{B}$-coordinates.*

Suppose that $A = CBC^{-1}$. The above observation gives us another way of computing Ax for a vector x in $\mathbf{R}^n$. Recall that $CBC^{-1}x = C(B(C^{-1}x))$, so that multiplying CBC^{-1} by x means *first* multiplying by C^{-1}, *then* by B, *then* by C. See this example in Section 3.4.

Recipe: Computing Ax in terms of B. Suppose that $A = CBC^{-1}$, where C is an invertible matrix with columns $v_1, v_2, \ldots, v_n$. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$, a basis for $\mathbf{R}^n$. Let x be a vector in $\mathbf{R}^n$. To compute Ax, one does the following:

1. Multiply x by C^{-1}, which changes to the $\mathcal{B}$-coordinates: $[x]_{\mathcal{B}} = C^{-1}x$.
2. Multiply this by B: $B[x]_{\mathcal{B}} = BC^{-1}x$.
3. Interpreting this vector as a $\mathcal{B}$-coordinate vector, we multiply it by C to change back to the usual coordinates: $Ax = CBC^{-1}x = CB[x]_{\mathcal{B}}$.



To summarize: if $A = CBC^{-1}$, then A and B do the same thing, only in different coordinate systems.

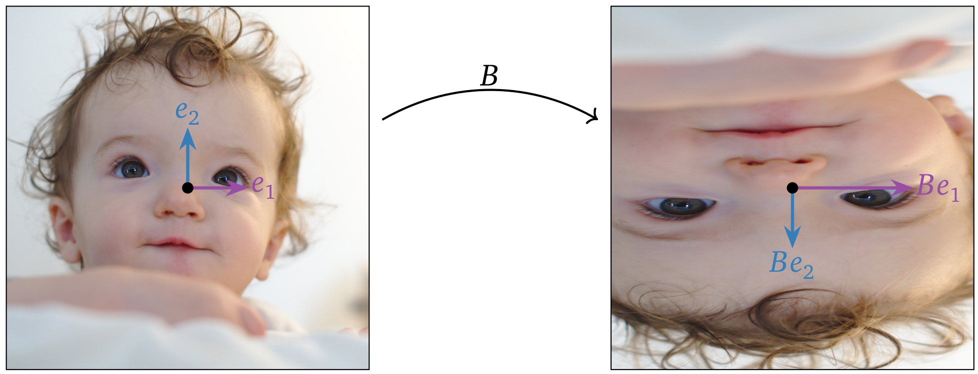
The following example is the heart of this section.

Example. Consider the matrices

$$A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

One can verify that $A = CBC^{-1}$: see this example in Section 5.4. Let $v_1 = \binom{1}{1}$ and $v_2 = \binom{1}{-1}$, the columns of C, and let $\mathcal{B} = \{v_1, v_2\}$, a basis of $\mathbf{R}^2$.

The matrix B is diagonal: it scales the x-direction by a factor of 2 and the y-direction by a factor of -1.



To compute Ax, first we multiply by C^{-1} to find the $\mathcal{B}$-coordinates of x, then we multiply by B, then we multiply by C again. For instance, let $x = \binom{0}{-2}$.

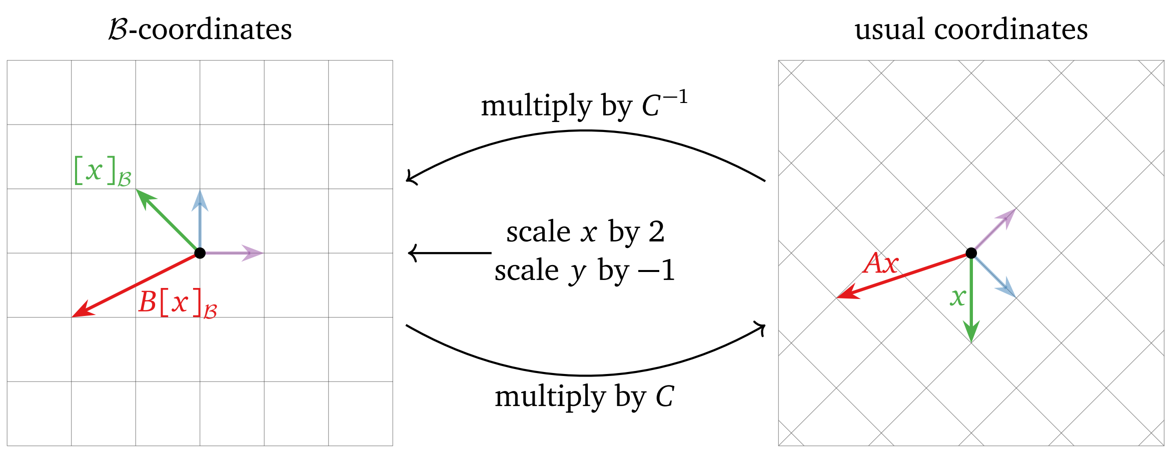
1. We see from the $\mathcal{B}$-coordinate grid below that $x = -v_1 + v_2$. Therefore, $C^{-1}x = [x]_{\mathcal{B}} = \binom{-1}{1}$.

2. Multiplying by B scales the coordinates: $B[x]_{\mathcal{B}} = \binom{-2}{-1}$.

3. Interpreting $\binom{-2}{-1}$ as a $\mathcal{B}$-coordinate vector, we multiply by C to get

$$Ax = C\begin{pmatrix}-2\\-1\end{pmatrix} = -2v_1 - v_2 = \begin{pmatrix}-3\\-1\end{pmatrix}.$$

Of course, this vector lies at $(-2,-1)$ on the $\mathcal{B}$-coordinate grid.



Now let $x = \frac{1}{2}\binom{5}{-3}$.

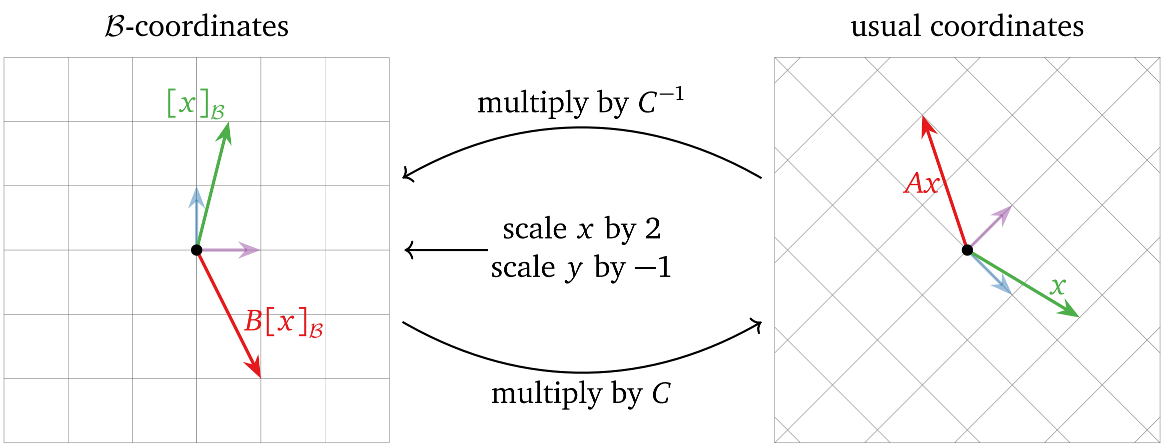
1. We see from the $\mathcal{B}$-coordinate grid that $x = \frac{1}{2}v_1 + 2v_2$. Therefore, $C^{-1}x = [x]_{\mathcal{B}} = \binom{1/2}{2}$.

2. Multiplying by B scales the coordinates: $B[x]_{\mathcal{B}} = \binom{1}{-2}$.

3. Interpreting $\binom{1}{-2}$ as a $\mathcal{B}$-coordinate vector, we multiply by C to get

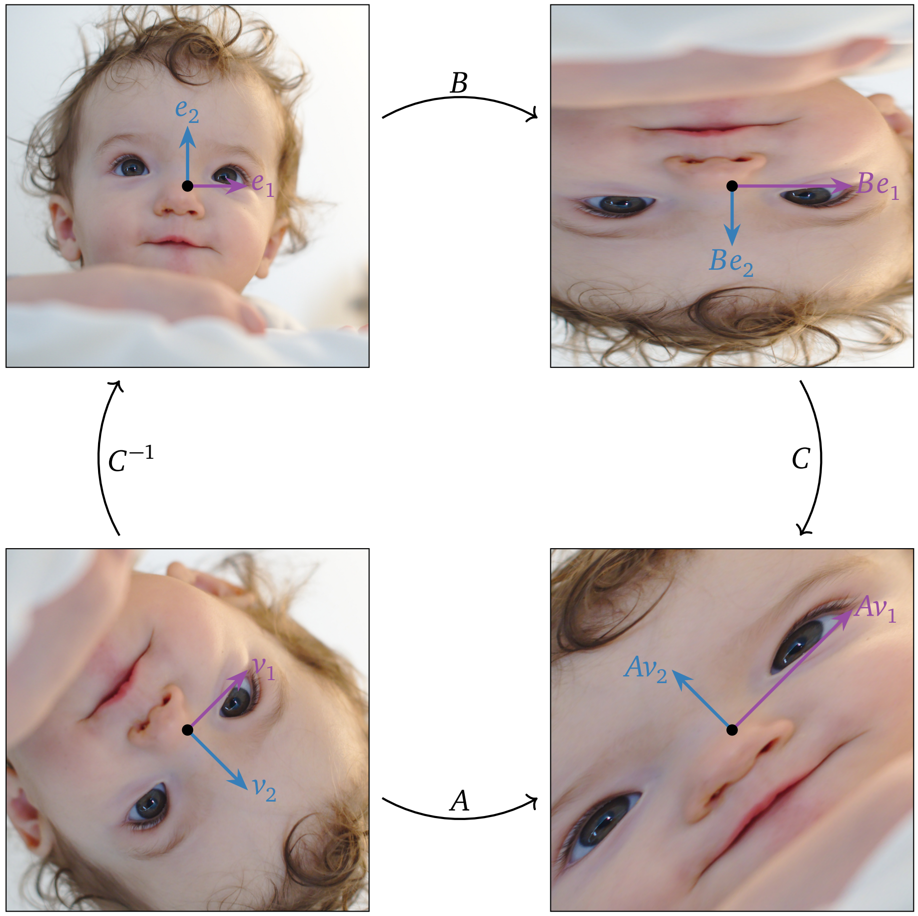
$$Ax = C\begin{pmatrix}1\\-2\end{pmatrix} = v_1 - 2v_2 = \begin{pmatrix}-1\\3\end{pmatrix}.$$

   This vector lies at $(1,-2)$ on the $\mathcal{B}$-coordinate grid.



To summarize:

- B scales the e_1-direction by 2 and the e_2-direction by -1.
- A scales the v_1-direction by 2 and the v_2-direction by -1.



Use this link to view the online demo

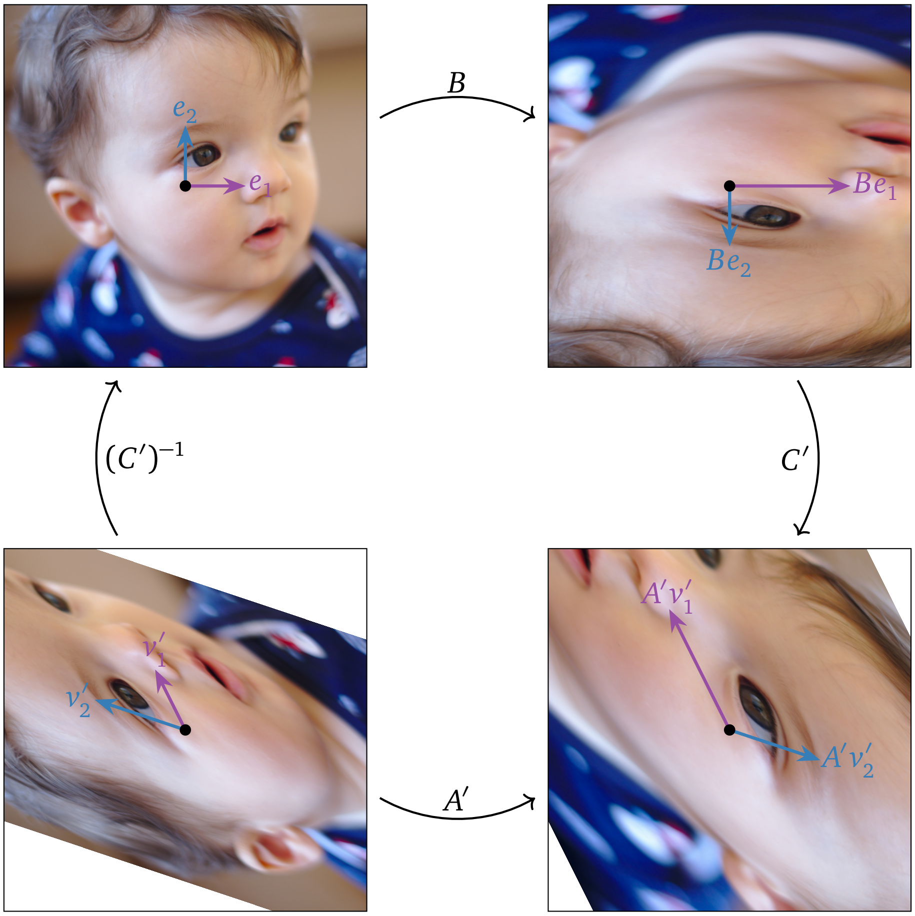
The geometric relationship between the similar matrices A and B acting on $\mathbf{R}^2$. *Click and drag the heads of* x *and* $[x]_{\mathcal{B}}$. *Study this picture until you can reliably predict where the other three vectors will be after moving one of them: this is the essence of the geometry of similar matrices.*

Interactive: Another matrix similar to B**.** Consider the matrices

$$A' = \frac{1}{5}\begin{pmatrix} -8 & -9 \\ 6 & 13 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \qquad C' = \frac{1}{2}\begin{pmatrix} -1 & -3 \\ 2 & 1 \end{pmatrix}.$$

Then $A' = C'B(C')^{-1}$, as one can verify. Let $v_1' = \frac{1}{2}\binom{-1}{2}$ and $v_2' = \frac{1}{2}\binom{-3}{1}$, the columns of C', and let $\mathcal{B}' = \{v_1', v_2'\}$. Then A' does the same thing as B, as in the previous example, except A' uses the $\mathcal{B}'$-coordinate system. In other words:

- B scales the e_1-direction by 2 and the e_2-direction by -1.
- A' scales the v_1'-direction by 2 and the v_2'-direction by -1.



Use this link to view the online demo

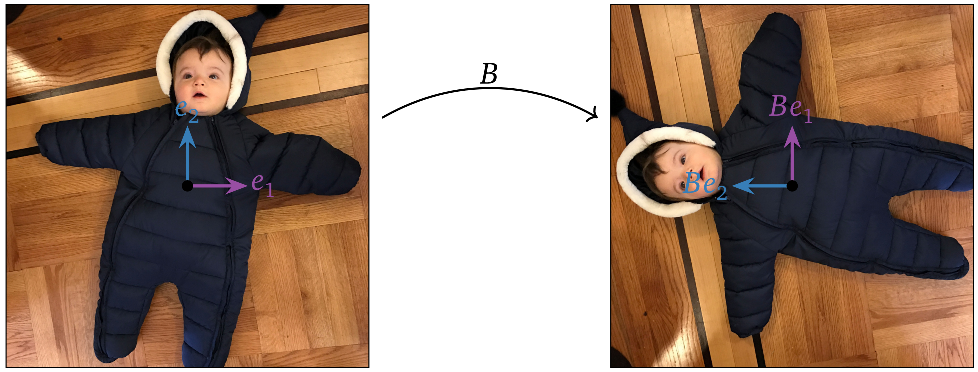
The geometric relationship between the similar matrices A' and B acting on $\mathbf{R}^2$. Click and drag the heads of x and $[x]_{\mathcal{B}'}$.

Example (A matrix similar to a rotation matrix)**.** Consider the matrices

$$A = \frac{1}{6}\begin{pmatrix} 7 & -17 \\ 5 & -7 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 2 & -1/2 \\ 1 & 1/2 \end{pmatrix}.$$

One can verify that $A = CBC^{-1}$. Let $v_1 = \binom{2}{1}$ and $v_2 = \frac{1}{2}\binom{-1}{1}$, the columns of C, and let $\mathcal{B} = \{v_1, v_2\}$, a basis of $\mathbf{R}^2$.

The matrix B rotates the plane counterclockwise by 90°.

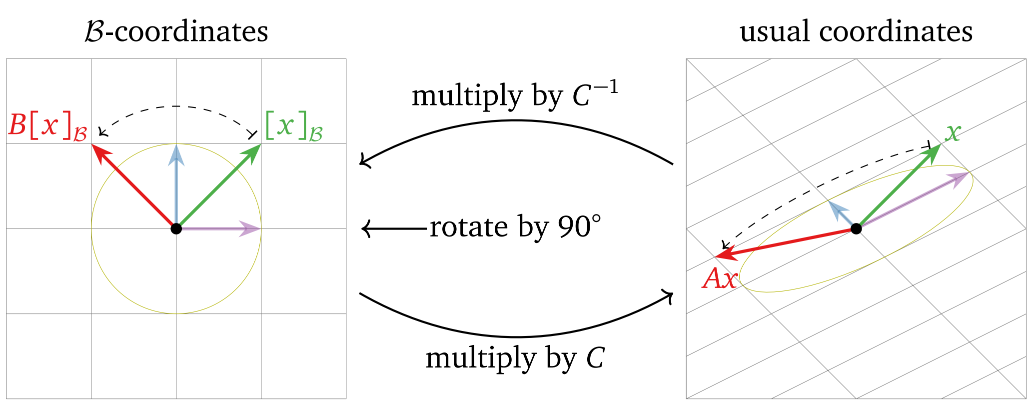


To compute Ax, first we multiply by C^{-1} to find the $\mathcal{B}$-coordinates of x, then we multiply by B, then we multiply by C again. For instance, let $x = \frac{3}{2}\begin{pmatrix}1\\1\end{pmatrix}$.

1. We see from the $\mathcal{B}$-coordinate grid below that $x = v_1 + v_2$. Therefore, $C^{-1}x = [x]_{\mathcal{B}} = \begin{pmatrix}1\\1\end{pmatrix}$.
2. Multiplying by B rotates by 90°: $B[x]_{\mathcal{B}} = \begin{pmatrix}-1\\1\end{pmatrix}$.
3. Interpreting $\begin{pmatrix}-1\\1\end{pmatrix}$ as a $\mathcal{B}$-coordinate vector, we multiply by C to get

$$Ax = C\begin{pmatrix}-1\\1\end{pmatrix} = -v_1 + v_2 = \frac{1}{2}\begin{pmatrix}-5\\-1\end{pmatrix}.$$

Of course, this vector lies at $(-1, 1)$ on the $\mathcal{B}$-coordinate grid.



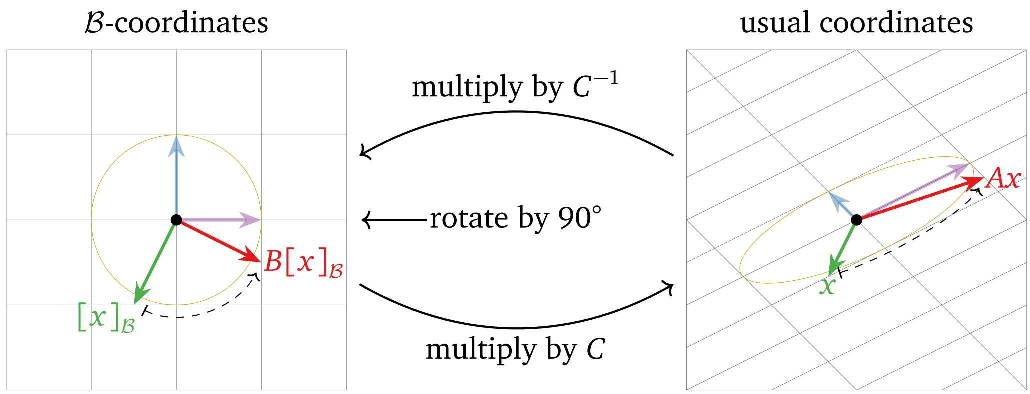
Now let $x = \frac{1}{2}\begin{pmatrix}-1\\-2\end{pmatrix}$.

1. We see from the $\mathcal{B}$-coordinate grid that $x = -\frac{1}{2}v_1 - v_2$. Therefore, $C^{-1}x = [x]_{\mathcal{B}} = \begin{pmatrix}-1/2\\-1\end{pmatrix}$.
2. Multiplying by B rotates by 90°: $B[x]_{\mathcal{B}} = \begin{pmatrix}1\\-1/2\end{pmatrix}$.

3. Interpreting $\binom{1}{-1/2}$ as a $\mathcal{B}$-coordinate vector, we multiply by C to get

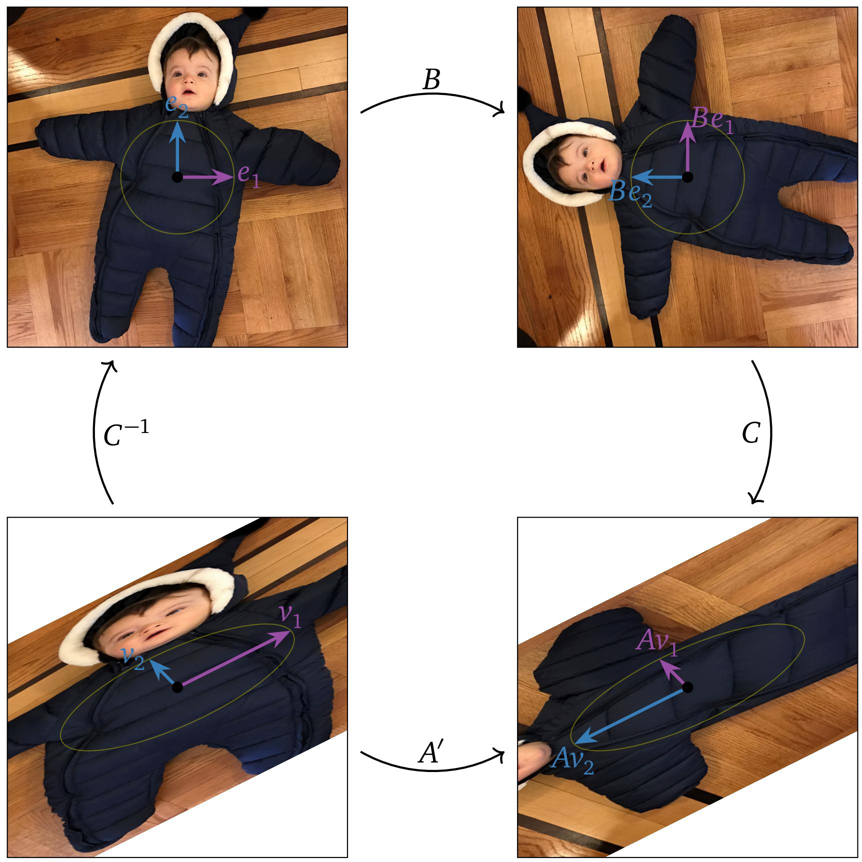
$$Ax = C\begin{pmatrix}1\\-1/2\end{pmatrix} = v_1 - \frac{1}{2}v_2 = \frac{1}{4}\begin{pmatrix}9\\3\end{pmatrix}.$$

This vector lies at $(1, -\frac{1}{2})$ on the $\mathcal{B}$-coordinate grid.



To summarize:

- B rotates counterclockwise around the circle centered at the origin and passing through e_1 and e_2.
- A rotates counterclockwise around the ellipse centered at the origin and passing through v_1 and v_2.



Use this link to view the online demo

The geometric relationship between the similar matrices A and B acting on $\mathbf{R}^2$. *Click and drag the heads of* x *and* $[x]_{\mathcal{B}}$.

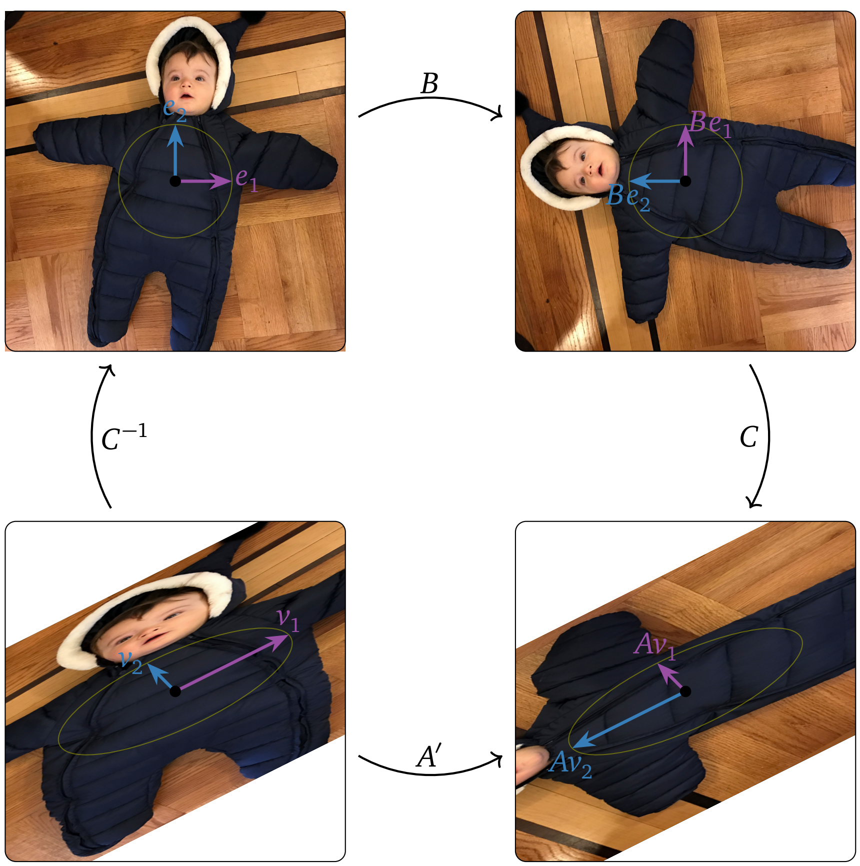
To summarize and generalize the previous example:

A Matrix Similar to a Rotation Matrix. Let

$$B = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad C = \begin{pmatrix} | & | \\ v_1 & v_2 \\ | & | \end{pmatrix} \qquad A = CBC^{-1},$$

where C is assumed invertible. Then:

- B rotates the plane by an angle of θ around the circle centered at the origin and passing through e_1 and e_2, in the direction from e_1 to e_2.
- A rotates the plane by an angle of θ around the ellipse centered at the origin and passing through v_1 and v_2, in the direction from v_1 to v_2.



Interactive: Similar 3×3 matrices. Consider the matrices

$$A = \begin{pmatrix} -1 & 0 & 0 \\ -1 & 0 & 2 \\ -1 & 1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad C = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}.$$

Then $A = CBC^{-1}$, as one can verify. Let v_1, v_2, v_3 be the columns of C, and let $\mathcal{B} = \{v_1, v_2, v_3\}$, a basis of $\mathbf{R}^3$. Then A does the same thing as B, except A uses the $\mathcal{B}$-coordinate system. In other words:

- B scales the e_1, e_2-plane by -1 and the e_3-direction by 2.
- A scales the v_1, v_2-plane by -1 and the v_3-direction by 2.

Use this link to view the online demo

The geometric relationship between the similar matrices A and B acting on $\mathbf{R}^3$. Click and drag the heads of x and $[x]_{\mathcal{B}}$.

5.3.3 Eigenvalues of Similar Matrices

Since similar matrices behave in the same way with respect to different coordinate systems, we should expect their eigenvalues and eigenvectors to be closely related.

Fact. *Similar matrices have the same characteristic polynomial.*

Proof. Suppose that $A = CBC^{-1}$, where A, B, C are $n \times n$ matrices. We calculate

$$\begin{aligned} A - \lambda I_n &= CBC^{-1} - \lambda CC^{-1} = CBC^{-1} - C\lambda C^{-1} \\ &= CBC^{-1} - C\lambda I_n C^{-1} = C(B - \lambda I_n)C^{-1}. \end{aligned}$$

Therefore,

$$\det(A - \lambda I_n) = \det(C(B - \lambda I_n)C^{-1}) = \det(C)\det(B - \lambda I_n)\det(C)^{-1} = \det(B - \lambda I_n).$$

Here we have used the multiplicativity property in Section 4.1 and its corollary in Section 4.1. □

Since the eigenvalues of a matrix are the roots of its characteristic polynomial, we have shown:

> Similar matrices have the same eigenvalues.

By this theorem in Section 5.2, similar matrices also have the same trace and determinant.

Note. The converse of the fact is false. Indeed, the matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

both have characteristic polynomial $f(\lambda) = (\lambda - 1)^2$, but they are not similar, because the only matrix that is similar to I_2 is I_2 itself.

Given that similar matrices have the same eigenvalues, one might guess that they have the same eigen*vectors* as well. Upon reflection, this is not what one should expect: indeed, the eigenvectors should only match up after changing from one coordinate system to another. This is the content of the next fact, remembering that C and C^{-1} change between the usual coordinates and the $\mathcal{B}$-coordinates.

Fact. *Suppose that $A = CBC^{-1}$. Then*

$$\begin{aligned} v \text{ is an eigenvector of } A \quad &\implies \quad C^{-1}v \text{ is an eigenvector of } B \\ v \text{ is an eigenvector of } B \quad &\implies \quad Cv \text{ is an eigenvector of } A. \end{aligned}$$

The eigenvalues of $v \,/\, C^{-1}v$ or $v \,/\, Cv$ are the same.

Proof. Suppose that v is an eigenvector of A with eigenvalue λ, so that $Av = \lambda v$. Then

$$B(C^{-1}v) = C^{-1}(CBC^{-1}v) = C^{-1}(Av) = C^{-1}\lambda v = \lambda(C^{-1}v),$$

so that $C^{-1}v$ is an eigenvector of B with eigenvalue λ. Likewise if v is an eigenvector of B with eigenvalue λ, then $Bv = \lambda v$, and we have

$$A(Cv) = (CBC^{-1})Cv = CBv = C(\lambda v) = \lambda(Cv),$$

so that Cv is an eigenvalue of A with eigenvalue λ. □

If $A = CBC^{-1}$, then C^{-1} takes the λ-eigenspace of A to the λ-eigenspace of B, and C takes the λ-eigenspace of B to the λ-eigenspace of A.

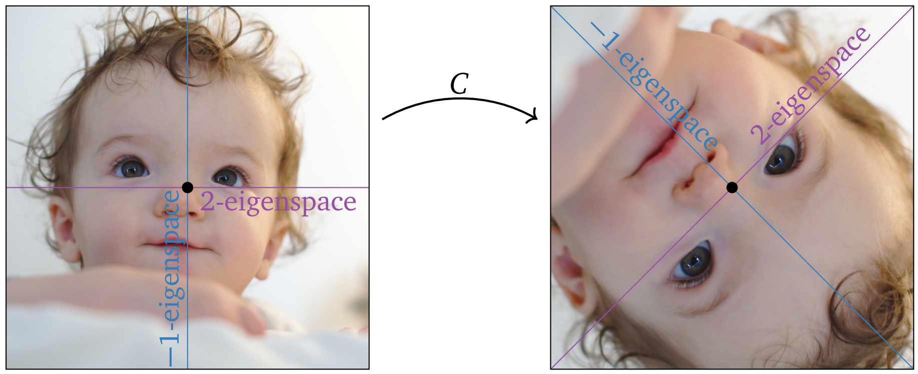
Example. We continue with the above example: let

$$A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

so $A = CBC^{-1}$. Let $v_1 = \binom{1}{1}$ and $v_2 = \binom{1}{-1}$, the columns of C. Recall that:

- B scales the e_1-direction by 2 and the e_2-direction by -1.
- A scales the v_1-direction by 2 and the v_2-direction by -1.

This means that the x-axis is the 2-eigenspace of B, and the y-axis is the -1-eigenspace of B; likewise, the "v_1-axis" is the 2-eigenspace of A, and the "v_2-axis" is the -1-eigenspace of A. This is consistent with the fact, as multiplication by C changes e_1 into $Ce_1 = v_1$ and e_2 into $Ce_2 = v_2$.



Use this link to view the online demo

The eigenspaces of A are the lines through v_1 and v_2. These are the images under C of the coordinate axes, which are the eigenspaces of B.

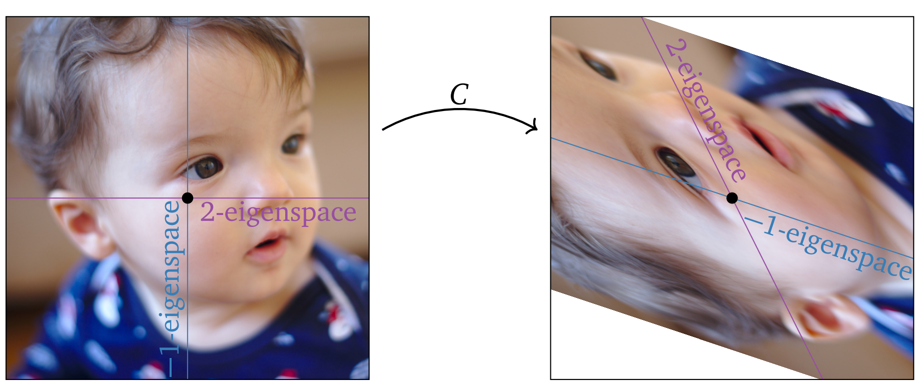
Interactive: Another matrix similar to B. Continuing with this example, let

$$A' = \frac{1}{5}\begin{pmatrix} -8 & -9 \\ 6 & 13 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \qquad C' = \frac{1}{2}\begin{pmatrix} -1 & -3 \\ 2 & 1 \end{pmatrix},$$

so $A' = C'B(C')^{-1}$. Let $v_1' = \frac{1}{2}\binom{-1}{2}$ and $v_2' = \frac{1}{2}\binom{-3}{1}$, the columns of C'. Then:

- B scales the e_1-direction by 2 and the e_2-direction by -1.
- A' scales the v_1'-direction by 2 and the v_2'-direction by -1.

As before, the x-axis is the 2-eigenspace of B, and the y-axis is the -1-eigenspace of B; likewise, the "v_1'-axis" is the 2-eigenspace of A', and the "v_2'-axis" is the -1-eigenspace of A'. This is consistent with the fact, as multiplication by C' changes e_1 into $C'e_1 = v_1'$ and e_2 into $C'e_2 = v_2'$.



Use this link to view the online demo

The eigenspaces of A' are the lines through v_1' and v_2'. These are the images under C' of the coordinate axes, which are the eigenspaces of B.

Interactive: Similar 3×3 matrices. Continuing with this example, let

$$A = \begin{pmatrix} -1 & 0 & 0 \\ -1 & 0 & 2 \\ -1 & 1 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad C = \begin{pmatrix} -1 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix},$$

so $A = CBC^{-1}$. Let v_1, v_2, v_3 be the columns of C. Then:

- B scales the e_1, e_2-plane by -1 and the e_3-direction by 2.
- A scales the v_1, v_2-plane by -1 and the v_3-direction by 2.

In other words, the xy-plane is the -1-eigenspace of B, and the z-axis is the 2-eigenspace of B; likewise, the "v_1, v_2-plane" is the -1-eigenspace of A, and the "v_3-axis" is the 2-eigenspace of A. This is consistent with the fact, as multiplication by C changes e_1 into $Ce_1 = v_1$, e_2 into $Ce_2 = v_2$, and e_3 into $Ce_3 = v_3$.

Use this link to view the online demo

The -1-eigenspace of A is the green plane, and the 2-eigenspace of A is the violet line. These are the images under C of the xy-plane and the z-axis, respectively, which are the eigenspaces of B.

5.4 Diagonalization

Objectives

1. Learn two main criteria for a matrix to be diagonalizable.
2. Develop a library of examples of matrices that are and are not diagonalizable.
3. Understand what diagonalizability and multiplicity have to say about similarity.
4. *Recipes:* diagonalize a matrix, quickly compute powers of a matrix by diagonalization.

5. *Pictures:* the geometry of diagonal matrices, why a shear is not diagonalizable.
6. *Theorem:* the diagonalization theorem (two variants).
7. *Vocabulary words:* **diagonalizable**, **algebraic multiplicity**, **geometric multiplicity**.

Diagonal matrices are the easiest kind of matrices to understand: they just scale the coordinate directions by their diagonal entries. In Section 5.3, we saw that similar matrices behave in the same way, with respect to different coordinate systems. Therefore, if a matrix is similar to a diagonal matrix, it is also relatively easy to understand. This section is devoted to the question: "When is a matrix similar to a diagonal matrix?"

5.4.1 Diagonalizability

Before answering the above question, first we give it a name.

Definition. An $n \times n$ matrix A is **diagonalizable** if it is similar to a diagonal matrix: that is, if there exists an invertible $n \times n$ matrix C and a diagonal matrix D such that

$$A = CDC^{-1}.$$

Example. Any diagonal matrix is D is diagonalizable because it is similar to itself. For instance,

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = I_3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} I_3^{-1}.$$

Example. Most of the examples in Section 5.3 involve diagonalizable matrices:

$$\begin{pmatrix} -12 & 15 \\ -10 & 13 \end{pmatrix} \text{ is diagonalizable because it equals } \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 3 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}^{-1}$$

$$\begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix} \text{ is diagonalizable because it equals } \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1}$$

$$\frac{1}{5}\begin{pmatrix} -8 & -9 \\ 6 & 13 \end{pmatrix} \text{ is diagonalizable because it equals } \frac{1}{2}\begin{pmatrix} -1 & -3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}\left(\frac{1}{2}\begin{pmatrix} -1 & -3 \\ 2 & 1 \end{pmatrix}\right)^{-1}$$

$$\begin{pmatrix} -1 & 0 & 0 \\ -1 & 0 & 2 \\ -1 & 1 & 1 \end{pmatrix} \text{ is diagonalizable because it equals } \begin{pmatrix} -1 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} -1 & 1 & 0 \\ 1 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}^{-1}.$$

Example. If a matrix A is diagonalizable, and if B is similar to A, then B is diagonalizable as well by this proposition in Section 5.3.

Powers of diagonalizable matrices Multiplying diagonal matrices together just multiplies their diagonal entries:

$$\begin{pmatrix} x_1 & 0 & 0 \\ 0 & x_2 & 0 \\ 0 & 0 & x_3 \end{pmatrix}\begin{pmatrix} y_1 & 0 & 0 \\ 0 & y_2 & 0 \\ 0 & 0 & y_3 \end{pmatrix} = \begin{pmatrix} x_1y_1 & 0 & 0 \\ 0 & x_2y_2 & 0 \\ 0 & 0 & x_3y_3 \end{pmatrix}.$$

Therefore, it is easy to take powers of a diagonal matrix:

$$\begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix}^n = \begin{pmatrix} x^n & 0 & 0 \\ 0 & y^n & 0 \\ 0 & 0 & z^n \end{pmatrix}.$$

By this fact in Section 5.3, if $A = CDC^{-1}$ then $A^n = CD^nC^{-1}$, so it is also easy to take powers of *diagonalizable* matrices. This will be very important in applications to difference equations in Section 5.6.

Recipe: Compute powers of a diagonalizable matrix. If $A = CDC^{-1}$, where D is a diagonal matrix, then $A^n = CD^nC^{-1}$:

$$A = C\begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix}C^{-1} \implies A^n = C\begin{pmatrix} x^n & 0 & 0 \\ 0 & y^n & 0 \\ 0 & 0 & z^n \end{pmatrix}C^{-1}.$$

Example. Let

$$A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1}.$$

Find a formula for A^n in which the entries are functions of n, where n is any positive whole number.

Solution. We have

$$\begin{aligned} A^n &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}^n\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2^n & 0 \\ 0 & (-1)^n \end{pmatrix}\frac{1}{-2}\begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2^n & (-1)^n \\ 2^n & (-1)^{n+1} \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 2^n + (-1)^n & 2^n + (-1)^{n+1} \\ 2^n + (-1)^{n+1} & 2^n + (-1)^n \end{pmatrix}, \end{aligned}$$

where we used $(-1)^{n+2} = (-1)^2(-1)^n = (-1)^n$.

A fundamental question about a matrix is whether or not it is diagonalizable. The following is the primary criterion for diagonalizability. It shows that diagonalizability is an eigenvalue problem.

Diagonalization Theorem. *An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.*

In this case, $A = CDC^{-1}$ for

$$C = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \qquad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where $v_1, v_2, \ldots, v_n$ are linearly independent eigenvectors, and $\lambda_1, \lambda_2, \ldots, \lambda_n$ are the corresponding eigenvalues, in the same order.

Proof. First suppose that A has n linearly independent eigenvectors $v_1, v_2, \ldots, v_n$, with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Define C as above, so C is invertible by the invertible matrix theorem in Section 5.1. Let $D = C^{-1}AC$, so $A = CDC^{-1}$. Multiplying by standard coordinate vectors picks out the columns of C: we have $Ce_i = v_i$, so $e_i = C^{-1}v_i$. We multiply by the standard coordinate vectors to find the columns of D:

$$De_i = C^{-1}ACe_i = C^{-1}Av_i = C^{-1}\lambda_i v_i = \lambda_i C^{-1}v_i = \lambda_i e_i.$$

Therefore, the columns of D are multiples of the standard coordinate vectors:

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

Now suppose that $A = CDC^{-1}$, where C has columns $v_1, v_2, \ldots, v_n$, and D is diagonal with diagonal entries $\lambda_1, \lambda_2, \ldots, \lambda_n$. Since C is invertible, its columns are linearly independent. We have to show that v_i is an eigenvector of A with eigenvalue λ_i. We know that the standard coordinate vector e_i is an eigenvector of D with eigenvalue λ_i, so:

$$Av_i = CDC^{-1}v_i = CDe_i = C\lambda_i e_i = \lambda_i Ce_i = \lambda_i v_i. \qquad \square$$

By this fact in Section 5.1, if an $n \times n$ matrix A has n *distinct* eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then a choice of corresponding eigenvectors $v_1, v_2, \ldots, v_n$ is automatically linearly independent.

> An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Easy Example. Apply the diagonalization theorem to the matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Solution. This diagonal matrix is in particular upper-triangular, so its eigenvalues are the diagonal entries $1, 2, 3$. The standard coordinate vectors are eigenvalues of a diagonal matrix:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Therefore, the diagonalization theorem says that $A = CDC^{-1}$, where the columns of C are the standard coordinate vectors, and the D is the diagonal matrix with entries $1, 2, 3$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1}.$$

This just tells us that A is similar to itself.

Actually, the diagonalization theorem is not completely trivial even for diagonal matrices. If we put our eigenvalues in the order $3, 2, 1$, then the corresponding eigenvectors are e_3, e_2, e_1, so we also have that $A = C'D'(C')^{-1}$, where C' is the matrix with columns e_3, e_2, e_1, and D' is the diagonal matrix with entries $3, 2, 1$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1}.$$

In particular, the matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are similar to each other.

Non-Uniqueness of Diagonalization. We saw in the above example that changing the order of the eigenvalues and eigenvectors produces a different diagonalization of the same matrix. There are generally many different ways to diagonalize a matrix, corresponding to different orderings of the eigenvalues of that matrix. The important thing is that the eigenvalues and eigenvectors have to be listed in the same order.

$$\begin{aligned} A &= \begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix}^{-1} \\ &= \begin{pmatrix} | & | & | \\ v_3 & v_2 & v_1 \\ | & | & | \end{pmatrix} \begin{pmatrix} \lambda_3 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix} \begin{pmatrix} | & | & | \\ v_3 & v_2 & v_1 \\ | & | & | \end{pmatrix}^{-1}. \end{aligned}$$

There are other ways of finding different diagonalizations of the same matrix. For instance, you can scale one of the eigenvectors by a constant c:

$$\begin{aligned} A &= \begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix}^{-1} \\ &= \begin{pmatrix} | & | & | \\ cv_1 & v_2 & v_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \begin{pmatrix} | & | & | \\ cv_1 & v_2 & v_3 \\ | & | & | \end{pmatrix}^{-1}, \end{aligned}$$

you can find a different basis entirely for an eigenspace of dimension at least 2, etc.

Example (A diagonalizable 2×2 matrix)**.** Diagonalize the matrix

$$A = \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix}.$$

Solution. We need to find the eigenvalues and eigenvectors of A. First we compute the characteristic polynomial:

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) = \lambda^2 - \lambda - 2 = (\lambda+1)(\lambda-2).$$

Therefore, the eigenvalues are -1 and 2. We need to compute eigenvectors for each eigenvalue. We start with $\lambda_1 = -1$:

$$(A + 1I_2)v = 0 \iff \begin{pmatrix} 3/2 & 3/2 \\ 3/2 & 3/2 \end{pmatrix} v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} v = 0.$$

The parametric form is $x = -y$, so $v_1 = \binom{-1}{1}$ is an eigenvector with eigenvalue λ_1. Now we find an eigenvector with eigenvalue $\lambda_2 = 2$:

$$(A - 2I_2)v = 0 \iff \begin{pmatrix} -3/2 & 3/2 \\ 3/2 & -3/2 \end{pmatrix} v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} v = 0.$$

The parametric form is $x = y$, so $v_2 = \binom{1}{1}$ is an eigenvector with eigenvalue 2.

The eigenvectors v_1, v_2 are linearly independent, so the diagonalization theorem says that

$$A = CDC^{-1} \quad \text{for} \quad C = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Alternatively, if we choose 2 as our first eigenvalue, then

$$A = C'D'(C')^{-1} \quad \text{for} \quad C' = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \qquad D' = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.$$

Use this link to view the online demo

The green line is the -1-eigenspace of A, and the violet line is the 2-eigenspace. There are two linearly independent (noncollinear) eigenvectors visible in the picture: choose any nonzero vector on the green line, and any nonzero vector on the violet line.

Example (A diagonalizable 2×2 matrix with a zero eigenvector)**.** Diagonalize the matrix

$$A = \begin{pmatrix} 2/3 & -4/3 \\ -2/3 & 4/3 \end{pmatrix}.$$

Solution. We need to find the eigenvalues and eigenvectors of A. First we compute the characteristic polynomial:

$$f(\lambda) = \lambda^2 - \operatorname{Tr}(A)\lambda + \det(A) = \lambda^2 - 2\lambda = \lambda(\lambda - 2).$$

Therefore, the eigenvalues are 0 and 2. We need to compute eigenvectors for each eigenvalue. We start with $\lambda_1 = 0$:

$$(A - 0I_2)v = 0 \iff \begin{pmatrix} 2/3 & -4/3 \\ -2/3 & 4/3 \end{pmatrix} v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} v = 0.$$

The parametric form is $x = 2y$, so $v_1 = \binom{2}{1}$ is an eigenvector with eigenvalue λ_1. Now we find an eigenvector with eigenvalue $\lambda_2 = 2$:

$$(A - 2I_2)v = 0 \iff \begin{pmatrix} -4/3 & -4/3 \\ -2/3 & -2/3 \end{pmatrix} v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} v = 0.$$

The parametric form is $x = -y$, so $v_2 = \binom{1}{-1}$ is an eigenvector with eigenvalue 2.

The eigenvectors v_1, v_2 are linearly independent, so the diagonalization theorem says that

$$A = CDC^{-1} \quad \text{for} \quad C = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \qquad D = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

Alternatively, if we choose 2 as our first eigenvalue, then

$$A = C'D'(C')^{-1} \quad \text{for} \quad C' = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \qquad D' = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

In the above example, the (non-invertible) matrix $A = \frac{1}{3}\begin{pmatrix} 2 & -4 \\ -2 & 4 \end{pmatrix}$ is similar to the diagonal matrix $D = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$. Since A is not invertible, zero is an eigenvalue by the invertible matrix theorem, so one of the diagonal entries of D is necessarily zero. Also see this example below.

Example (A diagonalizable 3 × 3 matrix)**.** Diagonalize the matrix

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

Solution. We need to find the eigenvalues and eigenvectors of A. First we compute the characteristic polynomial by expanding cofactors along the third column:

$$\begin{aligned} f(\lambda) = \det(A - \lambda I_3) &= (1-\lambda)\det\left(\begin{pmatrix} 4 & -3 \\ 2 & -1 \end{pmatrix} - \lambda I_2\right) \\ &= (1-\lambda)(\lambda^2 - 3\lambda + 2) = -(\lambda-1)^2(\lambda-2). \end{aligned}$$

Therefore, the eigenvalues are 1 and 2. We need to compute eigenvectors for each eigenvalue. We start with $\lambda_1 = 1$:

$$(A - I_3)v = 0 \iff \begin{pmatrix} 3 & -3 & 0 \\ 2 & -2 & 0 \\ 1 & -1 & 0 \end{pmatrix} v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} v = 0.$$

The parametric vector form is

$$\begin{cases} x = y \\ y = y \\ z = \qquad z \end{cases} \implies \begin{pmatrix} x \\ y \\ z \end{pmatrix} = y\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Hence a basis for the 1-eigenspace is

$$\mathcal{B}_1 = \{v_1, v_2\} \quad \text{where} \quad v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Now we compute the eigenspace for $\lambda_2 = 2$:

$$(A - 2I_3)v = 0 \iff \begin{pmatrix} 2 & -3 & 0 \\ 2 & -3 & 0 \\ 1 & -1 & -1 \end{pmatrix} v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} v = 0$$

The parametric form is $x = 3z, y = 2z$, so an eigenvector with eigenvalue 2 is

$$v_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

The eigenvectors v_1, v_2, v_3 are linearly independent: v_1, v_2 form a basis for the 1-eigenspace, and v_3 is not contained in the 1-eigenspace because its eigenvalue is 2. Therefore, the diagonalization theorem says that

$$A = CDC^{-1} \quad \text{for} \quad C = \begin{pmatrix} 1 & 0 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Use this link to view the online demo

The green plane is the 1*-eigenspace of A, and the violet line is the* 2*-eigenspace. There are three linearly independent eigenvectors visible in the picture: choose any two noncollinear vectors on the green plane, and any nonzero vector on the violet line.*

Here is the procedure we used in the above examples.

Recipe: Diagonalization. Let A be an $n \times n$ matrix. To diagonalize A:

1. Find the eigenvalues of A using the characteristic polynomial.
2. For each eigenvalue λ of A, compute a basis $\mathcal{B}_\lambda$ for the λ-eigenspace.
3. If there are fewer than n total vectors in all of the eigenspace bases $\mathcal{B}_\lambda$, then the matrix is not diagonalizable.
4. Otherwise, the n vectors $v_1, v_2, \ldots, v_n$ in the eigenspace bases are linearly independent, and $A = CDC^{-1}$ for

$$C = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_n \\ | & | & & | \end{pmatrix} \quad \text{and} \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where λ_i is the eigenvalue for v_i.

We will justify the linear independence assertion in part 4 in the proof of this theorem below.

Example (A shear is not diagonalizable)**.** Let

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

so $T(x) = Ax$ is a shear. The characteristic polynomial of A is $f(\lambda) = (\lambda - 1)^2$, so the only eigenvalue of A is 1. We compute the 1-eigenspace:

$$(A - I_2)v = 0 \iff \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 0 \iff y = 0.$$

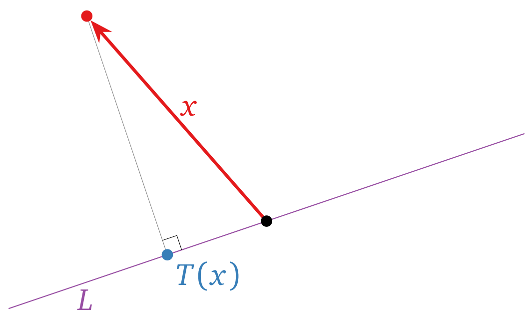
In other words, the 1-eigenspace is exactly the x-axis, so *all* of the eigenvectors of A lie on the x-axis. It follows that A does *not* admit two linearly independent eigenvectors, so by the diagonalization theorem, it is not diagonalizable.

In this example in Section 5.1, we studied the eigenvalues of a shear geometrically; we reproduce the interactive demo here.

Use this link to view the online demo

All eigenvectors of a shear lie on the x-axis.

Example (A projection is diagonalizable). Let L be a line through the origin in $\mathbf{R}^2$, and define $T: \mathbf{R}^2 \to \mathbf{R}^2$ to be the transformation that sends a vector x to the closest point on L to x, as in the picture below.



This is an example of an *orthogonal projection*. We will see in Section 6.3 that T is a linear transformation; let A be the matrix for T. Any vector on L is not moved by T because it is the closest point on L to itself: hence it is an eigenvector of A with eigenvalue 1. Let $L^\perp$ be the line perpendicular to L and passing through the origin. Any vector x on $L^\perp$ is closest to the zero vector on L, so a (nonzero) such vector is an eigenvector of A with eigenvalue 0. (See this example in Section 5.1 for a special case.) Since A has two distinct eigenvalues, it is diagonalizable; in fact, we know from the diagonalization theorem that A is similar to the matrix $\begin{pmatrix}1&0\\0&0\end{pmatrix}$.

Note that we never had to do any algebra! We know that A is diagonalizable for *geometric* reasons.

Use this link to view the online demo

The line L (violet) is the 1-eigenspace of A, and $L^\perp$ (green) is the 0-eigenspace. Since there are linearly independent eigenvectors, we know that A is diagonalizable.

Example (A non-diagonalizable 3×3 matrix). Let

$$A = \begin{pmatrix}1&1&0\\0&1&0\\0&0&2\end{pmatrix}.$$

The characteristic polynomial of A is $f(\lambda) = -(\lambda-1)^2(\lambda-2)$, so the eigenvalues of A are 1 and 2. We compute the 1-eigenspace:

$$(A-I_3)v = 0 \iff \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff y = z = 0.$$

In other words, the 1-eigenspace is the x-axis. Similarly,

$$(A-2I_3)v = 0 \iff \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0 \iff x = y = 0,$$

so the 2-eigenspace is the z-axis. In particular, all eigenvectors of A lie on the xz-plane, so there do not exist three linearly independent eigenvectors of A. By the diagonalization theorem, the matrix A is not diagonalizable.

Notice that A contains a 2×2 block on its diagonal that looks like a shear:

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

This makes one suspect that such a matrix is not diagonalizable.

Use this link to view the online demo

All eigenvectors of A lie on the x- and z-axes.

Example (A rotation matrix)**.** Let

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

so $T(x) = Ax$ is the linear transformation that rotates counterclockwise by 90°. We saw in this example in Section 5.1 that A does not have any eigenvectors at all. It follows that A is not diagonalizable.

Use this link to view the online demo

This rotation matrix has no eigenvectors.

The characteristic polynomial of A is $f(\lambda) = \lambda^2 + 1$, which of course does not have any real roots. If we allow *complex* numbers, however, then f has *two* roots, namely, $\pm i$, where $i = \sqrt{-1}$. Hence the matrix is diagonalizable if we allow ourselves to use complex numbers. We will treat this topic in detail in Section 5.5.

The following point is often a source of confusion.

Diagonalizability has nothing to do with invertibility. Of the following matrices, the first is diagonalizable and invertible, the second is diagonalizable but not invertible, the third is invertible but not diagonalizable, and the fourth is neither invertible nor diagonalizable, as the reader can verify:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Remark (Non-diagonalizable 2×2 matrices with an eigenvalue)**.** As in the above example, one can check that the matrix

$$A_\lambda = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

is not diagonalizable for any number λ. We claim that any non-diagonalizable 2×2 matrix B with a real eigenvalue λ is similar to A_λ. Therefore, up to similarity, these are the only such examples.

To prove this, let B be such a matrix. Let v_1 be an eigenvector with eigenvalue λ, and let v_2 be any vector in $\mathbf{R}^2$ that is not collinear with v_1, so that $\{v_1, v_2\}$ forms a basis for $\mathbf{R}^2$. Let C be the matrix with columns v_1, v_2, and consider $A = C^{-1}BC$. We have $Ce_1 = v_1$ and $Ce_2 = v_2$, so $C^{-1}v_1 = e_1$ and $C^{-1}v_2 = e_2$. We can compute the first column of A as follows:

$$Ae_1 = C^{-1}BCe_1 = C^{-1}Bv_1 = C^{-1}\lambda v_1 = \lambda C^{-1}v_1 = \lambda e_1.$$

Therefore, A has the form

$$A = \begin{pmatrix} \lambda & b \\ 0 & d \end{pmatrix}.$$

Since A is similar to B, it also has only one eigenvalue λ; since A is upper-triangular, this implies $d = \lambda$, so

$$A = \begin{pmatrix} \lambda & b \\ 0 & \lambda \end{pmatrix}.$$

As B is not diagonalizable, we know A is not diagonal (B is similar to A), so $b \neq 0$. Now we observe that

$$\begin{pmatrix} 1/b & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \lambda & b \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} 1/b & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \lambda/b & 1 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} b & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = A_\lambda.$$

We have shown that B is similar to A, which is similar to A_λ, so B is similar to A_λ by the transitivity property of similar matrices.

5.4.2 The Geometry of Diagonalizable Matrices

A diagonal matrix is easy to understand geometrically, as it just scales the coordinate axes:

$$\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}=1\cdot\begin{pmatrix}1\\0\\0\end{pmatrix}\qquad\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}\begin{pmatrix}0\\1\\0\end{pmatrix}=2\cdot\begin{pmatrix}0\\1\\0\end{pmatrix}$$

$$\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}=3\cdot\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

Therefore, we know from Section 5.3 that a diagonalizable matrix simply scales the "axes" with respect to a different coordinate system. Indeed, if $v_1, v_2, \ldots, v_n$ are linearly independent eigenvectors of an $n\times n$ matrix A, then A scales the v_i-direction by the eigenvalue λ_i.

In the following examples, we visualize the action of a diagonalizable matrix A in terms of its *dynamics*. In other words, we start with a collection of vectors (drawn as points), and we see where they move when we multiply them by A repeatedly.

Example (Eigenvalues $|\lambda_1| > 1$, $|\lambda_2| < 1$)**.** Describe how the matrix

$$A=\frac{1}{10}\begin{pmatrix}11&6\\9&14\end{pmatrix}$$

acts on the plane.

Solution. First we diagonalize A. The characteristic polynomial is

$$f(\lambda)=\lambda^2-\mathrm{Tr}(A)\lambda+\det(A)=\lambda^2-\frac{5}{2}\lambda+1=(\lambda-2)\left(\lambda-\frac{1}{2}\right).$$

We compute the 2-eigenspace:

$$(A-2I_3)v=0\iff\frac{1}{10}\begin{pmatrix}-9&6\\9&-6\end{pmatrix}v=0\xrightarrow{\text{RREF}}\begin{pmatrix}1&-2/3\\0&0\end{pmatrix}v=0.$$

The parametric form of this equation is $x=2/3y$, so one eigenvector is $v_1=\binom{2/3}{1}$. For the 1/2-eigenspace, we have:

$$\left(A-\frac{1}{2}I_3\right)v=0\iff\frac{1}{10}\begin{pmatrix}6&6\\9&9\end{pmatrix}v=0\xrightarrow{\text{RREF}}\begin{pmatrix}1&1\\0&0\end{pmatrix}v=0.$$

The parametric form of this equation is $x=-y$, so an eigenvector is $v_2=\binom{-1}{1}$. It follows that $A=CDC^{-1}$, where

$$C=\begin{pmatrix}2/3&-1\\1&1\end{pmatrix}\qquad D=\begin{pmatrix}2&0\\0&1/2\end{pmatrix}.$$

The diagonal matrix D scales the x-coordinate by 2 and the y-coordinate by 1/2. Therefore, it moves vectors closer to the x-axis and farther from the y-axis. In fact, since $(2x)(y/2) = xy$, multiplication by D does not move a point off of a hyperbola $xy = C$.

The matrix A does the same thing, in the v_1, v_2-coordinate system: multiplying a vector by A scales the v_1-coordinate by 2 and the v_2-coordinate by 1/2. Therefore, A moves vectors closer to the 2-eigenspace and farther from the 1/2-eigenspace.

Use this link to view the online demo

Dynamics of the matrices A and D. Click "multiply" to multiply the colored points by D on the left and A on the right.

Example (Eigenvalues $|\lambda_1| > 1$, $|\lambda_2| > 1$)**.** Describe how the matrix

$$A = \frac{1}{5}\begin{pmatrix} 13 & -2 \\ -3 & 12 \end{pmatrix}$$

acts on the plane.

Solution. First we diagonalize A. The characteristic polynomial is

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) = \lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3).$$

Next we compute the 2-eigenspace:

$$(A - 2I_3)v = 0 \iff \frac{1}{5}\begin{pmatrix} 3 & -2 \\ -3 & 2 \end{pmatrix} v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -2/3 \\ 0 & 0 \end{pmatrix} v = 0.$$

The parametric form of this equation is $x = 2/3y$, so one eigenvector is $v_1 = \binom{2/3}{1}$. For the 3-eigenspace, we have:

$$(A - 3I_3)v = 0 \iff \frac{1}{5}\begin{pmatrix} -2 & -2 \\ -3 & -3 \end{pmatrix} v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} v = 0.$$

The parametric form of this equation is $x = -y$, so an eigenvector is $v_2 = \binom{-1}{1}$. It follows that $A = CDC^{-1}$, where

$$C = \begin{pmatrix} 2/3 & -1 \\ 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

The diagonal matrix D scales the x-coordinate by 2 and the y-coordinate by 3. Therefore, it moves vectors farther from both the x-axis and the y-axis, but faster in the y-direction than the x-direction.

The matrix A does the same thing, in the v_1, v_2-coordinate system: multiplying a vector by A scales the v_1-coordinate by 2 and the v_2-coordinate by 3. Therefore, A moves vectors farther from the 2-eigenspace and the 3-eigenspace, but faster in the v_2-direction than the v_1-direction.

Use this link to view the online demo

Dynamics of the matrices A and D. Click "multiply" to multiply the colored points by D on the left and A on the right.

Example (Eigenvalues $|\lambda_1| < 1$, $|\lambda_2| < 1$)**.** Describe how the matrix

$$A' = \frac{1}{30}\begin{pmatrix} 12 & 2 \\ 3 & 13 \end{pmatrix}$$

acts on the plane.

Solution. This is the inverse of the matrix A from the previous example. In that example, we found $A = CDC^{-1}$ for

$$C = \begin{pmatrix} 2/3 & -1 \\ 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Therefore, remembering that taking inverses reverses the order of multiplication, we have

$$A' = A^{-1} = (CDC^{-1})^{-1} = (C^{-1})^{-1}D^{-1}C^{-1} = C\begin{pmatrix} 1/2 & 0 \\ 0 & 1/3 \end{pmatrix}C^{-1}.$$

The diagonal matrix D^{-1} does the opposite of what D does: it scales the x-coordinate by $1/2$ and the y-coordinate by $1/3$. Therefore, it moves vectors closer to both coordinate axes, but faster in the y-direction. The matrix A' does the same thing, but with respect to the v_1, v_2-coordinate system.

Use this link to view the online demo

Dynamics of the matrices A' and D^{-1}. Click "multiply" to multiply the colored points by D^{-1} on the left and A' on the right.

Example (Eigenvalues $|\lambda_1| = 1$, $|\lambda_2| < 1$)**.** Describe how the matrix

$$A = \frac{1}{6}\begin{pmatrix} 5 & -1 \\ -2 & 4 \end{pmatrix}$$

acts on the plane.

Solution. First we diagonalize A. The characteristic polynomial is

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = (\lambda - 1)\left(\lambda - \frac{1}{2}\right).$$

Next we compute the 1-eigenspace:

$$(A - I_3)v = 0 \iff \frac{1}{6}\begin{pmatrix} -1 & -1 \\ -2 & -2 \end{pmatrix}v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}v = 0.$$

The parametric form of this equation is $x = -y$, so one eigenvector is $v_1 = \binom{-1}{1}$. For the 1/2-eigenspace, we have:

$$\left(A - \frac{1}{2}I_3\right)v = 0 \iff \frac{1}{6}\begin{pmatrix} 2 & -1 \\ -2 & 1 \end{pmatrix} v = 0 \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & -1/2 \\ 0 & 0 \end{pmatrix} v = 0.$$

The parametric form of this equation is $x = 1/2y$, so an eigenvector is $v_2 = \binom{1/2}{1}$. It follows that $A = CDC^{-1}$, where

$$C = \begin{pmatrix} -1 & 1/2 \\ 1 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

The diagonal matrix D scales the y-coordinate by 1/2 and does not move the x-coordinate. Therefore, it simply moves vectors closer to the x-axis along vertical lines. The matrix A does the same thing, in the v_1, v_2-coordinate system: multiplying a vector by A scales the v_2-coordinate by 1/2 and does not change the v_1-coordinate. Therefore, A "sucks vectors into the 1-eigenspace" along lines parallel to v_2.

Use this link to view the online demo

Dynamics of the matrices A and D. Click "multiply" to multiply the colored points by D on the left and A on the right.

Interactive: A diagonalizable 3×3 **matrix.** The diagonal matrix

$$D = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3/2 \end{pmatrix}$$

scales the x-coordinate by 1/2, the y-coordinate by 2, and the z-coordinate by 3/2. Looking straight down at the xy-plane, the points follow parabolic paths taking them away from the x-axis and toward the y-axis. The z-coordinate is scaled by 3/2, so points fly away from the xy-plane in that direction.

If $A = CDC^{-1}$ for some invertible matrix C, then A does the same thing as D, but with respect to the coordinate system defined by the columns of C.

Use this link to view the online demo

Dynamics of the matrices A and D. Click "multiply" to multiply the colored points by D on the left and A on the right.

5.4.3 Algebraic and Geometric Multiplicity

In this subsection, we give a variant of the diagonalization theorem that provides another criterion for diagonalizability. It is stated in the language of *multiplicities* of eigenvalues.

In algebra, we define the *multiplicity* of a root λ_0 of a polynomial $f(\lambda)$ to be the number of factors of $\lambda-\lambda_0$ that divide $f(\lambda)$. For instance, in the polynomial

$$f(\lambda)=-\lambda^3+4\lambda^2-5\lambda+2=-(\lambda-1)^2(\lambda-2),$$

the root $\lambda_0=2$ has multiplicity 1, and the root $\lambda_0=1$ has multiplicity 2.

Definition. Let A be an $n\times n$ matrix, and let λ be an eigenvalue of A.

1. The **algebraic multiplicity** of λ is its multiplicity as a root of the characteristic polynomial of A.

2. The **geometric multiplicity** of λ is the dimension of the λ-eigenspace.

Since the λ-eigenspace of A is $\mathrm{Nul}(A-\lambda I_n)$, its dimension is the number of free variables in the system of equations $(A-\lambda I_n)x=0$, i.e., the number of columns without pivots in the matrix $A-\lambda I_n$.

Example. The shear matrix

$$A=\begin{pmatrix}1&1\\0&1\end{pmatrix}$$

has only one eigenvalue $\lambda=1$. The characteristic polynomial of A is $f(\lambda)=(\lambda-1)^2$, so 1 has algebraic multiplicity 2, as it is a double root of f. On the other hand, we showed in this example that the 1-eigenspace of A is the x-axis, so the geometric multiplicity of 1 is equal to 1. This matrix is not diagonalizable.

Use this link to view the online demo

Eigenspace of the shear matrix, with multiplicities.

The identity matrix

$$I_2=\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

also has characteristic polynomial $(\lambda-1)^2$, so the eigenvalue 1 has algebraic multiplicity 2. Since every nonzero vector in $\mathbf{R}^2$ is an eigenvector of I_2 with eigenvalue 1, the 1-eigenspace is all of $\mathbf{R}^2$, so the geometric multiplicity is 2 as well. This matrix is diagonal.

Use this link to view the online demo

Eigenspace of the identity matrix, with multiplicities.

Example. Continuing with this example, let

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

The characteristic polynomial is $f(\lambda) = -(\lambda-1)^2(\lambda-2)$, so that 1 and 2 are the eigenvalues, with algebraic multiplicities 2 and 1, respectively. We computed that the 1-eigenspace is a plane and the 2-eigenspace is a line, so that 1 and 2 also have geometric multiplicities 2 and 1, respectively. This matrix is diagonalizable.

Use this link to view the online demo

The green plane is the 1*-eigenspace of A, and the violet line is the* 2*-eigenspace. Hence the geometric multiplicity of the* 1*-eigenspace is* 2*, and the geometric multiplicity of the* 2*-eigenspace is* 1.

In this example, we saw that the matrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

also has characteristic polynomial $f(\lambda) = -(\lambda-1)^2(\lambda-2)$, so that 1 and 2 are the eigenvalues, with algebraic multiplicities 2 and 1, respectively. In this case, however, both eigenspaces are *lines,* so that both eigenvalues have geometric multiplicity 1. This matrix is not diagonalizable.

Use this link to view the online demo

Both eigenspaces of A are lines, so they both have geometric multiplicity 1.

We saw in the above examples that the algebraic and geometric multiplicities need not coincide. However, they do satisfy the following fundamental inequality, the proof of which is beyond the scope of this text.

Theorem (Algebraic and Geometric Multiplicity)**.** *Let A be a square matrix and let λ be an eigenvalue of A. Then*

$$1 \leq \text{(the geometric multiplicity of } \lambda) \leq \text{(the algebraic multiplicity of } \lambda).$$

In particular, if the algebraic multiplicity of λ is equal to 1, then so is the geometric multiplicity.

If A has an eigenvalue λ with algebraic multiplicity 1, then the λ-eigenspace is a *line.*

We can use the theorem to give another criterion for diagonalizability (in addition to the diagonalization theorem).

Diagonalization Theorem, Variant. *Let A be an $n \times n$ matrix. The following are equivalent:*

1. *A is diagonalizable.*

2. *The sum of the geometric multiplicities of the eigenvalues of A is equal to n.*

3. *The sum of the algebraic multiplicities of the eigenvalues of A is equal to n, and for each eigenvalue, the geometric multiplicity equals the algebraic multiplicity.*

Proof. We will show $1 \implies 2 \implies 3 \implies 1$. First suppose that A is diagonalizable. Then A has n linearly independent eigenvectors $v_1, v_2, \ldots, v_n$. This implies that the sum of the geometric multiplicities is *at least* n: for instance, if v_1, v_2, v_3 have the same eigenvalue λ, then the geometric multiplicity of λ is at least 3 (as the λ-eigenspace contains three linearly independent vectors), and so on. But the sum of the algebraic multiplicities is greater than or equal to the sum of the geometric multiplicities by the theorem, and the sum of the algebraic multiplicities is at most n because the characteristic polynomial has degree n. Therefore, the sum of the geometric multiplicities equals n.

Now suppose that the sum of the geometric multiplicities equals n. As above, this forces the sum of the algebraic multiplicities to equal n as well. As the algebraic multiplicities are all greater than or equal to the geometric multiplicities in any case, this implies that they are in fact equal.

Finally, suppose that the third condition is satisfied. Then the sum of the geometric multiplicities equals n. Suppose that the distinct eigenvectors are $\lambda_1, \lambda_2, \ldots, \lambda_k$, and that $\mathcal{B}_i$ is a basis for the λ_i-eigenspace, which we call V_i. We claim that the collection $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ of all vectors in all of the eigenspace bases $\mathcal{B}_i$ is linearly independent. Consider the vector equation

$$0 = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n.$$

Grouping the eigenvectors with the same eigenvalues, this sum has the form

$$0 = (\text{something in } V_1) + (\text{something in } V_2) + \cdots + (\text{something in } V_k).$$

Since eigenvectors with distinct eigenvalues are linearly independent, each "something in V_i" is equal to zero. But this implies that all coefficients $c_1, c_2, \ldots, c_n$ are equal to zero, since the vectors in each $\mathcal{B}_i$ are linearly independent. Therefore, A has n linearly independent eigenvectors, so it is diagonalizable. □

The first part of the third statement simply says that the characteristic polynomial of A factors completely into linear polynomials over the real numbers: in other words, there are no complex (non-real) roots. The second part of the third statement says in particular that for any diagonalizable matrix, the algebraic and geometric multiplicities coincide.

> Let A be a square matrix and let λ be an eigenvalue of A. If the algebraic multiplicity of λ does not equal the geometric multiplicity, then A is not diagonalizable.

The examples at the beginning of this subsection illustrate the theorem. Here we give some general consequences for diagonalizability of 2×2 and 3×3 matrices.

Diagonalizability of 2×2 Matrices. Let A be a 2×2 matrix. There are four cases:

1. A has two different eigenvalues. In this case, each eigenvalue has algebraic and geometric multiplicity equal to one. This implies A is diagonalizable. For example:
$$A=\begin{pmatrix}1&7\\0&2\end{pmatrix}.$$

2. A has one eigenvalue λ of algebraic and geometric multiplicity 2. To say that the geometric multiplicity is 2 means that $\mathrm{Nul}(A-\lambda I_2)=\mathbf{R}^2$, i.e., that every vector in $\mathbf{R}^2$ is in the null space of $A-\lambda I_2$. This implies that $A-\lambda I_2$ is the zero matrix, so that A is the diagonal matrix λI_2. In particular, A is diagonalizable. For example:
$$A=\begin{pmatrix}1&0\\0&1\end{pmatrix}.$$

3. A has one eigenvalue λ of algebraic multiplicity 2 and geometric multiplicity 1. In this case, A is not diagonalizable, by part 3 of the theorem. For example, a shear:
$$A=\begin{pmatrix}1&1\\0&1\end{pmatrix}.$$

4. A has no eigenvalues. This happens when the characteristic polynomial has no real roots. In particular, A is not diagonalizable. For example, a rotation:
$$A=\begin{pmatrix}1&-1\\1&1\end{pmatrix}.$$

Diagonalizability of 3×3 Matrices. Let A be a 3×3 matrix. We can analyze the diagonalizability of A on a case-by-case basis, as in the previous example.

1. A has three different eigenvalues. In this case, each eigenvalue has algebraic and geometric multiplicity equal to one. This implies A is diagonalizable. For example:
$$A=\begin{pmatrix}1&7&4\\0&2&3\\0&0&-1\end{pmatrix}.$$

2. A has two distinct eigenvalues λ_1,λ_2. In this case, one has algebraic multiplicity one and the other has algebraic multiplicity two; after reordering, we can assume λ_1 has multiplicity 1 and λ_2 has multiplicity 2. This implies that λ_1 has geometric multiplicity 1, so *A is diagonalizable if and only if the λ_2-eigenspace is a plane.* For example:
$$A=\begin{pmatrix}1&7&4\\0&2&0\\0&0&2\end{pmatrix}.$$

On the other hand, if the geometric multiplicity of λ_2 is 1, then A is not diagonalizable. For example:

$$A = \begin{pmatrix} 1 & 7 & 4 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

3. A has only one eigenvalue λ. If the algebraic multiplicity of λ is 1, then A is not diagonalizable. This happens when the characteristic polynomial has two complex (non-real) roots. For example:

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Otherwise, the algebraic multiplicity of λ is equal to 3. In this case, if the geometric multiplicity is 1:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

or 2:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

then A is not diagonalizable. If the geometric multiplicity is 3, then $\mathrm{Nul}(A - \lambda I_3) = \mathbf{R}^3$, so that $A - \lambda I_3$ is the zero matrix, and hence $A = \lambda I_3$. Therefore, in this case A is necessarily diagonal, as in:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Similarity and multiplicity Recall from this fact in Section 5.3 that similar matrices have the same eigenvalues. It turns out that *both* notions of multiplicity of an eigenvalue are preserved under similarity.

Theorem. *Let A and B be similar $n \times n$ matrices, and let λ be an eigenvalue of A and B. Then:*

1. *The algebraic multiplicity of λ is the same for A and B.*

2. *The geometric multiplicity of λ is the same for A and B.*

Proof. Since A and B have the same characteristic polynomial, the multiplicity of λ as a root of the characteristic polynomial is the same for both matrices, which proves the first statement. For the second, suppose that $A = CBC^{-1}$ for an invertible matrix C. By this fact in Section 5.3, the matrix C takes eigenvectors of B to eigenvectors of A, both with eigenvalue λ.

Let $\{v_1, v_2, \ldots, v_k\}$ be a basis of the λ-eigenspace of B. We claim that $\{Cv_1, Cv_2, \ldots, Cv_k\}$ is linearly independent. Suppose that

$$c_1 Cv_1 + c_2 Cv_2 + \cdots + c_k Cv_k = 0.$$

Regrouping, this means

$$C\bigl(c_1 v_1 + c_2 v_2 + \cdots + c_k v_k\bigr) = 0.$$

By the invertible matrix theorem in Section 5.1, the null space of C is trivial, so this implies

$$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k = 0.$$

Since $v_1, v_2, \ldots, v_k$ are linearly independent, we get $c_1 = c_2 = \cdots = c_k = 0$, as desired.

By the previous paragraph, the dimension of the λ-eigenspace of A is greater than or equal to the dimension of the λ-eigenspace of B. By symmetry (B is similar to A as well), the dimensions are equal, so the geometric multiplicities coincide. □

For instance, the four matrices in this example are not similar to each other, because the algebraic and/or geometric multiplicities of the eigenvalues do not match up. Or, combined with the above theorem, we see that a diagonalizable matrix cannot be similar to a non-diagonalizable one, because the algebraic and geometric multiplicities of such matrices cannot both coincide.

Example. Continuing with this example, let

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

This is a diagonalizable matrix that is similar to

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \quad \text{using the matrix} \quad C = \begin{pmatrix} 1 & 0 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}.$$

The 1-eigenspace of D is the xy-plane, and the 2-eigenspace is the z-axis. The matrix C takes the xy-plane to the 1-eigenspace of A, which is again a plane, and the z-axis to the 2-eigenspace of A, which is again a line. This shows that the geometric multiplicities of A and D coincide.

Use this link to view the online demo

The matrix C takes the xy-plane to the 1-eigenspace of A (the grid) and the z-axis to the 2-eigenspace (the green line).

The converse of the theorem is false: there exist matrices whose eigenvectors have the same algebraic and geometric multiplicities, but which are not similar. See the example below. However, for 2×2 and 3×3 matrices whose characteristic polynomial has no complex (non-real) roots, the converse of the theorem is true. (We will handle the case of complex roots in Section 5.5.)

Example (Matrices that look similar but are not)**.** Show that the matrices

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

have the same eigenvalues with the same algebraic and geometric multiplicities, but are not similar.

Solution. These matrices are upper-triangular. They both have characteristic polynomial $f(\lambda) = \lambda^4$, so they both have one eigenvalue 0 with algebraic multiplicity 4. The 0-eigenspace is the null space, which has dimension 2 in each case because A and B have two columns without pivots. Hence 0 has geometric multiplicity 2 in each case.

To show that A and B are not similar, we note that

$$A^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B^2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

as the reader can verify. If $A = CBC^{-1}$ then by this important note, we have

$$A^2 = CB^2C^{-1} = C0C^{-1} = 0,$$

which is not the case.

On the other hand, suppose that A and B are *diagonalizable* matrices with the same characteristic polynomial. Since the geometric multiplicities of the eigenvalues coincide with the algebraic multiplicities, which are the same for A and B, we conclude that there exist n linearly independent eigenvectors of each matrix, all of which have the same eigenvalues. This shows that A and B are both similar to the same diagonal matrix. Using the transitivity property of similar matrices, this shows:

> *Diagonalizable* matrices are similar if and only if they have the same characteristic polynomial, or equivalently, the same eigenvalues with the same algebraic multiplicities.

Example. Show that the matrices

$$A = \begin{pmatrix} 1 & 7 & 2 \\ 0 & -1 & 3 \\ 0 & 0 & 4 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 4 & 0 \\ -5 & -4 & -1 \end{pmatrix}$$

are similar.

Solution. Both matrices have the three distinct eigenvalues $1, -1, 4$. Hence they are both diagonalizable, and are similar to the diagonal matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

By the transitivity property of similar matrices, this implies that A and B are similar to each other.

Example (Diagonal matrices with the same entries are similar)**.** Any two diagonal matrices with the same diagonal entries (possibly in a different order) are similar to each other. Indeed, such matrices have the same characteristic polynomial. We saw this phenomenon in this example, where we noted that

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1}.$$

5.5 Complex Eigenvalues

Objectives

1. Learn to find complex eigenvalues and eigenvectors of a matrix.
2. Learn to recognize a rotation-scaling matrix, and compute by how much the matrix rotates and scales.
3. Understand the geometry of 2×2 and 3×3 matrices with a complex eigenvalue.
4. *Recipes:* a 2×2 matrix with a complex eigenvalue is similar to a rotation-scaling matrix, the eigenvector trick for 2×2 matrices.
5. *Pictures:* the geometry of matrices with a complex eigenvalue.
6. *Theorems:* the rotation-scaling theorem, the block diagonalization theorem.
7. *Vocabulary word:* **rotation-scaling matrix**.

In Section 5.4, we saw that an $n \times n$ matrix whose characteristic polynomial has n distinct real roots is *diagonalizable*: it is similar to a diagonal matrix, which is much simpler to analyze. The other possibility is that a matrix has *complex* roots, and that is the focus of this section. It turns out that such a matrix is similar (in the 2×2 case) to a *rotation-scaling matrix*, which is also relatively easy to understand.

In a certain sense, this entire section is analogous to Section 5.4, with rotation-scaling matrices playing the role of diagonal matrices.

See Appendix A for a review of the complex numbers.

5.5.1 Matrices with Complex Eigenvalues

As a consequence of the fundamental theorem of algebra as applied to the characteristic polynomial, we see that:

> Every $n \times n$ matrix has exactly n complex eigenvalues, counted with multiplicity.

We can compute a corresponding (complex) eigenvector in exactly the same way as before: by row reducing the matrix $A - \lambda I_n$. Now, however, we have to do arithmetic with complex numbers.

Example (A 2×2 matrix)**.** Find the complex eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Solution. The characteristic polynomial of A is

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A) = \lambda^2 - 2\lambda + 2.$$

The roots of this polynomial are

$$\lambda = \frac{2 \pm \sqrt{4-8}}{2} = 1 \pm i.$$

First we compute an eigenvector for $\lambda = 1 + i$. We have

$$A - (1+i)I_2 = \begin{pmatrix} 1-(1+i) & -1 \\ 1 & 1-(1+i) \end{pmatrix} = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}.$$

Now we row reduce, noting that the second row is i times the first:

$$\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \xrightarrow{R_2 = R_2 - iR_1} \begin{pmatrix} -i & -1 \\ 0 & 0 \end{pmatrix} \xrightarrow{R_1 = R_1 \div -i} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}.$$

The parametric form is $x = iy$, so that an eigenvector is $v_1 = \binom{i}{1}$. Next we compute an eigenvector for $\lambda = 1-i$. We have

$$A-(1-i)I_2 = \begin{pmatrix} 1-(1-i) & -1 \\ 1 & 1-(1-i) \end{pmatrix} = \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}.$$

Now we row reduce, noting that the second row is $-i$ times the first:

$$\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \xrightarrow{R_2=R_2+iR_1} \begin{pmatrix} i & -1 \\ 0 & 0 \end{pmatrix} \xrightarrow{R_1=R_1\div i} \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}.$$

The parametric form is $x = -iy$, so that an eigenvector is $v_2 = \binom{-i}{1}$.

We can verify our answers:

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} i-1 \\ i+1 \end{pmatrix} = (1+i)\begin{pmatrix} i \\ 1 \end{pmatrix}$$
$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -i \\ 1 \end{pmatrix} = \begin{pmatrix} -i-1 \\ -i+1 \end{pmatrix} = (1-i)\begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

Example (A 3 × 3 matrix)**.** Find the eigenvalues and eigenvectors, real and complex, of the matrix

$$A = \begin{pmatrix} 4/5 & -3/5 & 0 \\ 3/5 & 4/5 & 0 \\ 1 & 2 & 2 \end{pmatrix}.$$

Solution. We compute the characteristic polynomial by expanding cofactors along the third row:

$$f(\lambda) = \det\begin{pmatrix} 4/5-\lambda & -3/5 & 0 \\ 3/5 & 4-5-\lambda & 0 \\ 1 & 2 & 2-\lambda \end{pmatrix} = (2-\lambda)\left(\lambda^2 - \frac{8}{5}\lambda + 1\right).$$

This polynomial has one real root at 2, and two complex roots at

$$\lambda = \frac{8/5 \pm \sqrt{64/25-4}}{2} = \frac{4\pm 3i}{5}.$$

Therefore, the eigenvalues are

$$\lambda = 2, \quad \frac{4+3i}{5}, \quad \frac{4-3i}{5}.$$

We eyeball that $v_1 = e_3$ is an eigenvector with eigenvalue 2, since the third column is $2e_3$.

Next we find an eigenvector with eigenvaluue $(4+3i)/5$. We have

$$A - \frac{4+3i}{5}I_3 = \begin{pmatrix} -3i/5 & -3/5 & 0 \\ 3/5 & -3i/5 & 0 \\ 1 & 2 & 2-(4+3i)/5 \end{pmatrix} \xrightarrow[R_2=R_2\times 5/3]{R_1=R_1\times -5/3} \begin{pmatrix} i & 1 & 0 \\ 1 & -i & 0 \\ 1 & 2 & \frac{6-3i}{5} \end{pmatrix}.$$

We row reduce, noting that the second row is $-i$ times the first:

$$\begin{pmatrix} i & 1 & 0 \\ 1 & -i & 0 \\ 1 & 2 & \frac{6-3i}{5} \end{pmatrix} \xrightarrow{R_2=R_2+iR_1} \begin{pmatrix} i & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 2 & \frac{6-3i}{5} \end{pmatrix}$$

$$\xrightarrow{R_3=R_3+iR_1} \begin{pmatrix} i & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 2+i & \frac{6-3i}{5} \end{pmatrix}$$

$$\xrightarrow{R_2\longleftrightarrow R_3} \begin{pmatrix} i & 1 & 0 \\ 0 & 2+i & \frac{6-3i}{5} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[R_2=R_2\div(2+i)]{R_1=R_1\div i} \begin{pmatrix} 1 & -i & 0 \\ 0 & 1 & \frac{9-12i}{25} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{R_1=R_1+iR_2} \begin{pmatrix} 1 & 0 & \frac{12+9i}{25} \\ 0 & 1 & \frac{9-12i}{25} \\ 0 & 0 & 0 \end{pmatrix}.$$

The free variable is z; the parametric form of the solution is

$$\begin{cases} x = -\dfrac{12+9i}{25}z \\ y = -\dfrac{9-12i}{25}z. \end{cases}$$

Taking $z=25$ gives the eigenvector

$$v_2 = \begin{pmatrix} -12-9i \\ -9+12i \\ 25 \end{pmatrix}.$$

A similar calculation (replacing all occurences of i by $-i$) shows that an eigenvector with eigenvalue $(4-3i)/5$ is

$$v_3 = \begin{pmatrix} -12+9i \\ -9-12i \\ 25 \end{pmatrix}.$$

We can verify our calculations:

$$\begin{pmatrix} 4/5 & -3/5 & 0 \\ 3/5 & 4/5 & 0 \\ 1 & 2 & 2 \end{pmatrix}\begin{pmatrix} -12+9i \\ -9-12i \\ 25 \end{pmatrix} = \begin{pmatrix} -21/5+72i/5 \\ -72/5-21i/5 \\ 20-15i \end{pmatrix} = \frac{4+3i}{5}\begin{pmatrix} -12+9i \\ -9-12i \\ 25 \end{pmatrix}$$

$$\begin{pmatrix} 4/5 & -3/5 & 0 \\ 3/5 & 4/5 & 0 \\ 1 & 2 & 2 \end{pmatrix}\begin{pmatrix} -12-9i \\ -9+12i \\ 25 \end{pmatrix} = \begin{pmatrix} -21/5-72i/5 \\ -72/5+21i/5 \\ 20+15i \end{pmatrix} = \frac{4-3i}{5}\begin{pmatrix} -12-9i \\ -9+12i \\ 25 \end{pmatrix}.$$

If A is a matrix with real entries, then its characteristic polynomial has real coefficients, so this note implies that its complex eigenvalues come in conjugate pairs. In the first example, we notice that

$$1+i \text{ has an eigenvector } v_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}$$
$$1-i \text{ has an eigenvector } v_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

In the second example,

$$\frac{4+3i}{5} \text{ has an eigenvector } v_1 = \begin{pmatrix} -12-9i \\ -9+12i \\ 25 \end{pmatrix}$$
$$\frac{4-3i}{5} \text{ has an eigenvector } v_2 = \begin{pmatrix} -12+9i \\ -9-12i \\ 25 \end{pmatrix}$$

In these cases, an eigenvector for the conjugate eigenvalue is simply the *conjugate eigenvector* (the eigenvector obtained by conjugating each entry of the first eigenvector). This is always true. Indeed, if $Av = \lambda v$ then

$$A\bar{v} = \overline{Av} = \overline{\lambda v} = \bar{\lambda}\bar{v},$$

which exactly says that $\bar{v}$ is an eigenvector of A with eigenvalue $\bar{\lambda}$.

Let A be a matrix with real entries. If

$$\lambda \text{ is a complex eigenvalue with eigenvector } v,$$
$$\text{then } \bar{\lambda} \text{ is a complex eigenvalue with eigenvector } \bar{v}.$$

In other words, both eigenvalues and eigenvectors come in conjugate pairs.

Since it can be tedious to divide by complex numbers while row reducing, it is useful to learn the following trick, which works equally well for matrices with real entries.

Eigenvector Trick for 2×2 Matrices. Let A be a 2×2 matrix, and let λ be a (real or complex) eigenvalue. Then

$$A - \lambda I_2 = \begin{pmatrix} z & w \\ \star & \star \end{pmatrix} \implies \begin{pmatrix} -w \\ z \end{pmatrix} \text{ is an eigenvector with eigenvalue } \lambda,$$

assuming the first row of $A - \lambda I_2$ is nonzero.

Indeed, since λ is an eigenvalue, we know that $A - \lambda I_2$ is not an invertible matrix. It follows that the rows are collinear (otherwise the determinant is nonzero), so that the second row is automatically a (complex) multiple of the first:

$$\begin{pmatrix} z & w \\ \star & \star \end{pmatrix} = \begin{pmatrix} z & w \\ cz & cw \end{pmatrix}.$$

It is obvious that $\binom{-w}{z}$ is in the null space of this matrix, as is $\binom{w}{-z}$, for that matter. Note that we never had to compute the second row of $A - \lambda I_2$, let alone row reduce!

Example (A 2×2 matrix, the easy way)**.** Find the complex eigenvalues and eigenvectors of the matrix

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Solution. Since the characteristic polynomial of a 2×2 matrix A is $f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(A)$, its roots are

$$\lambda = \frac{\mathrm{Tr}(A) \pm \sqrt{\mathrm{Tr}(A)^2 - 4\det(A)}}{2} = \frac{2 \pm \sqrt{4-8}}{2} = 1 \pm i.$$

To find an eigenvector with eigenvalue $1 + i$, we compute

$$A - (1+i)I_2 = \begin{pmatrix} -i & -1 \\ \star & \star \end{pmatrix} \xrightarrow{\text{eigenvector}} v_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}.$$

The eigenvector for the conjugate eigenvalue is the complex conjugate:

$$v_2 = \overline{v}_1 = \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

In this example we found the eigenvectors $\binom{i}{1}$ and $\binom{-i}{1}$ for the eigenvalues $1+i$ and $1-i$, respectively, but in this example we found the eigenvectors $\binom{1}{-i}$ and $\binom{1}{i}$ for the same eigenvalues of the same matrix. These vectors do not look like multiples of each other at first—but since we now have complex numbers at our disposal, we can see that they actually are multiples:

$$-i\begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -i \end{pmatrix} \qquad i\begin{pmatrix} -i \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

5.5.2 Rotation-Scaling Matrices

The most important examples of matrices with complex eigenvalues are rotation-scaling matrices, i.e., scalar multiples of rotation matrices.

Definition. A **rotation-scaling matrix** is a 2×2 matrix of the form

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

where a and b are real numbers, not both equal to zero.

The following proposition justifies the name.

Proposition. *Let*

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

be a rotation-scaling matrix. Then:

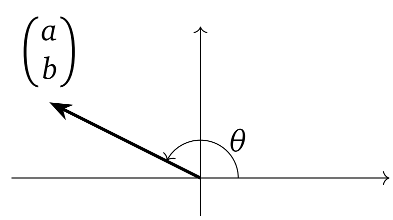
1. *A is a product of a rotation matrix*

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad \text{with a scaling matrix} \quad \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}.$$

2. *The scaling factor r is*

$$r = \sqrt{\det(A)} = \sqrt{a^2 + b^2}.$$

3. *The rotation angle θ is the counterclockwise angle from the positive x-axis to the vector $\binom{a}{b}$:*

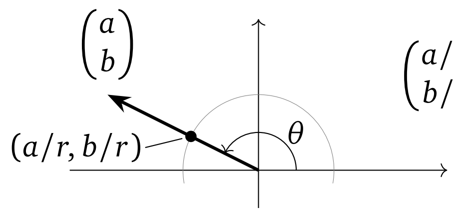


4. *The eigenvalues of A are $\lambda = a \pm bi$.*

Proof. Set $r = \sqrt{\det(A)} = \sqrt{a^2 + b^2}$. The point $(a/r, b/r)$ has the property that

$$\left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2 = \frac{a^2 + b^2}{r^2} = 1.$$

In other words $(a/r, b/r)$ lies on the unit circle. Therefore, it has the form $(\cos\theta, \sin\theta)$, where θ is the counterclockwise angle from the positive x-axis to the vector $\binom{a/r}{b/r}$, or since it is on the same line, to $\binom{a}{b}$:



$$\begin{pmatrix} a/r \\ b/r \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}.$$

It follows that

$$A = r\begin{pmatrix} a/r & -b/r \\ b/r & a/r \end{pmatrix} = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

as desired.

For the last statement, we compute the eigenvalues of A as the roots of the characteristic polynomial:

$$\lambda = \frac{\mathrm{Tr}(A) \pm \sqrt{\mathrm{Tr}(A)^2 - 4\det(A)}}{2} = \frac{2a \pm \sqrt{4a^2 - 4(a^2+b^2)}}{2} = a \pm bi. \qquad \square$$

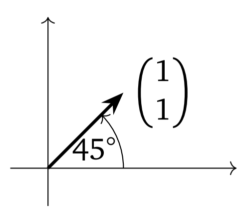
Geometrically, a rotation-scaling matrix does exactly what the name says: it rotates and scales (in either order).

Example (A rotation-scaling matrix)**.** What does the matrix

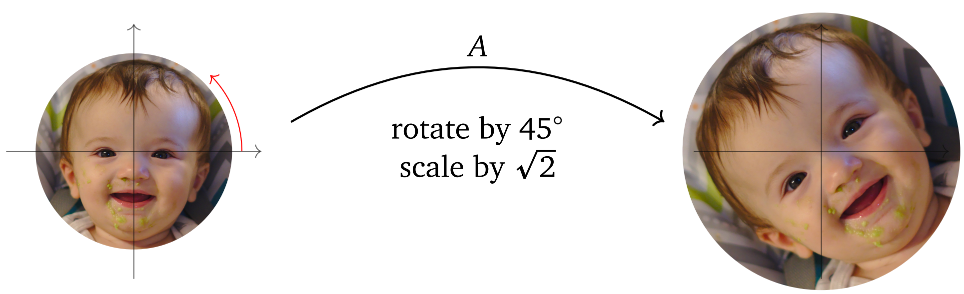
$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

do geometrically?

Solution. This is a rotation-scaling matrix with $a = b = 1$. Therefore, it scales by a factor of $\sqrt{\det(A)} = \sqrt{2}$ and rotates counterclockwise by 45°:



Here is a picture of A:



An interactive figure is included below.

Use this link to view the online demo

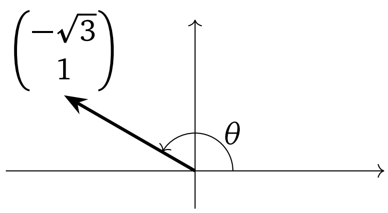
Multiplication by the matrix A rotates the plane by 45° and dilates by a factor of $\sqrt{2}$. Move the input vector x to see how the output vector b changes.

Example (A rotation-scaling matrix)**.** What does the matrix

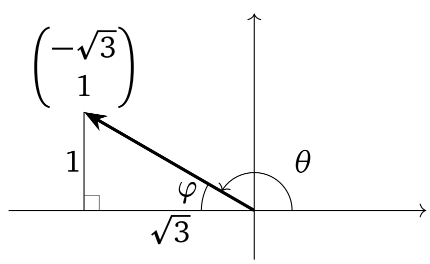
$$A = \begin{pmatrix} -\sqrt{3} & -1 \\ 1 & -\sqrt{3} \end{pmatrix}$$

do geometrically?

Solution. This is a rotation-scaling matrix with $a = -\sqrt{3}$ and $b = 1$. Therefore, it scales by a factor of $\sqrt{\det(A)} = \sqrt{3+1} = 2$ and rotates counterclockwise by the angle θ in the picture:



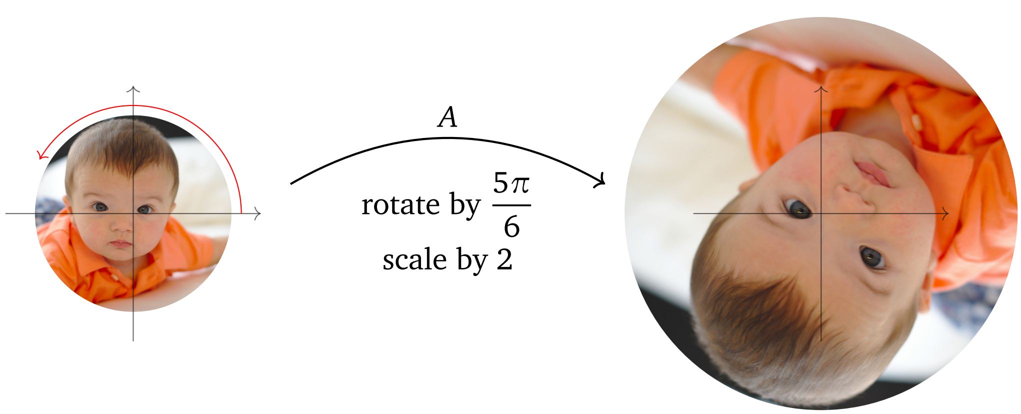
To compute this angle, we do a bit of trigonometry:



$$\varphi = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6}$$

$$\theta = \pi - \varphi = \frac{5\pi}{6}.$$

Therefore, A rotates counterclockwise by $5\pi/6$ and scales by a factor of 2.

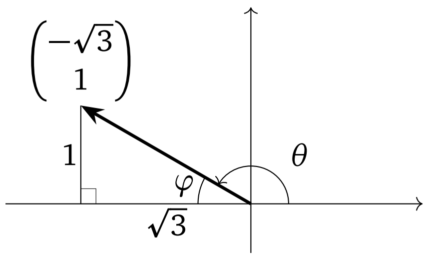


An interactive figure is included below.

Use this link to view the online demo

Multiplication by the matrix A rotates the plane by $5\pi/6$ and dilates by a factor of 2. Move the input vector x to see how the output vector b changes.

The matrix in the second example has second column $\binom{-\sqrt{3}}{1}$, which is rotated counterclockwise from the positive x-axis by an angle of $5\pi/6$. This rotation angle is *not* equal to $\tan^{-1}\bigl(1/(-\sqrt{3})\bigr) = -\frac{\pi}{6}$. The problem is that arctan always outputs values between $-\pi/2$ and $\pi/2$: it does not account for points in the second or third quadrants. This is why we drew a triangle and used its (positive) edge lengths to compute the angle φ:



$$\varphi = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6}$$
$$\theta = \pi - \varphi = \frac{5\pi}{6}.$$

Alternatively, we could have observed that $\binom{-\sqrt{3}}{1}$ lies in the second quadrant, so that the angle θ in question is

$$\theta = \tan^{-1}\left(\frac{1}{-\sqrt{3}}\right) + \pi.$$

When finding the rotation angle of a vector $\binom{a}{b}$, do not blindly compute $\tan^{-1}(b/a)$, since this will give the wrong answer when $\binom{a}{b}$ is in the second or third quadrant. Instead, draw a picture.

5.5.3 Geometry of 2×2 Matrices with a Complex Eigenvalue

Let A be a 2×2 matrix with a complex, non-real eigenvalue λ. Then A also has the eigenvalue $\overline{\lambda} \neq \lambda$. In particular, A has distinct eigenvalues, so it is diagonalizable using the complex numbers. We often like to think of our matrices as describing transformations of $\mathbf{R}^n$ (as opposed to $\mathbf{C}^n$). Because of this, the following construction is useful. It gives something like a diagonalization, except that all matrices involved have *real* entries.

Rotation-Scaling Theorem. *Let A be a 2×2 real matrix with a complex (non-real) eigenvalue λ, and let v be an eigenvector. Then $A = CBC^{-1}$ for*

$$C = \begin{pmatrix} | & | \\ \mathrm{Re}(v) & \mathrm{Im}(v) \\ | & | \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \mathrm{Re}(\lambda) & \mathrm{Im}(\lambda) \\ -\mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{pmatrix}.$$

In particular, A is similar to a rotation-scaling matrix that scales by a factor of $|\lambda| = \sqrt{\det(B)}$.

Proof. First we need to show that $\mathrm{Re}(v)$ and $\mathrm{Im}(v)$ are linearly independent, since otherwise C is not invertible. If not, then there exist real numbers x, y, not both equal to zero, such that $x\,\mathrm{Re}(v) + y\,\mathrm{Im}(v) = 0$. Then

$$\begin{aligned}(y+ix)v &= (y+ix)\big(\mathrm{Re}(v) + i\,\mathrm{Im}(v)\big) \\ &= y\,\mathrm{Re}(v) - x\,\mathrm{Im}(v) + (x\,\mathrm{Re}(v) + y\,\mathrm{Im}(v))\,i \\ &= y\,\mathrm{Re}(v) - x\,\mathrm{Im}(v).\end{aligned}$$

Now, $(y+ix)v$ is also an eigenvector of A with eigenvalue λ, as it is a scalar multiple of v. But we just showed that $(y+ix)v$ is a vector with real entries, and any *real* eigenvector of a real matrix has a real eigenvalue. Therefore, $\mathrm{Re}(v)$ and $\mathrm{Im}(v)$ must be linearly independent after all.

Let $\lambda = a + bi$ and $v = \binom{x+yi}{z+wi}$. We observe that

$$\begin{aligned}Av = \lambda v = (a+bi)\begin{pmatrix} x+yi \\ z+wi \end{pmatrix} &= \begin{pmatrix} (ax-by)+(ay+bx)i \\ (az-bw)+(aw+bz)i \end{pmatrix} \\ &= \begin{pmatrix} ax-by \\ az-bw \end{pmatrix} + i\begin{pmatrix} ay+bx \\ aw+bz \end{pmatrix}.\end{aligned}$$

On the other hand, we have

$$A\left(\begin{pmatrix} x \\ z \end{pmatrix} + i\begin{pmatrix} y \\ w \end{pmatrix}\right) = A\begin{pmatrix} x \\ z \end{pmatrix} + iA\begin{pmatrix} y \\ w \end{pmatrix} = A\,\mathrm{Re}(v) + iA\,\mathrm{Im}(v).$$

Matching real and imaginary parts gives

$$A\,\mathrm{Re}(v) = \begin{pmatrix} ax-by \\ az-bw \end{pmatrix} \qquad A\,\mathrm{Im}(v) = \begin{pmatrix} ay+bx \\ aw+bz \end{pmatrix}.$$

Now we compute $CBC^{-1}\,\mathrm{Re}(v)$ and $CBC^{-1}\,\mathrm{Im}(v)$. Since $Ce_1 = \mathrm{Re}(v)$ and

$Ce_2 = \mathrm{Im}(v)$, we have $C^{-1}\,\mathrm{Re}(v) = e_1$ and $C^{-1}\,\mathrm{Im}(v) = e_2$, so

$$\begin{aligned} CBC^{-1}\,\mathrm{Re}(v) &= CBe_1 = C\begin{pmatrix} a \\ -b \end{pmatrix} = a\,\mathrm{Re}(v) - b\,\mathrm{Im}(v) \\ &= a\begin{pmatrix} x \\ z \end{pmatrix} - b\begin{pmatrix} y \\ w \end{pmatrix} = \begin{pmatrix} ax - by \\ az - bw \end{pmatrix} = A\mathrm{Re}(v) \\ CBC^{-1}\,\mathrm{Im}(v) &= CBe_2 = C\begin{pmatrix} b \\ a \end{pmatrix} = b\,\mathrm{Re}(v) + a\,\mathrm{Im}(v) \\ &= b\begin{pmatrix} x \\ z \end{pmatrix} + a\begin{pmatrix} y \\ w \end{pmatrix} = \begin{pmatrix} ay + bx \\ aw + bz \end{pmatrix} = A\mathrm{Im}(v). \end{aligned}$$

Therefore, $A\mathrm{Re}(v) = CBC^{-1}\,\mathrm{Re}(v)$ and $A\mathrm{Im}(v) = CBC^{-1}\,\mathrm{Im}(v)$.

Since $\mathrm{Re}(v)$ and $\mathrm{Im}(v)$ are linearly independent, they form a basis for $\mathbf{R}^2$. Let w be any vector in $\mathbf{R}^2$, and write $w = c\,\mathrm{Re}(v) + d\,\mathrm{Im}(v)$. Then

$$\begin{aligned} Aw &= A\big(c\,\mathrm{Re}(v) + d\,\mathrm{Im}(v)\big) \\ &= cA\mathrm{Re}(v) + dA\mathrm{Im}(v) \\ &= cCBC^{-1}\,\mathrm{Re}(v) + dCBC^{-1}\,\mathrm{Im}(v) \\ &= CBC^{-1}\big(c\,\mathrm{Re}(v) + d\,\mathrm{Im}(v)\big) \\ &= CBC^{-1}w. \end{aligned}$$

This proves that $A = CBC^{-1}$. □

Here Re and Im denote the real and imaginary parts, respectively:

$$\mathrm{Re}(a+bi) = a \quad \mathrm{Im}(a+bi) = b \quad \mathrm{Re}\begin{pmatrix} x+yi \\ z+wi \end{pmatrix} = \begin{pmatrix} x \\ z \end{pmatrix} \quad \mathrm{Im}\begin{pmatrix} x+yi \\ z+wi \end{pmatrix} = \begin{pmatrix} y \\ w \end{pmatrix}.$$

The rotation-scaling matrix in question is the matrix

$$B = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{with} \quad a = \mathrm{Re}(\lambda),\; b = -\mathrm{Im}(\lambda).$$

Geometrically, the rotation-scaling theorem says that a 2×2 matrix with a complex eigenvalue behaves similarly to a rotation-scaling matrix. See this important note in Section 5.3.

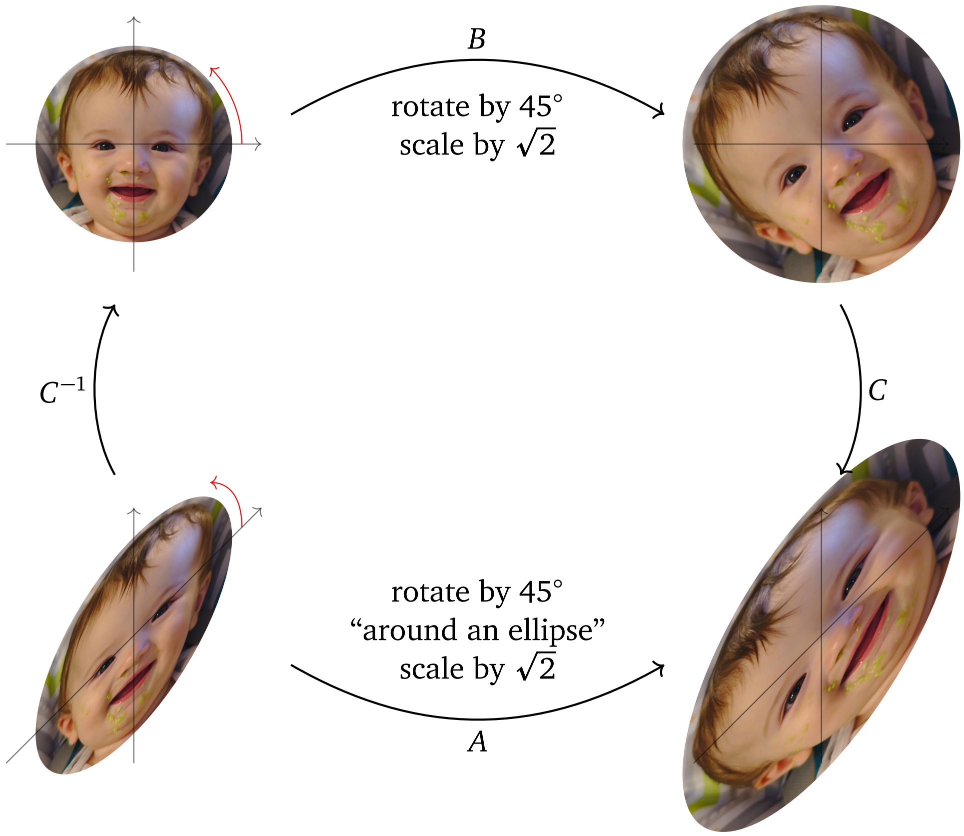
One should regard the rotation-scaling theorem as a close analogue of the diagonalization theorem in Section 5.4, with a rotation-scaling matrix playing the role of a diagonal matrix. Before continuing, we restate the theorem as a recipe:

Recipe: A 2×2 matrix with a complex eigenvalue. Let A be a 2×2 real matrix.

1. Compute the characteristic polynomial
$$f(\lambda) = \lambda^2 - \text{Tr}(A)\lambda + \det(A),$$
then compute its roots using the quadratic formula.
2. If the eigenvalues are complex, choose *one* of them, and call it λ.
3. Find a corresponding (complex) eigenvalue v using the trick.
4. Then $A = CBC^{-1}$ for
$$C = \begin{pmatrix} | & | \\ \text{Re}(v) & \text{Im}(v) \\ | & | \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \text{Re}(\lambda) & \text{Im}(\lambda) \\ -\text{Im}(\lambda) & \text{Re}(\lambda) \end{pmatrix}.$$
This scales by a factor of $|\lambda|$.

Example. What does the matrix
$$A = \begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}$$
do geometrically?

Solution. The eigenvalues of A are
$$\lambda = \frac{\text{Tr}(A) \pm \sqrt{\text{Tr}(A)^2 - 4\det(A)}}{2} = \frac{2 \pm \sqrt{4-8}}{2} = 1 \pm i.$$
We choose the eigenvalue $\lambda = 1 - i$ and find a corresponding eigenvector, using the trick:
$$A - (1-i)I_2 = \begin{pmatrix} 1+i & -1 \\ \star & \star \end{pmatrix} \xrightarrow{\text{eigenvector}} v = \begin{pmatrix} 1 \\ 1+i \end{pmatrix}.$$
According to the rotation-scaling theorem, we have $A = CBC^{-1}$ for
$$C = \left(\text{Re}\begin{pmatrix} 1 \\ 1+i \end{pmatrix} \ \ \text{Im}\begin{pmatrix} 1 \\ 1+i \end{pmatrix} \right) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$
$$B = \begin{pmatrix} \text{Re}(\lambda) & \text{Im}(\lambda) \\ -\text{Im}(\lambda) & \text{Re}(\lambda) \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$
The matrix B is the rotation-scaling matrix in the above example: it rotates counterclockwise by an angle of 45° and scales by a factor of $\sqrt{2}$. The matrix A does the same thing, but with respect to the $\text{Re}(v), \text{Im}(v)$-coordinate system:



To summarize:

- B rotates around the circle centered at the origin and passing through e_1 and e_2, in the direction from e_1 to e_2, then scales by $\sqrt{2}$.
- A rotates around the ellipse centered at the origin and passing through $\mathrm{Re}(v)$ and $\mathrm{Im}(v)$, in the direction from $\mathrm{Re}(v)$ to $\mathrm{Im}(v)$, then scales by $\sqrt{2}$.

The reader might want to refer back to this example in Section 5.3.

Use this link to view the online demo

The geometric action of A and B on the plane. Click "multiply" to multiply the colored points by B on the left and A on the right.

If instead we had chosen $\overline{\lambda} = 1+i$ as our eigenvalue, then we would have found the eigenvector $\overline{v} = \binom{1}{1-i}$. In this case we would have $A = C'B'(C')^{-1}$, where

$$C' = \begin{pmatrix} \mathrm{Re}\begin{pmatrix} 1 \\ 1-i \end{pmatrix} & \mathrm{Im}\begin{pmatrix} 1 \\ 1-i \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

$$B' = \begin{pmatrix} \mathrm{Re}(\overline{\lambda}) & \mathrm{Im}(\overline{\lambda}) \\ -\mathrm{Im}(\overline{\lambda}) & \mathrm{Re}(\overline{\lambda}) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

So, A is also similar to a *clockwise* rotation by 45°, followed by a scale by $\sqrt{2}$.

Example. What does the matrix

$$A = \begin{pmatrix} -\sqrt{3}+1 & -2 \\ 1 & -\sqrt{3}-1 \end{pmatrix}$$

do geometrically?

Solution. The eigenvalues of A are

$$\lambda = \frac{\mathrm{Tr}(A) \pm \sqrt{\mathrm{Tr}(A)^2 - 4\det(A)}}{2} = \frac{-2\sqrt{3} \pm \sqrt{12-16}}{2} = -\sqrt{3} \pm i.$$

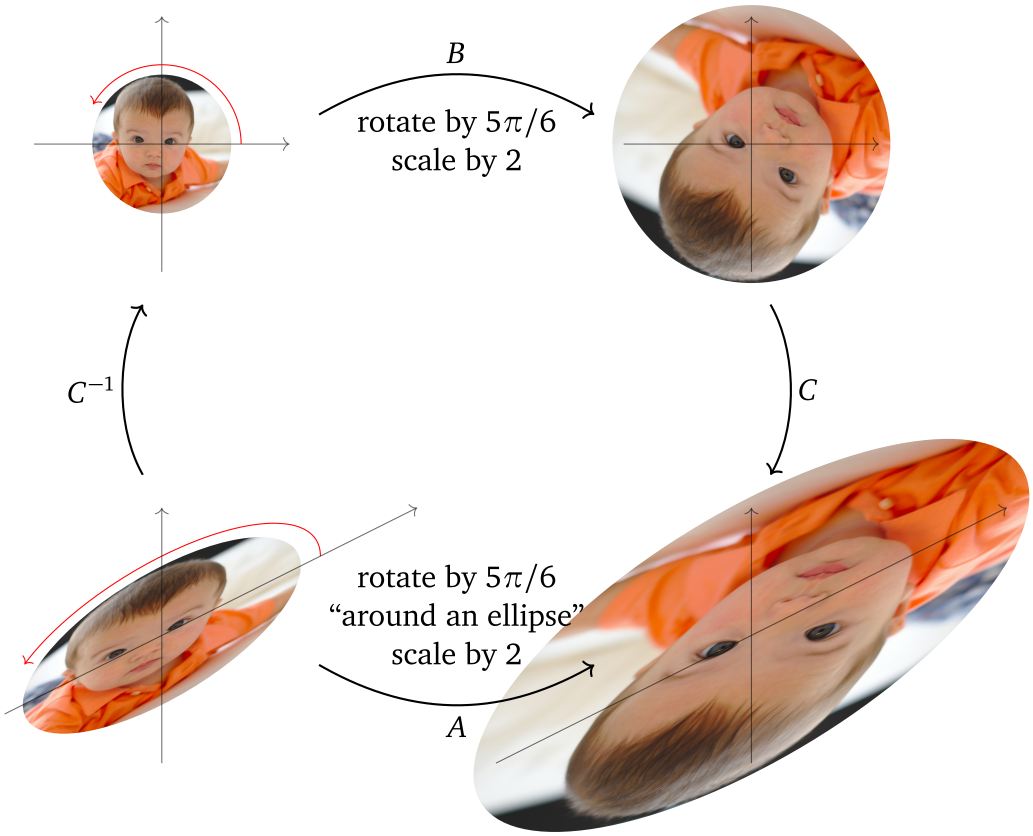
We choose the eigenvalue $\lambda = -\sqrt{3} - i$ and find a corresponding eigenvector, using the trick:

$$A - (-\sqrt{3} - i)I_2 = \begin{pmatrix} 1+i & -2 \\ \star & \star \end{pmatrix} \xrightarrow{\text{eigenvector}} v = \begin{pmatrix} 2 \\ 1+i \end{pmatrix}.$$

According to the rotation-scaling theorem, we have $A = CBC^{-1}$ for

$$C = \left(\mathrm{Re}\begin{pmatrix} 2 \\ 1+i \end{pmatrix} \ \ \mathrm{Im}\begin{pmatrix} 2 \\ 1+i \end{pmatrix} \right) = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$
$$B = \begin{pmatrix} \mathrm{Re}(\lambda) & \mathrm{Im}(\lambda) \\ -\mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{pmatrix} = \begin{pmatrix} -\sqrt{3} & -1 \\ 1 & -\sqrt{3} \end{pmatrix}.$$

The matrix B is the rotation-scaling matrix in the above example: it rotates counterclockwise by an angle of $5\pi/6$ and scales by a factor of 2. The matrix A does the same thing, but with respect to the $\mathrm{Re}(v), \mathrm{Im}(v)$-coordinate system:



To summarize:

- B rotates around the circle centered at the origin and passing through e_1 and e_2, in the direction from e_1 to e_2, then scales by 2.
- A rotates around the ellipse centered at the origin and passing through $\mathrm{Re}(v)$ and $\mathrm{Im}(v)$, in the direction from $\mathrm{Re}(v)$ to $\mathrm{Im}(v)$, then scales by 2.

The reader might want to refer back to this example in Section 5.3.

Use this link to view the online demo

The geometric action of A and B on the plane. Click "multiply" to multiply the colored points by B on the left and A on the right.

If instead we had chosen $\overline{\lambda} = -\sqrt{3} - i$ as our eigenvalue, then we would have found the eigenvector $\overline{v} = \binom{2}{1-i}$. In this case we would have $A = C'B'(C')^{-1}$, where

$$C' = \begin{pmatrix} \mathrm{Re}\begin{pmatrix} 2 \\ 1-i \end{pmatrix} & \mathrm{Im}\begin{pmatrix} 2 \\ 1-i \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix}$$

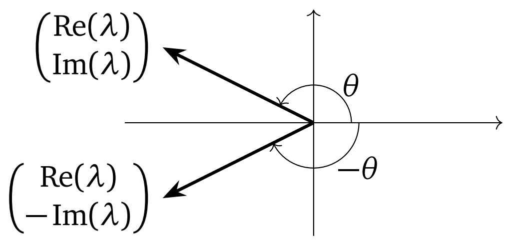
$$B' = \begin{pmatrix} \mathrm{Re}(\overline{\lambda}) & \mathrm{Im}(\overline{\lambda}) \\ -\mathrm{Im}(\overline{\lambda}) & \mathrm{Re}(\overline{\lambda}) \end{pmatrix} = \begin{pmatrix} -\sqrt{3} & 1 \\ -1 & -\sqrt{3} \end{pmatrix}.$$

So, A is also similar to a *clockwise* rotation by $5\pi/6$, followed by a scale by 2.

We saw in the above examples that the rotation-scaling theorem can be applied in two different ways to any given matrix: one has to choose one of the two conjugate eigenvalues to work with. Replacing λ by $\overline{\lambda}$ has the effect of replacing v by $\overline{v}$, which just negates all imaginary parts, so we also have $A = C'B'(C')^{-1}$ for

$$C' = \begin{pmatrix} | & | \\ \mathrm{Re}(v) & -\mathrm{Im}(v) \\ | & | \end{pmatrix} \quad \text{and} \quad B' = \begin{pmatrix} \mathrm{Re}(\lambda) & -\mathrm{Im}(\lambda) \\ \mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{pmatrix}.$$

The matrices B and B' are similar to each other. The only difference between them is the direction of rotation, since $\binom{\mathrm{Re}(\lambda)}{-\mathrm{Im}(\lambda)}$ and $\binom{\mathrm{Re}(\lambda)}{\mathrm{Im}(\lambda)}$ are mirror images of each other over the x-axis:



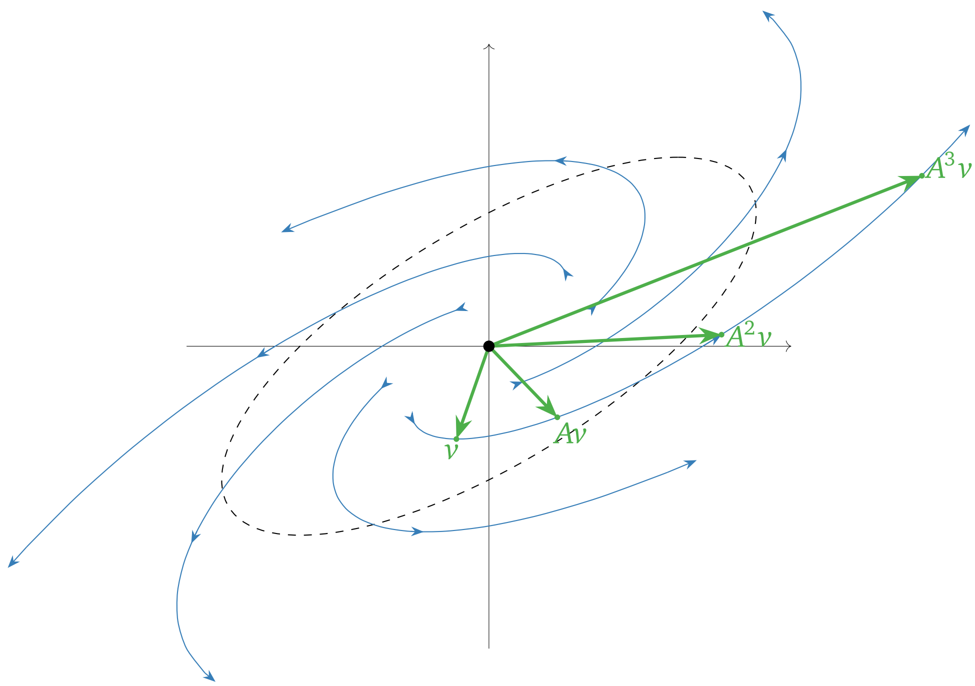
The discussion that follows is closely analogous to the exposition in this subsection in Section 5.4, in which we studied the dynamics of diagonalizable 2×2 matrices.

Dynamics of a 2×2 Matrix with a Complex Eigenvalue. Let A be a 2×2 matrix with a complex (non-real) eigenvalue λ. By the rotation-scaling theorem, the matrix A is similar to a matrix that rotates by some amount and scales by $|\lambda|$. Hence, A rotates around an ellipse and scales by $|\lambda|$. There are three different cases.

$|\lambda| > 1$: when the scaling factor is greater than 1, then vectors tend to get longer, i.e., farther from the origin. In this case, repeatedly multiplying a vector by A makes the vector "spiral out". For example,

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{pmatrix} \qquad \lambda = \frac{\sqrt{3}-i}{\sqrt{2}} \qquad |\lambda| = \sqrt{2} > 1$$

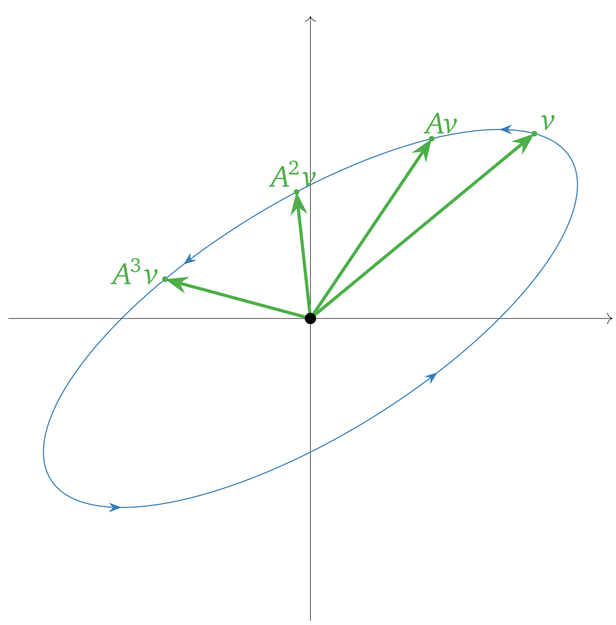
gives rise to the following picture:



$|\lambda| = 1$: when the scaling factor is equal to 1, then vectors do not tend to get longer or shorter. In this case, repeatedly multiplying a vector by A simply "rotates around an ellipse". For example,

$$A = \frac{1}{2}\begin{pmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{pmatrix} \qquad \lambda = \frac{\sqrt{3}-i}{2} \qquad |\lambda| = 1$$

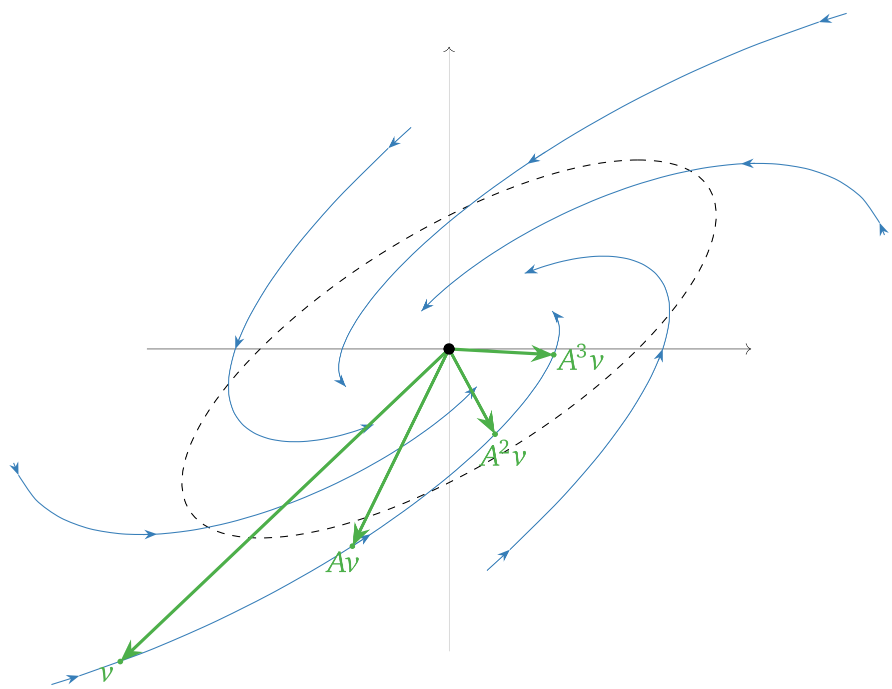
gives rise to the following picture:



$|\lambda| < 1$: when the scaling factor is less than 1, then vectors tend to get shorter, i.e., closer to the origin. In this case, repeatedly multiplying a vector by A makes the vector "spiral in". For example,

$$A = \frac{1}{2\sqrt{2}}\begin{pmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{pmatrix} \qquad \lambda = \frac{\sqrt{3}-i}{2\sqrt{2}} \qquad |\lambda| = \frac{1}{\sqrt{2}} < 1$$

gives rise to the following picture:



Interactive: $|\lambda| > 1$.

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{pmatrix} \qquad B = \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \qquad C = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$

$$\lambda = \frac{\sqrt{3}-i}{\sqrt{2}} \qquad |\lambda| = \sqrt{2} > 1$$

Use this link to view the online demo

The geometric action of A and B on the plane. Click "multiply" to multiply the colored points by B on the left and A on the right.

Interactive: $|\lambda| = 1$.

$$A = \frac{1}{2}\begin{pmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{pmatrix} \qquad B = \frac{1}{2}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \qquad C = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$

$$\lambda = \frac{\sqrt{3}-i}{2} \qquad |\lambda| = 1$$

Use this link to view the online demo

The geometric action of A and B on the plane. Click "multiply" to multiply the colored points by B on the left and A on the right.

Interactive: $|\lambda| < 1$.

$$A = \frac{1}{2\sqrt{2}}\begin{pmatrix} \sqrt{3}+1 & -2 \\ 1 & \sqrt{3}-1 \end{pmatrix} \qquad B = \frac{1}{2\sqrt{2}}\begin{pmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{pmatrix} \qquad C = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$$

$$\lambda = \frac{\sqrt{3}-i}{2\sqrt{2}} \qquad |\lambda| = \frac{1}{\sqrt{2}} < 1$$

Use this link to view the online demo

The geometric action of A and B on the plane. Click "multiply" to multiply the colored points by B on the left and A on the right.

Remark (Classification of 2×2 matrices up to similarity)**.** At this point, we can write down the "simplest" possible matrix which is similar to any given 2×2 matrix A. There are four cases:

1. A has two real eigenvalues λ_1, λ_2. In this case, A is diagonalizable, so A is similar to the matrix
$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$
This representation is unique up to reordering the eigenvalues.

2. A has one real eigenvalue λ of geometric multiplicity 2. In this case, we saw in this example in Section 5.4 that A is *equal* to the matrix
$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}.$$

3. A has one real eigenvalue λ of geometric multiplicity 1. In this case, A is not diagonalizable, and we saw in this remark in Section 5.4 that A is similar to the matrix
$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

4. A has no real eigenvalues. In this case, A has a complex eigenvalue λ, and A is similar to the rotation-scaling matrix

$$\begin{pmatrix} \mathrm{Re}(\lambda) & \mathrm{Im}(\lambda) \\ -\mathrm{Im}(\lambda) & \mathrm{Re}(\lambda) \end{pmatrix}$$

by the rotation-scaling theorem. By this proposition, the eigenvalues of a rotation-scaling matrix $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ are $a \pm bi$, so that two rotation-scaling matrices $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} c & -d \\ d & c \end{smallmatrix}\right)$ are similar if and only if $a = c$ and $b = \pm d$.

5.5.4 Block Diagonalization

For matrices larger than 2×2, there is a theorem that combines the diagonalization theorem in Section 5.4 and the rotation-scaling theorem. It says essentially that a matrix is similar to a matrix with parts that look like a diagonal matrix, and parts that look like a rotation-scaling matrix.

Block Diagonalization Theorem. *Let A be a real $n \times n$ matrix. Suppose that for each (real or complex) eigenvalue, the algebraic multiplicity equals the geometric multiplicity. Then $A = CBC^{-1}$, where B and C are as follows:*

- *The matrix B is **block diagonal**, where the blocks are 1×1 blocks containing the real eigenvalues (with their multiplicities), or 2×2 blocks containing the matrices*

$$\begin{pmatrix} \mathrm{Re}(\lambda) & \mathrm{Im}(\lambda) \\ -\mathrm{Im}(\lambda) & \mathrm{Re}(\Lambda) \end{pmatrix}$$

  for each non-real eigenvalue λ (with multiplicity).

- *The columns of C form bases for the eigenspaces for the real eigenvectors, or come in pairs $\big(\mathrm{Re}(v)\ \mathrm{Im}(v)\big)$ for the non-real eigenvectors.*

The block diagonalization theorem is proved in the same way as the diagonalization theorem in Section 5.4 and the rotation-scaling theorem. It is best understood in the case of 3×3 matrices.

Block Diagonalization of a 3×3 Matrix with a Complex Eigenvalue. Let A be a 3×3 matrix with a complex eigenvalue λ_1. Then $\overline{\lambda}_1$ is another eigenvalue, and there is one real eigenvalue λ_2. Since there are three distinct eigenvalues, they have algebraic and geometric multiplicity one, so the block diagonalization theorem applies to A.

Let v_1 be a (complex) eigenvector with eigenvalue λ_1, and let v_2 be a (real) eigenvector with eigenvalue λ_2. Then the block diagonalization theorem says that $A = CBC^{-1}$ for

$$C = \begin{pmatrix} | & | & | \\ \mathrm{Re}(v_1) & \mathrm{Im}(v_1) & v_2 \\ | & | & | \end{pmatrix} \qquad B = \begin{pmatrix} \mathrm{Re}(\lambda_1) & \mathrm{Im}(\lambda_1) & 0 \\ -\mathrm{Im}(\lambda_1) & \mathrm{Re}(\lambda_1) & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

Example (Geometry of a 3×3 matrix with a complex eigenvalue)**.** What does the matrix

$$A = \frac{1}{29}\begin{pmatrix} 33 & -23 & 9 \\ 22 & 33 & -23 \\ 19 & 14 & 50 \end{pmatrix}$$

do geometrically?

Solution. First we find the (real and complex) eigenvalues of A. We compute the characteristic polynomial using whatever method we like:

$$f(\lambda) = \det(A - \lambda I_3) = -\lambda^3 + 4\lambda^2 - 6\lambda + 4.$$

We search for a real root using the rational root theorem. The possible rational roots are $\pm 1, \pm 2, \pm 4$; we find $f(2) = 0$, so that $\lambda - 2$ divides $f(\lambda)$. Performing polynomial long division gives

$$f(\lambda) = -(\lambda - 2)\big(\lambda^2 - 2\lambda + 2\big).$$

The quadratic term has roots

$$\lambda = \frac{2 \pm \sqrt{4-8}}{2} = 1 \pm i,$$

so that the complete list of eigenvalues is $\lambda_1 = 1 - i$, $\overline{\lambda_1} = 1 + i$, and $\lambda_2 = 2$.

Now we compute some eigenvectors, starting with $\lambda_1 = 1 - i$. We row reduce (probably with the aid of a computer):

$$A - (1-i)I_3 = \frac{1}{29}\begin{pmatrix} 4+29i & -23 & 9 \\ 22 & 4+29i & -23 \\ 19 & 14 & 21+29i \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 7/5 + i/5 \\ 0 & 1 & -2/5 + 9i/5 \\ 0 & 0 & 0 \end{pmatrix}.$$

The free variable is z, and the parametric form is

$$\begin{cases} x = -\left(\frac{7}{5} + \frac{1}{5}i\right)z \\ y = \left(\frac{2}{5} - \frac{9}{5}i\right)z \end{cases} \xrightarrow[\text{eigenvector}]{z=5} v_1 = \begin{pmatrix} -7-i \\ 2-9i \\ 5 \end{pmatrix}.$$

For $\lambda_2 = 2$, we have

$$A - 2I_3 = \frac{1}{29}\begin{pmatrix} -25 & -23 & 9 \\ 22 & -25 & -23 \\ 19 & 14 & -8 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -2/3 \\ 0 & 1 & 1/3 \\ 0 & 0 & 0 \end{pmatrix}.$$

The free variable is z, and the parametric form is

$$\begin{cases} x = \frac{2}{3}z \\ y = -\frac{1}{3}z \end{cases} \xrightarrow[\text{eigenvector}]{z=3} v_2 = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}.$$

According to the block diagonalization theorem, we have $A = CBC^{-1}$ for

$$C = \begin{pmatrix} | & | & | \\ \mathrm{Re}(v_1) & \mathrm{Im}(v_1) & v_2 \\ | & | & | \end{pmatrix} = \begin{pmatrix} -7 & -1 & 2 \\ 2 & -9 & -1 \\ 5 & 0 & 3 \end{pmatrix}$$
$$B = \begin{pmatrix} \mathrm{Re}(\lambda_1) & \mathrm{Im}(\lambda_1) & 0 \\ -\mathrm{Im}(\lambda_1) & \mathrm{Re}(\lambda_1) & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

The matrix B is a combination of the rotation-scaling matrix $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ from this example, and a diagonal matrix. More specifically, B acts on the xy-coordinates by rotating counterclockwise by 45° and scaling by $\sqrt{2}$, and it scales the z-coordinate by 2. This means that points above the xy-plane spiral out away from the z-axis and move up, and points below the xy-plane spiral out away from the z-axis and move down.

The matrix A does the same thing as B, but with respect to the $\{\mathrm{Re}(v_1), \mathrm{Im}(v_1), v_2\}$-coordinate system. That is, A acts on the $\mathrm{Re}(v_1), \mathrm{Im}(v_1)$-plane by spiraling out, and A acts on the v_2-coordinate by scaling by a factor of 2. See the demo below.

Use this link to view the online demo

The geometric action of A and B on $\mathbf{R}^3$. *Click “multiply” to multiply the colored points by B on the left and A on the right. (We have scaled C by* $1/6$ *so that the vectors x and y have roughly the same size.)*

5.6 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.

5.6.1 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. 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 $\mathbf{R}^n$.

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

$$v_t = Av_{t-1} = A^2 v_{t-2} = \cdots = A^t v_0,$$

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. 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 (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:

$$\begin{aligned} x_{t+1} &= \qquad 6y_t + 8z_t \\ y_{t+1} &= \tfrac{1}{2}x_t \\ z_{t+1} &= \qquad \tfrac{1}{2}y_t. \end{aligned}$$

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:

Use this link to view the online demo

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.*

5.6.2 Stochastic Matrices and the Steady State

In this subsection, we discuss difference equations representing *probabilities*, like the Red Box example. 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. 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.

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. Continuing with the Red Box example, 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% will be returned to kiosk 1
> 30% will be returned to kiosk 2
> 40% 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.

> 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. *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. 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.

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 (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

$$f(\lambda) = \lambda^2 - \mathrm{Tr}(A)\lambda + \det(\lambda) = \lambda^2 - \frac{3}{2}\lambda + \frac{1}{2} = (\lambda - 1)(\lambda - 1/2).$$

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 $v_1 = Av_0$, $v_2 = Av_1$, ... 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 $\mathbf{R}^2$; for any vector $x = a_1u_1 + a_2u_2$ in $\mathbf{R}^2$, 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

$$A^t x = a_1u_1 + \frac{a_2}{2^t}u_2 \longrightarrow a_1u_1$$

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.

Use this link to view the online demo

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. Continuing with the Red Box example, 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 $\mathbf{R}^3$; for any vector $x = a_1u_1 + a_2u_2 + a_3u_3$ in $\mathbf{R}^3$, 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

$$A^t x = a_1u_1 - (0.2)^t a_2u_2 + (0.1)^t a_3u_3 \longrightarrow a_1u_1$$

as $t \to \infty$. This shows that $A^t x$ 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.

Use this link to view the online demo

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.

5.6.3 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.

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:

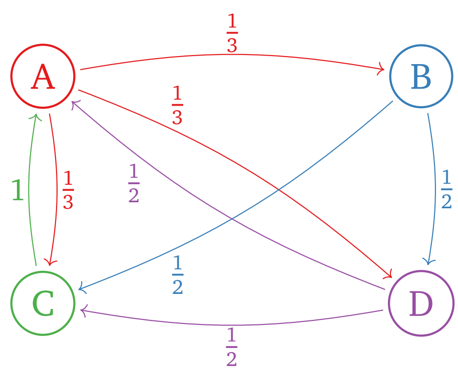
- If a very important page links to your page (and not to a zillion other ones as well), then your page is considered important.
- If a zillion unimportant pages link to your page, then your page is still important.
- If only one unknown page links to yours, your page is not important.

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.

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. Consider the following internet with only four pages. Links are indicated by arrows.



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

$$\begin{pmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} c + \frac{1}{2}d \\ \frac{1}{3}a \\ \frac{1}{3}a + \frac{1}{2}b + \frac{1}{2}d \\ \frac{1}{3}a + \frac{1}{2}b \end{pmatrix} = \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

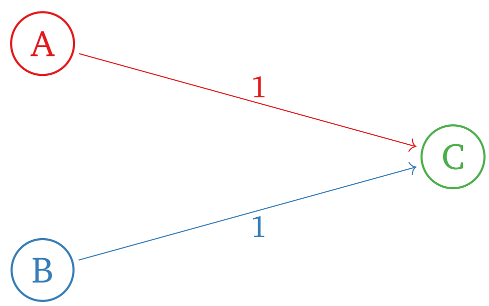
The matrix on the left is the **importance matrix**, and the final equality expresses the importance rule.

The above example illustrates the key observation.

Key Observation. 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 (A page with no links)**.** Consider the following internet with three pages:



The importance matrix is

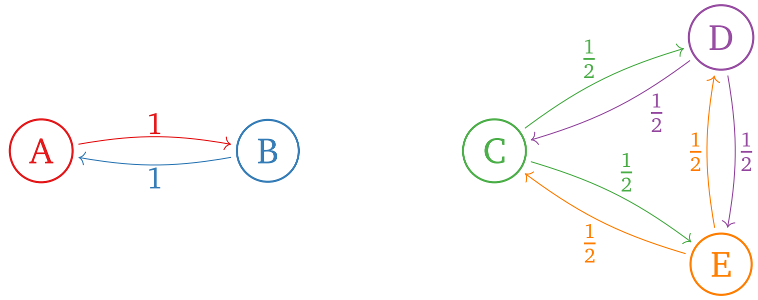
$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

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 (Disconnected Internet)**.** Consider the following internet:



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 $p = 0.15$.)

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

$$M = (1-p)\cdot A' + p\cdot B \quad \text{where} \quad B = \frac{1}{n}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

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.

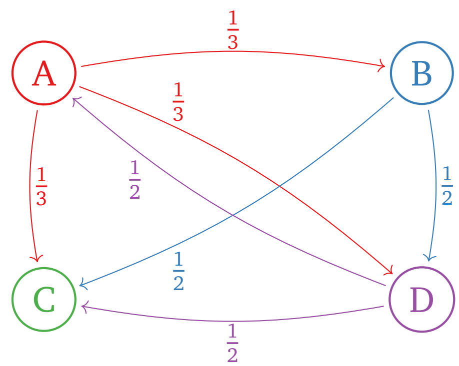
Fact. *The Google Matrix is a positive stochastic matrix.*

If we declare that the ranks of all of the pages must sum to 1, then we find:

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. What is the PageRank vector for the following internet? (Use the damping factor $p = 0.15$.)



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

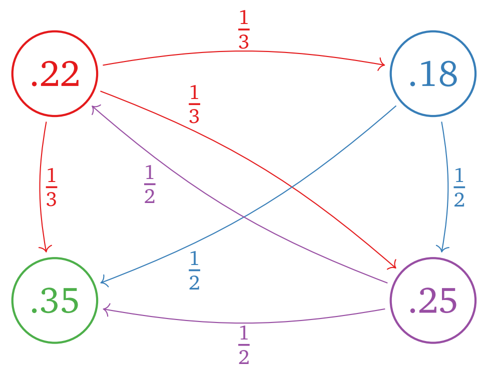
$$M = 0.85 \cdot \begin{pmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{1}{4} & 0 \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & \frac{1}{4} & 0 \end{pmatrix} + 0.15 \cdot \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

$$\approx \begin{pmatrix} 0.0375 & 0.0375 & 0.2500 & 0.4625 \\ 0.3208 & 0.0375 & 0.2500 & 0.0375 \\ 0.3208 & 0.4625 & 0.2500 & 0.4625 \\ 0.3208 & 0.4625 & 0.2500 & 0.0375 \end{pmatrix}.$$

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:



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.

Chapter 6

Orthogonality

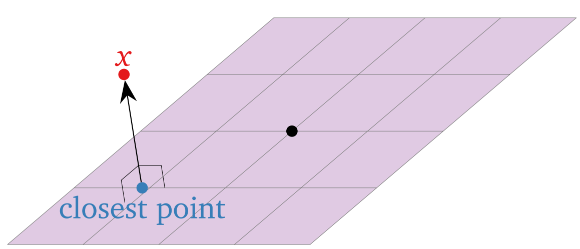
Let us recall one last time the structure of this book:

1. Solve the matrix equation $Ax = b$.
2. Solve the matrix equation $Ax = \lambda x$, where λ is a number.
3. Approximately solve the matrix equation $Ax = b$.

We have now come to the third part.

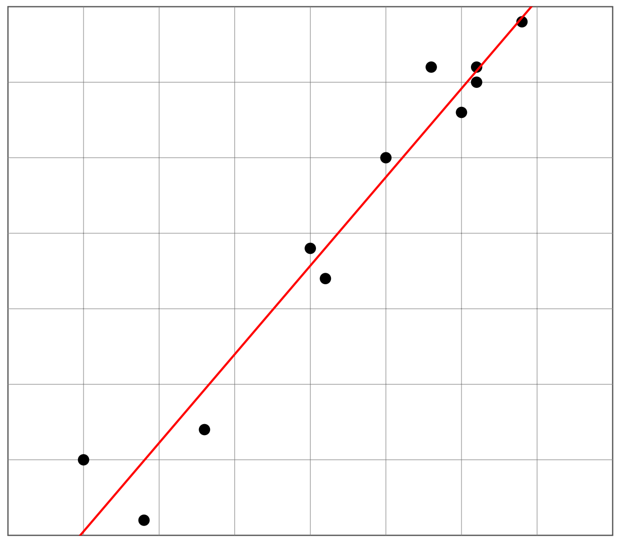
Primary Goal. Approximately solve the matrix equation $Ax = b$.

Finding approximate solutions of equations generally requires computing the closest vector on a subspace to a given vector. This becomes an *orthogonality* problem: one needs to know which vectors are perpendicular to the subspace.

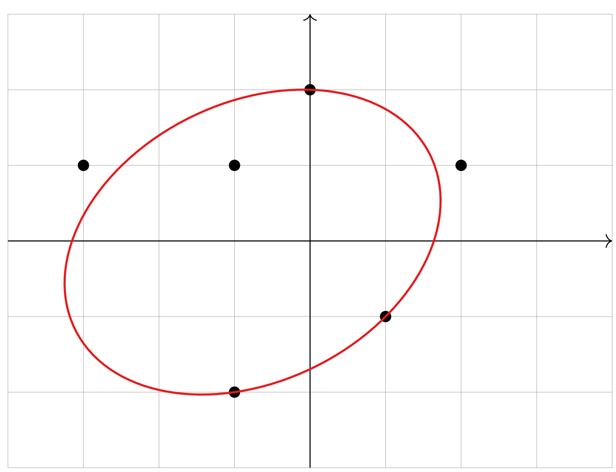


First we will define orthogonality and learn to find orthogonal complements of subspaces in Section 6.1 and Section 6.2. The core of this chapter is Section 6.3, in which we discuss *the orthogonal projection* of a vector onto a subspace; this is a method of calculating the closest vector on a subspace to a given vector. These calculations become easier in the presence of an orthogonal set, as we will see in Section 6.4. In Section 6.5 we will present the least-squares method of approximately solving systems of equations, and we will give applications to data modeling.

Example. In data modeling, one often asks: "what line is my data supposed to lie on?" This can be solved using a simple application of the least-squares method.



Example. Gauss invented the method of least squares to find a best-fit ellipse: he correctly predicted the (elliptical) orbit of the asteroid Ceres as it passed behind the sun in 1801.

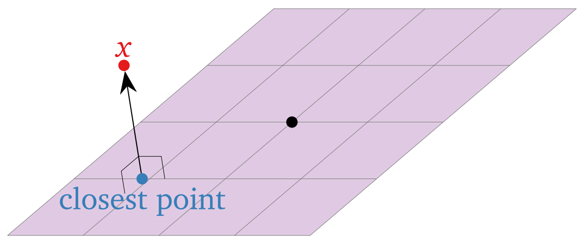


6.1 Dot Products and Orthogonality

Objectives

1. Understand the relationship between the dot product, length, and distance.
2. Understand the relationship between the dot product and orthogonality.
3. *Vocabulary words:* **dot product**, **length**, **distance**, **unit vector**, **unit vector in the direction of** x.
4. *Essential vocabulary word:* **orthogonal**.

In this chapter, it will be necessary to find the *closest* point on a subspace to a given point, like so:



The closest point has the property that the difference between the two points is *orthogonal*, or *perpendicular*, to the subspace. For this reason, we need to develop notions of orthogonality, length, and distance.

6.1.1 The Dot Product

The basic construction in this section is the *dot product*, which measures angles between vectors and computes the length of a vector.

Definition. The **dot product** of two vectors x, y in $\mathbf{R}^n$ is

$$x \cdot y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

Thinking of x, y as column vectors, this is the same as $x^T y$.

For example,

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 = 32.$$

Notice that the dot product of two *vectors* is a *scalar*.

You can do arithmetic with dot products mostly as usual, as long as you remember you can only dot two vectors together, and that the result is a scalar.

Properties of the Dot Product. Let x, y, z be vectors in $\mathbf{R}^n$ and let c be a scalar.

1. *Commutativity:* $x \cdot y = y \cdot x$.
2. *Distributivity with addition:* $(x + y) \cdot z = x \cdot z + y \cdot z$.
3. *Distributivity with scalar multiplication:* $(cx) \cdot y = c(x \cdot y)$.

The dot product of a vector with itself is an important special case:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1^2 + x_2^2 + \cdots + x_n^2.$$

Therefore, for any vector x, we have:

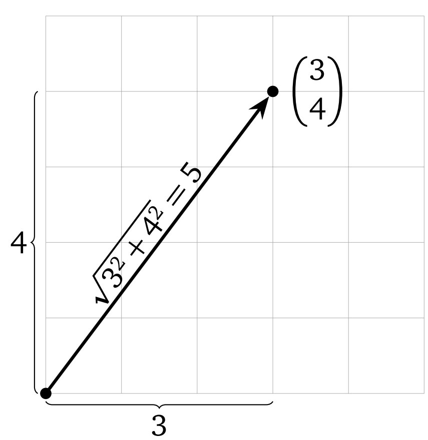
- $x \cdot x \geq 0$
- $x \cdot x = 0 \iff x = 0$.

This leads to a good definition of *length*.

Fact. *The* ***length*** *of a vector x in $\mathbf{R}^n$ is the number*

$$\|x\| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

It is easy to see why this is true for vectors in $\mathbf{R}^2$, by the Pythagorean theorem.



$$\left\| \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\| = \sqrt{3^2 + 4^2} = 5$$

For vectors in $\mathbf{R}^3$, one can check that $\|x\|$ really is the length of x, although now this requires *two* applications of the Pythagorean theorem.

Note that the length of a vector is the length of the *arrow*; if we think in terms of points, then the length is its distance from the origin.

Example. Suppose that $\|x\| = 2$, $\|y\| = 3$, and $x \cdot y = -4$. What is $\|2x + 3y\|$?

Solution. We compute

$$\begin{aligned}\|2x+3y\|^2 &= (2x+3y)\cdot(2x+3y)\\ &= 4x\cdot x + 6x\cdot y + 6x\cdot y + 9y\cdot y\\ &= 4\|x\|^2 + 9\|y\|^2 + 12x\cdot y\\ &= 4\cdot 4 + 9\cdot 9 - 12\cdot 4 = 49.\end{aligned}$$

Hence $\|2x + 3y\| = \sqrt{49} = 7$.

Fact. *If x is a vector and c is a scalar, then $\|cx\| = |c| \cdot \|x\|$.*

This says that scaling a vector by c scales its length by $|c|$. For example,

$$\left\|\begin{pmatrix}6\\8\end{pmatrix}\right\| = \left\|2\begin{pmatrix}3\\4\end{pmatrix}\right\| = 2\left\|\begin{pmatrix}3\\4\end{pmatrix}\right\| = 10.$$

Now that we have a good notion of length, we can define the *distance* between points in $\mathbf{R}^n$. Recall that the difference between two points x, y is naturally a vector, namely, the vector $y - x$ pointing from x to y.

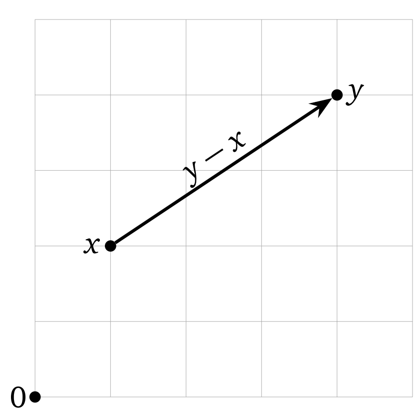
Definition. The **distance** between two points x, y in $\mathbf{R}^n$ is the length of the vector from x to y:

$$\text{dist}(x, y) = \|y - x\|.$$

Example. Find the distance from $(1, 2)$ to $(4, 4)$.

Solution. Let $x = (1, 2)$ and $y = (4, 4)$. Then

$$\text{dist}(x, y) = \|y - x\| = \left\|\begin{pmatrix}3\\2\end{pmatrix}\right\| = \sqrt{3^2 + 2^2} = \sqrt{13}.$$



Vectors with length 1 are very common in applications, so we give them a name.

Definition. A **unit vector** is a vector x with length $\|x\| = \sqrt{x \cdot x} = 1$.

The standard coordinate vectors $e_1, e_2, e_3, \ldots$ are unit vectors:

$$\|e_1\| = \left\| \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\| = \sqrt{1^2 + 0^2 + 0^2} = 1.$$

For any nonzero vector x, there is a unique unit vector pointing in the same direction. It is obtained by dividing by the length of x.

Fact. *Let x be a nonzero vector in $\mathbf{R}^n$. The* ***unit vector in the direction of*** *x is the vector $x/\|x\|$.*

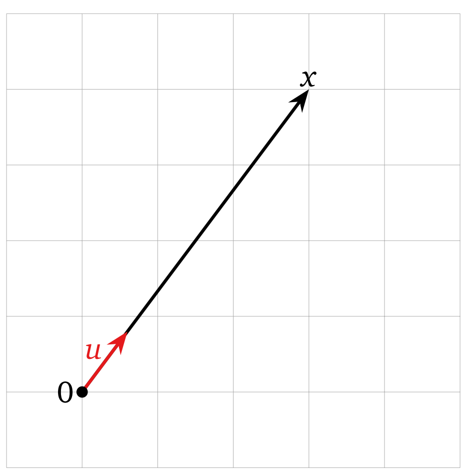
This is in fact a unit vector (noting that $\|x\|$ is a positive number, so $\big|1/\|x\|\big| = 1/\|x\|$):

$$\left\| \frac{x}{\|x\|} \right\| = \frac{1}{\|x\|}\|x\| = 1.$$

Example. What is the unit vector u in the direction of $x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$?

Solution. We divide by the length of x:

$$u = \frac{x}{\|x\|} = \frac{1}{\sqrt{3^2 + 4^2}} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 3/5 \\ 4/5 \end{pmatrix}.$$



6.1.2 Orthogonal Vectors

In this section, we show how the dot product can be used to define *orthogonality*, i.e., when two vectors are perpendicular to each other.

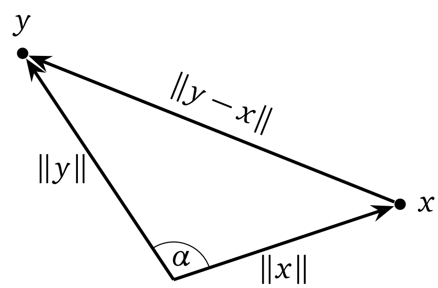
Definition. Two vectors x, y in $\mathbf{R}^n$ are **orthogonal** or **perpendicular** if $x \cdot y = 0$.

Notation: $x \perp y$ means $x \cdot y = 0$.

Since $0 \cdot x = 0$ for any vector x, the zero vector is orthogonal to every vector in $\mathbf{R}^n$.

We motivate the above definition using the *law of cosines* in $\mathbf{R}^2$. In our language, the law of cosines asserts that if x, y are two nonzero vectors, and if $\alpha > 0$ is the angle between them, then

$$\|y - x\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\,\|y\|\cos\alpha.$$



In particular, $\alpha = 90^\circ$ if and only if $\cos(\alpha) = 0$, which happens if and only if $\|y - x\|^2 = \|x\|^2 + \|y\|^2$. Therefore,

$$\begin{aligned} x \text{ and } y \text{ are perpendicular} &\iff \|x\|^2 + \|y\|^2 = \|y - x\|^2 \\ &\iff x \cdot x + y \cdot y = (y - x) \cdot (y - x) \\ &\iff x \cdot x + y \cdot y = y \cdot y + x \cdot x - 2x \cdot y \\ &\iff x \cdot y = 0. \end{aligned}$$

To reiterate:

$$x \perp y \quad \iff \quad x \cdot y = 0 \quad \iff \quad \|y - x\|^2 = \|x\|^2 + \|y\|^2.$$

Example. Find all vectors orthogonal to $v = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

Solution. We have to find all vectors x such that $x \cdot v = 0$. This means solving the equation

$$0 = x \cdot v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = x_1 + x_2 - x_3.$$

The parametric form for the solution set is $x_1 = -x_2 + x_3$, so the parametric vector form of the general solution is

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Therefore, the answer is the *plane*

$$\text{Span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

For instance,

$$\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \perp \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \qquad \text{because} \qquad \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = 0.$$

Example. Find all vectors orthogonal to both $v = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ and $w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Solution. We have to solve the system of two homogeneous equations

$$\begin{aligned} 0 = \ x \cdot v = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} &= x_1 + x_2 - x_3 \\ 0 = x \cdot w = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} &= x_1 + x_2 + x_3. \end{aligned}$$

In matrix form:

$$\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The parametric vector form of the solution set is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Therefore, the answer is the *line*

$$\mathrm{Span}\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}\right\}.$$

For instance,

$$\begin{pmatrix}-1\\1\\0\end{pmatrix}\cdot\begin{pmatrix}1\\1\\-1\end{pmatrix}=0 \quad\text{and}\quad \begin{pmatrix}-1\\1\\0\end{pmatrix}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}=0.$$

Remark (Angle between two vectors)**.** More generally, the law of cosines gives a formula for the angle α between two nonzero vectors:

$$\begin{aligned}2\|x\|\|y\|\cos(\alpha) &= \|x\|^2+\|y\|^2-\|y-x\|^2\\ &= x\cdot x+y\cdot y-(y-x)\cdot(y-x)\\ &= x\cdot x+y\cdot y-y\cdot y-x\cdot x+2x\cdot y\\ &= 2x\cdot y\\ &\implies \alpha=\cos^{-1}\left(\frac{x\cdot y}{\|x\|\|y\|}\right).\end{aligned}$$

In higher dimensions, we take this to be the *definition* of the angle between x and y.

6.2 Orthogonal Complements

Objectives

1. Understand the basic properties of orthogonal complements.
2. Learn to compute the orthogonal complement of a subspace.
3. *Recipes:* shortcuts for computing the orthogonal complements of common subspaces.
4. *Picture:* orthogonal complements in $\mathbf{R}^2$ and $\mathbf{R}^3$.
5. *Theorem:* row rank equals column rank.
6. *Vocabulary words:* **orthogonal complement**, **row space**.

It will be important to compute the set of *all* vectors that are orthogonal to a given set of vectors. It turns out that a vector is orthogonal to a set of vectors if and only if it is orthogonal to the span of those vectors, which is a subspace, so we restrict ourselves to the case of subspaces.

6.2.1 Definition of the Orthogonal Complement

Taking the orthogonal complement is an operation that is performed on *subspaces*.

Definition. Let W be a subspace of $\mathbf{R}^n$. Its **orthogonal complement** is the subspace

$$W^\perp = \{v \text{ in } \mathbf{R}^n \mid v \cdot w = 0 \text{ for all } w \text{ in } W\}.$$

The symbol $W^\perp$ is sometimes read "W perp."

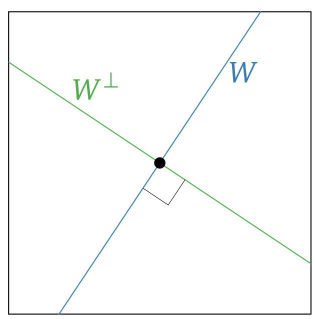
This is the set of all vectors v in $\mathbf{R}^n$ that are orthogonal to all of the vectors in W. We will show below that $W^\perp$ is indeed a subspace.

Note. We now have two similar-looking pieces of notation:

$$A^T \text{ is the transpose of a matrix } A.$$
$$W^\perp \text{ is the orthogonal complement of a subspace } W.$$

Try not to confuse the two.

Pictures of orthogonal complements The orthogonal complement of a line W through the origin in $\mathbf{R}^2$ is the perpendicular line $W^\perp$.

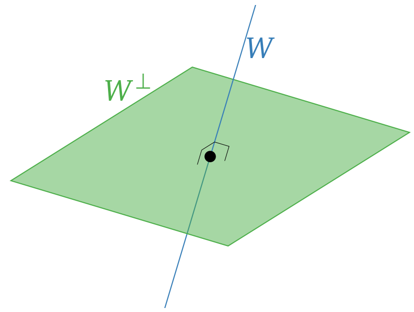


Interactive: Orthogonal complements in $\mathbf{R}^2$.

Use this link to view the online demo

The orthogonal complement of the line spanned by v is the perpendicular line. Click and drag the head of v to move it.

The orthogonal complement of a line W in $\mathbf{R}^3$ is the perpendicular plane $W^\perp$.

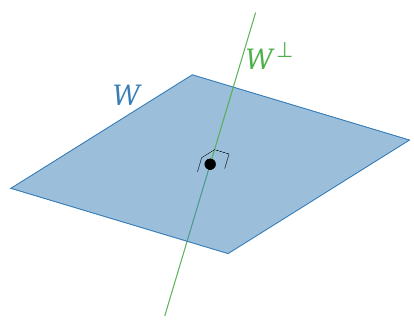


Interactive: Orthogonal complements in $\mathbf{R}^3$.

Use this link to view the online demo

The orthogonal complement of the line spanned by v is the perpendicular plane. Click and drag the head of v to move it.

The orthogonal complement of a plane W in $\mathbf{R}^3$ is the perpendicular line $W^\perp$.



Interactive: Orthogonal complements in $\mathbf{R}^3$.

Use this link to view the online demo

The orthogonal complement of the plane spanned by v, w is the perpendicular line. Click and drag the heads of v, w to change the plane.

We see in the above pictures that $(W^\perp)^\perp = W$.

Example. The orthogonal complement of $\mathbf{R}^n$ is $\{0\}$, since the zero vector is the only vector that is orthogonal to all of the vectors in $\mathbf{R}^n$.

For the same reason, we have $\{0\}^\perp = \mathbf{R}^n$.

6.2.2 Computing Orthogonal Complements

Since any subspace is a span, the following proposition gives a recipe for computing the orthogonal complement of any subspace. However, below we will give several shortcuts for computing the orthogonal complements of other common kinds of subspaces–in particular, null spaces. To compute the orthogonal complement of a general subspace, usually it is best to rewrite the subspace as the column space or null space of a matrix, as in this important note in Section 2.6.

Proposition (The orthogonal complement of a column space)**.** *Let A be a matrix and let $W = \mathrm{Col}(A)$. Then*

$$W^{\perp} = \mathrm{Nul}(A^T).$$

Proof. To justify the first equality, we need to show that a vector x is perpendicular to the all of the vectors in W if and only if it is perpendicular only to $v_1, v_2, \ldots, v_m$. Since the v_i are contained in W, we really only have to show that if $x \cdot v_1 = x \cdot v_2 = \cdots = x \cdot v_m = 0$, then x is perpendicular to every vector v in W. Indeed, any vector in W has the form $v = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m$ for suitable scalars $c_1, c_2, \ldots, c_m$, so

$$\begin{aligned} x \cdot v &= x \cdot (c_1 v_1 + c_2 v_2 + \cdots + c_m v_m) \\ &= c_1(x \cdot v_1) + c_2(x \cdot v_2) + \cdots + c_m(x \cdot v_m) \\ &= c_1(0) + c_2(0) + \cdots + c_m(0) = 0. \end{aligned}$$

Therefore, x is in $W^{\perp}$.

To prove the second equality, we let

$$A = \begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix}.$$

By the row-column rule for matrix multiplication in Section 2.3, for any vector x in $\mathbf{R}^n$ we have

$$Ax = \begin{pmatrix} v_1^T x \\ v_2^T x \\ \vdots \\ v_m^T x \end{pmatrix} = \begin{pmatrix} v_1 \cdot x \\ v_2 \cdot x \\ \vdots \\ v_m \cdot x \end{pmatrix}.$$

Therefore, x is in $\mathrm{Nul}(A)$ if and only if x is perpendicular to each vector $v_1, v_2, \ldots, v_m$. □

Since column spaces are the same as spans, we can rephrase the proposition as follows. Let $v_1, v_2, \ldots, v_m$ be vectors in $\mathbf{R}^n$, and let $W = \mathrm{Span}\{v_1, v_2, \ldots, v_m\}$. Then

$$W^{\perp} = \{\text{all vectors orthogonal to each } v_1, v_2, \ldots, v_m\} = \mathrm{Nul}\begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix}.$$

Again, it is important to be able to go easily back and forth between spans and column spaces. If you are handed a span, you can apply the proposition once you have rewritten your span as a column space.

By the proposition, computing the orthogonal complement of a span means *solving a system of linear equations*. For example, if

$$v_1 = \begin{pmatrix} 1 \\ 7 \\ 2 \end{pmatrix} \qquad v_2 = \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}$$

then $\mathrm{Span}\{v_1, v_2\}^\perp$ is the solution set of the homogeneous linear system associated to the matrix

$$\begin{pmatrix} —v_1^T— \\ —v_2^T— \end{pmatrix} = \begin{pmatrix} 1 & 7 & 2 \\ -2 & 3 & 1 \end{pmatrix}.$$

This is the solution set of the system of equations

$$\begin{cases} x_1 + 7x_2 + 2x_3 = 0 \\ -2x_1 + 3x_2 + x_3 = 0. \end{cases}$$

Example. Compute $W^\perp$, where

$$W = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ 7 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \right\}.$$

Solution. According to the proposition, we need to compute the null space of the matrix

$$\begin{pmatrix} 1 & 7 & 2 \\ -2 & 3 & 1 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -1/17 \\ 0 & 1 & 5/17 \end{pmatrix}.$$

The free variable is x_3, so the parametric form of the solution set is $x_1 = x_3/17$, $x_2 = -5x_3/17$, and the parametric vector form is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} 1/17 \\ -5/17 \\ 1 \end{pmatrix}.$$

Scaling by a factor of 17, we see that

$$W^\perp = \mathrm{Span}\left\{ \begin{pmatrix} 1 \\ -5 \\ 17 \end{pmatrix} \right\}.$$

We can check our work:

$$\begin{pmatrix} 1 \\ 7 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -5 \\ 17 \end{pmatrix} = 0 \qquad \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -5 \\ 17 \end{pmatrix} = 0.$$

Example. Find all vectors orthogonal to $v = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$.

Solution. According to the proposition, we need to compute the null space of the matrix

$$A = \begin{pmatrix} — v — \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}.$$

This matrix is in reduced-row echelon form. The parametric form for the solution set is $x_1 = -x_2 + x_3$, so the parametric vector form of the general solution is

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Therefore, the answer is the *plane*

$$\text{Span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Use this link to view the online demo

The set of all vectors perpendicular to v.

Example. Compute

$$\text{Span}\left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}^{\perp}.$$

Solution. According to the proposition, we need to compute the null space of the matrix

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The parametric vector form of the solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Therefore, the answer is the *line*

$$\text{Span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

Use this link to view the online demo

The orthogonal complement of the plane spanned by $v = (1,1,-1)$ *and* $w = (1,1,1)$.

In order to find shortcuts for computing orthogonal complements, we need the following basic facts. Looking back the the above examples, all of these facts should be believable.

Facts about Orthogonal Complements. *Let W be a subspace of $\mathbf{R}^n$. Then:*

1. $W^\perp$ *is also a subspace of* $\mathbf{R}^n$.
2. $(W^\perp)^\perp = W$.
3. $\dim(W) + \dim(W^\perp) = n$.

Proof. For the first assertion, we verify the three defining properties of subspaces.

1. The zero vector is in $W^\perp$ because the zero vector is orthogonal to every vector in $\mathbf{R}^n$.
2. Let u, v be in $W^\perp$, so $u \cdot x = 0$ and $v \cdot x = 0$ for every vector x in W. We must verify that $(u + v) \cdot x = 0$ for every x in W. Indeed, we have
$$(u + v) \cdot x = u \cdot x + v \cdot x = 0 + 0 = 0.$$
3. Let u be in $W^\perp$, so $u \cdot x = 0$ for every x in W, and let c be a scalar. We must verify that $(cu) \cdot x = 0$ for every x in W. Indeed, we have
$$(cu) \cdot x = c(u \cdot x) = c0 = 0.$$

Next we prove the third assertion. Let $v_1, v_2, \ldots, v_m$ be a basis for W, so $m = \dim(W)$, and let $v_{m+1}, v_{m+2}, \ldots, v_k$ be a basis for $W^\perp$, so $k - m = \dim(W^\perp)$. We need to show $k = n$. First we claim that $\{v_1, v_2, \ldots, v_m, v_{m+1}, v_{m+2}, \ldots, v_k\}$ is linearly independent. Suppose that $c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0$. Let $w = c_1v_1 + c_2v_2 + \cdots + c_mv_m$ and $w' = c_{m+1}v_{m+1} + c_{m+2}v_{m+2} + \cdots + c_kv_k$, so w is in W, w' is in W', and $w + w' = 0$. Then $w = -w'$ is in both W and $W^\perp$, which implies w is perpendicular to *itself*. In particular, $w \cdot w = 0$, so $w = 0$, and hence $w' = 0$. Therefore, all coefficients c_i are equal to zero, because $\{v_1, v_2, \ldots, v_m\}$ and $\{v_{m+1}, v_{m+2}, \ldots, v_k\}$ are linearly independent.

It follows from the previous paragraph that $k \leq n$. Suppose that $k < n$. Then the matrix
$$A = \begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_k^T — \end{pmatrix}$$

has more columns than rows (it is "wide"), so its null space is nonzero by this note in Section 3.2. Let x be a nonzero vector in $\mathrm{Nul}(A)$. Then

$$0 = Ax = \begin{pmatrix} v_1^T x \\ v_2^T x \\ \vdots \\ v_k^T x \end{pmatrix} = \begin{pmatrix} v_1 \cdot x \\ v_2 \cdot x \\ \vdots \\ v_k \cdot x \end{pmatrix}$$

by the row-column rule for matrix multiplication in Section 2.3. Since $v_1 \cdot x = v_2 \cdot x = \cdots = v_m \cdot x = 0$, it follows from this proposition that x is in $W^\perp$, and similarly, x is in $(W^\perp)^\perp$. As above, this implies x is orthogonal to itself, which contradicts our assumption that x is nonzero. Therefore, $k = n$, as desired.

Finally, we prove the second assertion. Clearly W is contained in $(W^\perp)^\perp$: this says that everything in W is perpendicular to the set of all vectors perpendicular to everything in W. Let $m = \dim(W)$. By 3, we have $\dim(W^\perp) = n - m$, so $\dim((W^\perp)^\perp) = n-(n-m) = m$. The only m-dimensional subspace of $(W^\perp)^\perp$ is all of $(W^\perp)^\perp$, so $(W^\perp)^\perp = W$. □

See these paragraphs for pictures of the second property. As for the third: for example, if W is a (2-dimensional) plane in $\mathbf{R}^4$, then $W^\perp$ is another (2-dimensional) plane. Explicitly, we have

$$\begin{aligned} \mathrm{Span}\{e_1, e_2\}^\perp &= \left\{ \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \text{ in } \mathbf{R}^4 \;\middle|\; \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0 \text{ and } \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} = 0 \right\} \\ &= \left\{ \begin{pmatrix} 0 \\ 0 \\ z \\ w \end{pmatrix} \text{ in } \mathbf{R}^4 \right\} = \mathrm{Span}\{e_3, e_4\}: \end{aligned}$$

the orthogonal complement of the xy-plane is the zw-plane.

Definition. The **row space** of a matrix A is the span of the rows of A, and is denoted $\mathrm{Row}(A)$.

If A is an $m \times n$ matrix, then the rows of A are vectors with n entries, so $\mathrm{Row}(A)$ is a subspace of $\mathbf{R}^n$. Equivalently, since the rows of A are the columns of A^T, the row space of A is the column space of A^T:

$$\mathrm{Row}(A) = \mathrm{Col}(A^T).$$

We showed in the above proposition that if A has rows $v_1^T, v_2^T, \ldots, v_m^T$, then

$$\mathrm{Row}(A)^\perp = \mathrm{Span}\{v_1, v_2, \ldots, v_m\}^\perp = \mathrm{Nul}(A).$$

Taking orthogonal complements of both sides and using the second fact gives

$$\mathrm{Row}(A) = \mathrm{Nul}(A)^\perp.$$

Replacing A by A^T and remembering that $\mathrm{Row}(A) = \mathrm{Col}(A^T)$ gives

$$\mathrm{Col}(A)^\perp = \mathrm{Nul}(A^T) \quad \text{and} \quad \mathrm{Col}(A) = \mathrm{Nul}(A^T)^\perp.$$

To summarize:

Recipes: Shortcuts for computing orthogonal complements. For any vectors $v_1, v_2, \ldots, v_m$, we have

$$\mathrm{Span}\{v_1, v_2, \ldots, v_m\}^\perp = \mathrm{Nul}\begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix}.$$

For any matrix A, we have

$$\mathrm{Row}(A)^\perp = \mathrm{Nul}(A) \qquad \mathrm{Nul}(A)^\perp = \mathrm{Row}(A)$$
$$\mathrm{Col}(A)^\perp = \mathrm{Nul}(A^T) \qquad \mathrm{Nul}(A^T)^\perp = \mathrm{Col}(A).$$

As mentioned in the beginning of this subsection, in order to compute the orthogonal complement of a general subspace, usually it is best to rewrite the subspace as the column space or null space of a matrix.

Example (Orthogonal complement of a subspace)**.** Compute the orthogonal complement of the subspace

$$W = \{(x, y, z) \text{ in } \mathbf{R}^3 \mid 3x + 2y = z\}.$$

Solution. Rewriting, we see that W is the solution set of the system of equations $3x + 2y - z = 0$, i.e., the null space of the matrix $A = \begin{pmatrix} 3 & 2 & -1 \end{pmatrix}$. Therefore,

$$W^\perp = \mathrm{Row}(A) = \mathrm{Span}\left\{\begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}\right\}.$$

No row reduction was needed!

Example (Orthogonal complement of an eigenspace)**.** Find the orthogonal complement of the 5-eigenspace of the matrix

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

Solution. The 5-eigenspace is

$$W = \mathrm{Nul}(A - 5I_3) = \mathrm{Nul}\begin{pmatrix} -3 & 4 & -1 \\ 3 & -3 & 0 \\ -2 & 4 & -2 \end{pmatrix},$$

so

$$W^\perp = \text{Row}\begin{pmatrix} -3 & 4 & -1 \\ 3 & -3 & 0 \\ -2 & 4 & -2 \end{pmatrix} = \text{Span}\left\{\begin{pmatrix} -3 \\ 4 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ -3 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 4 \\ -2 \end{pmatrix}\right\}.$$

These vectors are necessarily linearly dependent (why)?

6.2.3 Row rank and column rank

Suppose that A is an $m \times n$ matrix. Let us refer to the dimensions of $\text{Col}(A)$ and $\text{Row}(A)$ as the **row rank** and the **column rank** of A (note that the column rank of A is the same as the rank of A). The next theorem says that the row and column ranks are the same. This is surprising for a couple of reasons. First, $\text{Row}(A)$ lies in $\mathbf{R}^n$ and $\text{Col}(A)$ lies in $\mathbf{R}^m$. Also, the theorem implies that A and A^T have the same number of pivots, even though the reduced row echelon forms of A and A^T have nothing to do with each other otherwise.

Theorem. *Let A be a matrix. Then the row rank of A is equal to the column rank of A.*

Proof. By the rank theorem in Section 2.9, we have

$$\dim \text{Col}(A) + \dim \text{Nul}(A) = n.$$

On the other hand the third fact says that

$$\dim \text{Nul}(A)^\perp + \dim \text{Nul}(A) = n,$$

which implies $\dim \text{Col}(A) = \dim \text{Nul}(A)^\perp$. Since $\text{Nul}(A)^\perp = \text{Row}(A)$, we have

$$\dim \text{Col}(A) = \dim \text{Row}(A),$$

as desired. □

In particular, by this corollary in Section 2.7 both the row rank and the column rank are equal to the number of pivots of A.

6.3 Orthogonal Projection

Objectives

1. Understand the orthogonal decomposition of a vector with respect to a subspace.

2. Understand the relationship between orthogonal decomposition and orthogonal projection.

3. Understand the relationship between orthogonal decomposition and the closest vector on / distance to a subspace.
4. Learn the basic properties of orthogonal projections as linear transformations and as matrix transformations.
5. *Recipes:* orthogonal projection onto a line, orthogonal decomposition by solving a system of equations, orthogonal projection via a complicated matrix product.
6. *Pictures:* orthogonal decomposition, orthogonal projection.
7. *Vocabulary words:* **orthogonal decomposition**, **orthogonal projection**.

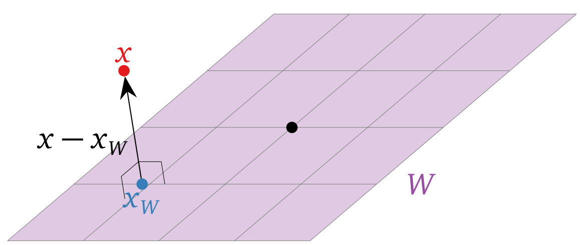
Let W be a subspace of $\mathbf{R}^n$ and let x be a vector in $\mathbf{R}^n$. In this section, we will learn to compute the *closest vector* x_W to x in W. The vector x_W is called the *orthogonal projection* of x onto W. This is exactly what we will use to almost solve matrix equations, as discussed in the introduction to Chapter 6.

6.3.1 Orthogonal Decomposition

We begin by fixing some notation.

Notation. Let W be a subspace of $\mathbf{R}^n$ and let x be a vector in $\mathbf{R}^n$. We denote the closest vector to x on W by x_W.

To say that x_W is the closest vector to x on W means that the difference $x - x_W$ is *orthogonal* to the vectors in W:



In other words, if $x_{W^\perp} = x - x_W$, then we have $x = x_W + x_{W^\perp}$, where x_W is in W and $x_{W^\perp}$ is in $W^\perp$. The first order of business is to prove that the closest vector always exists.

Theorem (Orthogonal decomposition)**.** *Let W be a subspace of $\mathbf{R}^n$ and let x be a vector in $\mathbf{R}^n$. Then we can write x uniquely as*

$$x = x_W + x_{W^\perp}$$

where x_W is the closest vector to x on W and $x_{W^\perp}$ is in $W^\perp$.

Proof. Let $m = \dim(W)$, so $n - m = \dim(W^\perp)$ by this fact in Section 6.2. Let $v_1, v_2, \ldots, v_m$ be a basis for W and let $v_{m+1}, v_{m+2}, \ldots, v_n$ be a basis for $W^\perp$. We showed in the proof of this fact in Section 6.2 that $\{v_1, v_2, \ldots, v_m, v_{m+1}, v_{m+2}, \ldots, v_n\}$ is linearly independent, so it forms a basis for $\mathbf{R}^n$. Therefore, we can write

$$x = (c_1v_1 + \cdots + c_mv_m) + (c_{m+1}v_{m+1} + \cdots + c_nv_n) = x_W + x_{W^\perp},$$

where $x_W = c_1v_1 + \cdots + c_mv_m$ and $x_{W^\perp} = c_{m+1}v_{m+1} + \cdots + c_nv_n$. Since $x_{W^\perp}$ is orthogonal to W, the vector x_W is the closest vector to x on W, so this proves that such a decomposition exists.

As for uniqueness, suppose that

$$x = x_W + x_{W^\perp} = y_W + y_{W^\perp}$$

for x_W, y_W in W and $x_{W^\perp}, y_{W^\perp}$ in $W^\perp$. Rearranging gives

$$x_W - y_W = y_{W^\perp} - x_{W^\perp}.$$

Since W and $W^\perp$ are subspaces, the left side of the equation is in W and the right side is in $W^\perp$. Therefore, $x_W - y_W$ is in W and in $W^\perp$, so it is orthogonal to itself, which implies $x_W - y_W = 0$. Hence $x_W = y_W$ and $x_{W^\perp} = y_{W^\perp}$, which proves uniqueness. □

Definition. Let W be a subspace of $\mathbf{R}^n$ and let x be a vector in $\mathbf{R}^n$. The expression

$$x = x_W + x_{W^\perp}$$

for x_W in W and $x_{W^\perp}$ in $W^\perp$, is called the **orthogonal decomposition** of x with respect to W, and the closest vector x_W is the **orthogonal projection** of x onto W.

Since x_W is the closest vector on W to x, the distance from x to the subspace W is the length of the vector from x_W to x, i.e., the length of $x_{W^\perp}$. To restate:

Closest vector and distance. Let W be a subspace of $\mathbf{R}^n$ and let x be a vector in $\mathbf{R}^n$.

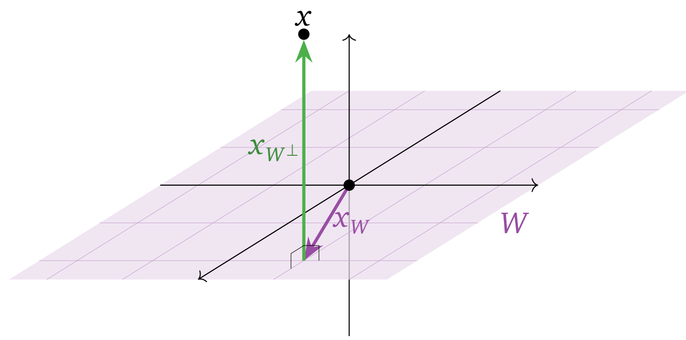
- The orthogonal projection x_W is the closest vector to x in W.
- The distance from x to W is $\|x_{W^\perp}\|$.

Example (Orthogonal decomposition with respect to the xy-plane)**.** Let W be the xy-plane in $\mathbf{R}^3$, so $W^\perp$ is the z-axis. It is easy to compute the orthogonal decomposition of a vector with respect to this W:

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \implies x_W = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \quad x_{W^\perp} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$$

$$x = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \implies x_W = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \quad x_{W^\perp} = \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}.$$

We see that the orthogonal decomposition in this case expresses a vector in terms of a "horizontal" component (in the xy-plane) and a "vertical" component (on the z-axis).



Use this link to view the online demo

Orthogonal decomposition of a vector with respect to the xy-plane in $\mathbf{R}^3$. Note that x_W is in the xy-plane and $x_{W^\perp}$ is in the z-axis. Click and drag the head of the vector x to see how the orthogonal decomposition changes.

Example (Orthogonal decomposition of a vector in W)**.** If x is in a subspace W, then the closest vector to x in W is itself, so $x = x_W$ and $x_{W^\perp} = 0$. Conversely, if $x = x_W$ then x is contained in W because x_W is contained in W.

Example (Orthogonal decomposition of a vector in $W^\perp$)**.** If W is a subspace and x is in $W^\perp$, then the orthogonal decomposition of x is $x = 0 + x$, where 0 is in W and x is in $W^\perp$. It follows that $x_W = 0$. Conversely, if $x_W = 0$ then the orthogonal decomposition of x is $x = x_W + x_{W^\perp} = 0 + x_{W^\perp}$, so $x = x_{W^\perp}$ is in $W^\perp$.

Interactive: Orthogonal decomposition in $\mathbf{R}^2$.

Use this link to view the online demo

Orthogonal decomposition of a vector with respect to a line W in $\mathbf{R}^2$. Note that x_W is in W and $x_{W^\perp}$ is in the line perpendicular to W. Click and drag the head of the vector x to see how the orthogonal decomposition changes.

Interactive: Orthogonal decomposition in $\mathbf{R}^3$.

Use this link to view the online demo

Orthogonal decomposition of a vector with respect to a plane W in $\mathbf{R}^3$. Note that x_W is in W and $x_{W^\perp}$ is in the line perpendicular to W. Click and drag the head of the vector x to see how the orthogonal decomposition changes.

Interactive: Orthogonal decomposition in $\mathbf{R}^3$.

Use this link to view the online demo

Orthogonal decomposition of a vector with respect to a line W in $\mathbf{R}^3$. Note that x_W is in W and $x_{W^\perp}$ is in the plane perpendicular to W. Click and drag the head of the vector x to see how the orthogonal decomposition changes.

Now we turn to the problem of computing x_W and $x_{W^\perp}$. Of course, since $x_{W^\perp} = x - x_W$, really all we need is to compute x_W. The following theorem gives a method for computing the orthogonal projection onto a column space. To compute the orthogonal projection onto a general subspace, usually it is best to rewrite the subspace as the column space of a matrix, as in this important note in Section 2.6.

Theorem. *Let A be an $m \times n$ matrix, let $W = \mathrm{Col}(A)$, and let x be a vector in $\mathbf{R}^m$. Then the matrix equation*

$$A^T Ac = A^T x$$

in the unknown vector c is consistent, and x_W is equal to Ac for any solution c.

Proof. Let $x = x_W + x_{W^\perp}$ be the orthogonal decomposition with respect to W. By definition x_W lies in $W = \mathrm{Col}(A)$ and so there is a vector c in $\mathbf{R}^n$ with $Ac = x_W$. Choose any such vector c. We know that $x - x_W = x - Ac$ lies in $W^\perp$, which is equal to $\mathrm{Nul}(A^T)$ by this important note in Section 6.2. We thus have

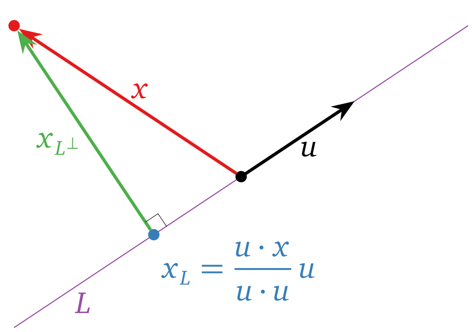
$$0 = A^T(x - Ac) = A^T x - A^T Ac$$

and so

$$A^T Ac = A^T x.$$

This exactly means that $A^T Ac = A^T x$ is consistent. If c is any solution to $A^T Ac = A^T x$ then by reversing the above logic, we conclude that $x_W = Ac$. □

Example (Orthogonal projection onto a line)**.** Let $L = \mathrm{Span}\{u\}$ be a line in $\mathbf{R}^n$ and let x be a vector in $\mathbf{R}^n$. By the theorem, to find x_L we must solve the matrix equation $u^T uc = u^T x$, where we regard u as an $n \times 1$ matrix (the column space of this matrix is exactly L!). But $u^T u = u \cdot u$ and $u^T x = u \cdot x$, so $c = (u \cdot x)/(u \cdot u)$ is a solution of $u^T uc = u^T x$, and hence $x_L = uc = (u \cdot x)/(u \cdot u)\,u$.



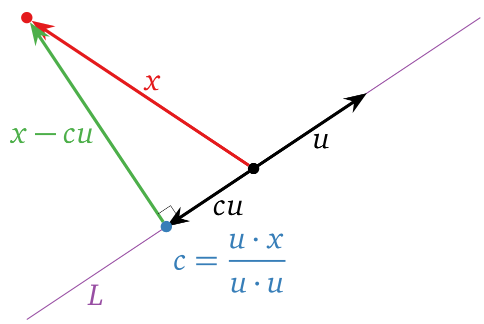
To reiterate:

Recipe: Orthogonal projection onto a line. If $L = \mathrm{Span}\{u\}$ is a line, then

$$x_L = \frac{u \cdot x}{u \cdot u}\, u \quad \text{and} \quad x_{L^\perp} = x - x_L$$

for any vector x.

Remark (Simple proof for the formula for projection onto a line)**.** In the special case where we are projecting a vector x in $\mathbf{R}^n$ onto a line $L = \mathrm{Span}\{u\}$, our formula for the projection can be derived very directly and simply. The vector x_L is a multiple of u, say $x_L = cu$. This multiple is chosen so that $x - x_L = x - cu$ is perpendicular to u, as in the following picture.



In other words,

$$(x - cu) \cdot u = 0.$$

Using the distributive property for the dot product and isolating the variable c gives us that

$$c = \frac{u \cdot x}{u \cdot u}$$

and so

$$x_L = cu = \frac{u \cdot x}{u \cdot u}\, u.$$

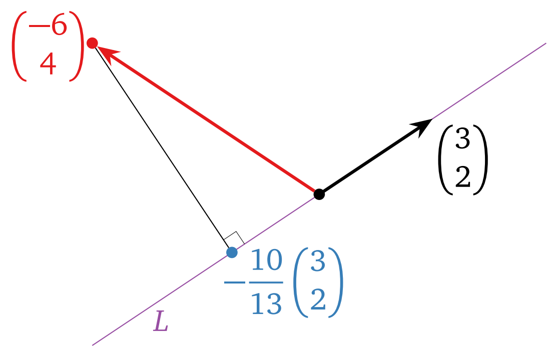
Example (Projection onto a line in $\mathbf{R}^2$)**.** Compute the orthogonal projection of $x = \binom{-6}{4}$ onto the line L spanned by $u = \binom{3}{2}$, and find the distance from x to L.

Solution. First we find

$$x_L = \frac{x \cdot u}{u \cdot u}\, u = \frac{-18 + 8}{9 + 4}\begin{pmatrix} 3 \\ 2 \end{pmatrix} = -\frac{10}{13}\begin{pmatrix} 3 \\ 2 \end{pmatrix} \qquad x_{L^\perp} = x - x_L = \frac{1}{13}\begin{pmatrix} -48 \\ 72 \end{pmatrix}.$$

The distance from x to L is

$$\|x_{L^\perp}\| = \frac{1}{13}\sqrt{48^2 + 72^2} \approx 6.656.$$



Use this link to view the online demo

Distance from the line L.

Example (Projection onto a line in $\mathbf{R}^3$)**.** Let

$$x = \begin{pmatrix}-2\\3\\-1\end{pmatrix} \qquad u = \begin{pmatrix}-1\\1\\1\end{pmatrix},$$

and let L be the line spanned by u. Compute x_L and $x_L^\perp$.

Solution.

$$x_L = \frac{x\cdot u}{u\cdot u}u = \frac{2+3-1}{1+1+1}\begin{pmatrix}-1\\1\\1\end{pmatrix} = \frac{4}{3}\begin{pmatrix}-1\\1\\1\end{pmatrix} \qquad x_{L^\perp} = x - x_L = \frac{1}{3}\begin{pmatrix}-2\\5\\-7\end{pmatrix}.$$

Use this link to view the online demo

Orthogonal projection onto the line L.

When A is a matrix with more than one column, computing the orthogonal projection of x onto $W = \mathrm{Col}(A)$ means solving the matrix equation $A^TAc = A^Tx$. In other words, we can compute the closest vector by *solving a system of linear* *equations*. To be explicit, we state the theorem as a recipe:

Recipe: Compute an orthogonal decomposition. Let W be a subspace of $\mathbf{R}^m$. Here is a method to compute the orthogonal decomposition of a vector x with respect to W:

1. Rewrite W as the column space of a matrix A. In other words, find a a spanning set for W, and let A be the matrix with those columns.
2. Compute the matrix $A^T A$ and the vector $A^T x$.
3. Form the augmented matrix for the matrix equation $A^T Ac = A^T x$ in the unknown vector c, and row reduce.
4. This equation is always consistent; choose one solution c. Then

$$x_W = Ac \qquad x_{W^\perp} = x - x_W.$$

Example (Projection onto the xy-plane)**.** Use the theorem to compute the orthogonal decomposition of a vector with respect to the xy-plane in $\mathbf{R}^3$.

Solution. A basis for the xy-plane is given by the two standard coordinate vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Let A be the matrix with columns e_1, e_2:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then

$$A^T A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 \qquad A^T \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

It follows that the unique solution c of $A^T Ac = I_2 c = A^T x$ is given by the first two coordinates of x, so

$$x_W = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \qquad x_{W^\perp} = x - x_W = \begin{pmatrix} 0 \\ 0 \\ x_3 \end{pmatrix}.$$

We have recovered this example.

Example (Projection onto a plane in $\mathbf{R}^3$)**.** Let

$$W = \text{Span}\left\{\begin{pmatrix}1\\0\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\right\} \qquad x = \begin{pmatrix}1\\2\\3\end{pmatrix}.$$

Compute x_W and the distance from x to W.

Solution. We have to solve the matrix equation $A^T Ac = A^T x$, where

$$A = \begin{pmatrix}1 & 1\\0 & 1\\-1 & 0\end{pmatrix}.$$

We have

$$A^T A = \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix} \qquad A^T x = \begin{pmatrix}-2\\3\end{pmatrix}.$$

We form an augmented matrix and row reduce:

$$\left(\begin{array}{cc|c}2 & 1 & -2\\1 & 2 & 3\end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c}1 & 0 & -7/3\\0 & 1 & 8/3\end{array}\right) \implies c = \frac{1}{3}\begin{pmatrix}-7\\8\end{pmatrix}.$$

It follows that

$$x_W = Ac = \frac{1}{3}\begin{pmatrix}1\\8\\7\end{pmatrix} \qquad x_{W^\perp} = x - x_W = \frac{1}{3}\begin{pmatrix}2\\-2\\2\end{pmatrix}.$$

The distance from x to W is

$$\|x_{W^\perp}\| = \frac{1}{3}\sqrt{4+4+4} \approx 1.155.$$

Use this link to view the online demo

Orthogonal projection onto the plane W.

Example (Projection onto another plane in $\mathbf{R}^3$)**.** Let

$$W = \left\{\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} \,\middle|\, x_1 - 2x_2 = x_3\right\} \quad \text{and} \quad x = \begin{pmatrix}1\\1\\1\end{pmatrix}.$$

Compute x_W.

Solution. *Method 1:* First we need to find a spanning set for W. We notice that W is the solution set of the homogeneous equation $x_1 - 2x_2 - x_3 = 0$, so $W = \text{Nul}\begin{pmatrix}1 & -2 & -1\end{pmatrix}$. We know how to compute a basis for a null space: we row reduce and find the parametric vector form. The matrix $\begin{pmatrix} 1 & -2 & -1 \end{pmatrix}$ is already in reduced row echelon form. The parametric form is $x_1 = 2x_2 + x_3$, so the parametric vector form is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

and hence a basis for V is given by

$$\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

We let A be the matrix whose columns are our basis vectors:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Hence $\mathrm{Col}(A) = \mathrm{Nul}\begin{pmatrix} 1 & -2 & -1 \end{pmatrix} = W$.

Now we can continue with step 1 of the recipe. We compute

$$A^T A = \begin{pmatrix} 5 & 2 \\ 2 & 2 \end{pmatrix} \qquad A^T x = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

We write the linear system $A^T Ac = A^T x$ as an augmented matrix and row reduce:

$$\left(\begin{array}{cc|c} 5 & 2 & 3 \\ 2 & 2 & 2 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & 2/3 \end{array}\right).$$

Hence we can take $c = \binom{1/3}{2/3}$, so

$$x_W = Ac = \begin{pmatrix} 2 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/3 \\ 2/3 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix}.$$

Use this link to view the online demo

Orthogonal projection onto the plane W.

Method 2: In this case, it is easier to compute $x_{W^\perp}$. Indeed, since $W = \mathrm{Nul}\begin{pmatrix} 1 & -2 & -1 \end{pmatrix}$, the orthogonal complement is the line

$$V = W^\perp = \mathrm{Col}\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}.$$

Using the formula for projection onto a line gives

$$x_{W^\perp} = x_V = \frac{\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\-2\\-1\end{pmatrix}}{\begin{pmatrix}1\\-2\\-1\end{pmatrix}\cdot\begin{pmatrix}1\\-2\\-1\end{pmatrix}}\begin{pmatrix}1\\-2\\-1\end{pmatrix} = \frac{1}{3}\begin{pmatrix}-1\\2\\1\end{pmatrix}.$$

Hence we have

$$x_W = x - x_{W^\perp} = \begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{1}{3}\begin{pmatrix}-1\\2\\1\end{pmatrix}. = \frac{1}{3}\begin{pmatrix}4\\1\\2\end{pmatrix},$$

as above.

Example (Projection onto a 3-space in $\mathbf{R}^4$)**.** Let

$$W = \operatorname{Span}\left\{\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\1\\-1\end{pmatrix}\right\} \qquad x = \begin{pmatrix}0\\1\\3\\4\end{pmatrix}.$$

Compute the orthogonal decomposition of x with respect to W.

Solution. We have to solve the matrix equation $A^TAc = A^Tx$, where

$$A = \begin{pmatrix}1&0&1\\0&1&1\\-1&0&1\\0&-1&-1\end{pmatrix}.$$

We compute

$$A^TA = \begin{pmatrix}2&0&0\\0&2&2\\0&2&4\end{pmatrix} \qquad A^Tx = \begin{pmatrix}-3\\-3\\0\end{pmatrix}.$$

We form an augmented matrix and row reduce:

$$\left(\begin{array}{ccc|c}2&0&0&-3\\0&2&2&-3\\0&2&4&0\end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{ccc|c}1&0&0&-3/2\\0&1&0&-3\\0&0&1&3/2\end{array}\right) \implies c = \frac{1}{2}\begin{pmatrix}-3\\-6\\3\end{pmatrix}.$$

It follows that

$$x_W = Ac = \frac{1}{2}\begin{pmatrix}0\\-3\\6\\3\end{pmatrix} \qquad x_{W^\perp} = \frac{1}{2}\begin{pmatrix}0\\5\\0\\5\end{pmatrix}.$$

In the context of the above recipe, if we start with a *basis* of W, then it turns out that the square matrix A^TA is automatically invertible! (It is always the case that A^TA is square and the equation $A^TAc = A^Tx$ is consistent, but A^TA need not be invertible in general.)

Corollary. *Let A be an $m\times n$ matrix with linearly independent columns and let $W = \mathrm{Col}(A)$. Then the $n\times n$ matrix A^TA is invertible, and for all vectors x in $\mathbf{R}^m$, we have*

$$x_W = A(A^TA)^{-1}A^Tx.$$

Proof. We will show that $\mathrm{Nul}(A^TA) = \{0\}$, which implies invertibility by the invertible matrix theorem in Section 5.1. Suppose that $A^TAc = 0$. Then $A^TAc = A^T0$, so $0_W = Ac$ by the theorem. But $0_W = 0$ (the orthogonal decomposition of the zero vector is just $0 = 0 + 0$), so $Ac = 0$, and therefore c is in $\mathrm{Nul}(A)$. Since the columns of A are linearly independent, we have $c = 0$, so $\mathrm{Nul}(A^TA) = 0$, as desired.

Let x be a vector in $\mathbf{R}^n$ and let c be a solution of $A^TAc = A^Tx$. Then $c = (A^TA)^{-1}A^Tx$, so $x_W = Ac = A(A^TA)^{-1}A^Tx$. □

The corollary applies in particular to the case where we have a subspace W of $\mathbf{R}^m$, and a basis $v_1, v_2, \ldots, v_n$ for W. To apply the corollary, we take A to be the $m\times n$ matrix with columns $v_1, v_2, \ldots, v_n$.

Example (Computing a projection)**.** Continuing with the above example, let

$$W = \mathrm{Span}\left\{\begin{pmatrix}1\\0\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\right\} \qquad x = \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}.$$

Compute x_W using the formula $x_W = A(A^TA)^{-1}A^Tx$.

Solution. Clearly the spanning vectors are noncollinear, so according to the corollary, we have $x_W = A(A^TA)^{-1}A^Tx$, where

$$A = \begin{pmatrix}1 & 1\\0 & 1\\-1 & 0\end{pmatrix}.$$

We compute

$$A^TA = \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix} \implies (A^TA)^{-1} = \frac{1}{3}\begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix},$$

so

$$x_W = A(A^TA)^{-1}A^Tx = \begin{pmatrix}1 & 1\\0 & 1\\-1 & 0\end{pmatrix}\frac{1}{3}\begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}\begin{pmatrix}1 & 0 & -1\\1 & 1 & 0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

$$= \frac{1}{3}\begin{pmatrix}2 & 1 & -1\\1 & 2 & 1\\-1 & 1 & 2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \frac{1}{3}\begin{pmatrix}2x_1 + x_2 - x_3\\x_1 + 2x_2 + x_3\\-x_1 + x_2 + 2x_3\end{pmatrix}.$$

So, for example, if $x = (1, 0, 0)$, this formula tells us that $x_W = (2, 1, -1)$.

6.3.2 Orthogonal Projection

In this subsection, we change perspective and think of the orthogonal projection x_W as a *function* of x. This function turns out to be a linear transformation with many nice properties, and is a good example of a linear transformation which is not originally defined as a matrix transformation.

Properties of Orthogonal Projections. *Let W be a subspace of $\mathbf{R}^n$, and define $T: \mathbf{R}^n \to \mathbf{R}^n$ by $T(x) = x_W$. Then:*

1. *T is a linear transformation.*
2. *$T(x) = x$ if and only if x is in W.*
3. *$T(x) = 0$ if and only if x is in $W^\perp$.*
4. *$T \circ T = T$.*
5. *The range of T is W.*

Proof.

1. We have to verify the defining properties of linearity in Section 3.3. Let x, y be vectors in $\mathbf{R}^n$, and let $x = x_W + x_{W^\perp}$ and $y = y_W + y_{W^\perp}$ be their orthogonal decompositions. Since W and $W^\perp$ are subspaces, the sums $x_W + y_W$ and $x_{W^\perp} + y_{W^\perp}$ are in W and $W^\perp$, respectively. Therefore, the orthogonal decomposition of $x + y$ is $(x_W + y_W) + (x_{W^\perp} + y_{W^\perp})$, so

$$T(x+y) = (x+y)_W = x_W + y_W = T(x) + T(y).$$

   Now let c be a scalar. Then cx_W is in W and $cx_{W^\perp}$ is in $W^\perp$, so the orthogonal decomposition of cx is $cx_W + cx_{W^\perp}$, and therefore,

$$T(cx) = (cx)_W = cx_W = cT(x).$$

   Since T satisfies the two defining properties in Section 3.3, it is a linear transformation.

2. See this example.
3. See this example.
4. For any x in $\mathbf{R}^n$ the vector $T(x)$ is in W, so $T \circ T(x) = T(T(x)) = T(x)$ by 2.
5. Any vector x in W is in the range of T, because $T(x) = x$ for such vectors. On the other hand, for any vector x in $\mathbf{R}^n$ the output $T(x) = x_W$ is in W, so W is the range of T.

□

We compute the standard matrix of the orthogonal projection in the same way as for any other transformation: by evaluating on the standard coordinate vectors. In this case, this means projecting the standard coordinate vectors onto the subspace.

Example (Matrix of a projection)**.** Let L be the line in $\mathbf{R}^2$ spanned by the vector $u = \binom{3}{2}$, and define $T : \mathbf{R}^2 \to \mathbf{R}^2$ by $T(x) = x_L$. Compute the standard matrix B for T.

Solution. The columns of B are $T(e_1) = (e_1)_L$ and $T(e_2) = (e_2)_L$. We have

$$\left.\begin{aligned} (e_1)_L &= \frac{u \cdot e_1}{u \cdot u} u = \frac{3}{13}\begin{pmatrix}3\\2\end{pmatrix} \\ (e_2)_L &= \frac{u \cdot e_2}{u \cdot u} u = \frac{2}{13}\begin{pmatrix}3\\2\end{pmatrix} \end{aligned}\right\} \implies B = \frac{1}{13}\begin{pmatrix}9 & 6\\ 6 & 4\end{pmatrix}.$$

Example (Matrix of a projection)**.** Let L be the line in $\mathbf{R}^2$ spanned by the vector

$$u = \begin{pmatrix}-1\\1\\1\end{pmatrix},$$

and define $T : \mathbf{R}^3 \to \mathbf{R}^3$ by $T(x) = x_L$. Compute the standard matrix B for T.

Solution. The columns of B are $T(e_1) = (e_1)_L$, $T(e_2) = (e_2)_L$, and $T(e_3) = (e_3)_L$. We have

$$\left.\begin{aligned} (e_1)_L &= \frac{u \cdot e_1}{u \cdot u} u = \frac{-1}{3}\begin{pmatrix}-1\\1\\1\end{pmatrix} \\ (e_2)_L &= \frac{u \cdot e_2}{u \cdot u} u = \frac{1}{3}\begin{pmatrix}-1\\1\\1\end{pmatrix} \\ (e_3)_L &= \frac{u \cdot e_3}{u \cdot u} u = \frac{1}{3}\begin{pmatrix}-1\\1\\1\end{pmatrix} \end{aligned}\right\} \implies B = \frac{1}{3}\begin{pmatrix}1 & -1 & -1\\ -1 & 1 & 1\\ -1 & 1 & 1\end{pmatrix}.$$

Example (Matrix of a projection)**.** Continuing with this example, let

$$W = \text{Span}\left\{\begin{pmatrix}1\\0\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\right\},$$

and define $T : \mathbf{R}^3 \to \mathbf{R}^3$ by $T(x) = x_W$. Compute the standard matrix B for T.

Solution. The columns of B are $T(e_1)=(e_1)_W$, $T(e_2)=(e_2)_W$, and $T(e_3)=(e_3)_W$. Let

$$A=\begin{pmatrix}1&1\\0&1\\-1&0\end{pmatrix}.$$

To compute each $(e_i)_W$, we solve the matrix equation $A^TAc=A^Te_i$ for c, then use the equality $(e_i)_W=Ac$. First we note that

$$A^TA=\begin{pmatrix}2&1\\1&2\end{pmatrix}; \qquad A^Te_i=\text{the } i\text{th column of } A^T=\begin{pmatrix}1&0&-1\\1&1&0\end{pmatrix}.$$

For e_1, we form an augmented matrix and row reduce:

$$\left(\begin{array}{cc|c}2&1&1\\1&2&1\end{array}\right)\xrightarrow{\text{RREF}}\left(\begin{array}{cc|c}1&0&1/3\\0&1&1/3\end{array}\right)\implies(e_1)_W=A\begin{pmatrix}1/3\\1/3\end{pmatrix}=\frac{1}{3}\begin{pmatrix}2\\1\\-1\end{pmatrix}.$$

We do the same for e_2:

$$\left(\begin{array}{cc|c}2&1&0\\1&2&1\end{array}\right)\xrightarrow{\text{RREF}}\left(\begin{array}{cc|c}1&0&-1/3\\0&1&2/3\end{array}\right)\implies(e_1)_W=A\begin{pmatrix}-1/3\\2/3\end{pmatrix}=\frac{1}{3}\begin{pmatrix}1\\2\\1\end{pmatrix}$$

and for e_3:

$$\left(\begin{array}{cc|c}2&1&-1\\1&2&0\end{array}\right)\xrightarrow{\text{RREF}}\left(\begin{array}{cc|c}1&0&-2/3\\0&1&1/3\end{array}\right)\implies(e_1)_W=A\begin{pmatrix}-2/3\\1/3\end{pmatrix}=\frac{1}{3}\begin{pmatrix}-1\\1\\2\end{pmatrix}.$$

It follows that

$$B=\frac{1}{3}\begin{pmatrix}2&1&-1\\1&2&1\\-1&1&2\end{pmatrix}.$$

In the previous example, we could have used the fact that

$$\left\{\begin{pmatrix}1\\0\\-1\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix}\right\}$$

forms a *basis* for W, so that

$$T(x)=x_W=\left[A(A^TA)^{-1}A^T\right]x \quad \text{for} \quad A=\begin{pmatrix}1&1\\0&1\\-1&0\end{pmatrix}$$

by the corollary. In this case, we have already expressed T as a matrix transformation with matrix $A(A^TA)^{-1}A^T$. See this example.

Let W be a subspace of $\mathbf{R}^n$ with basis $v_1, v_2, \ldots, v_m$, and let A be the matrix with columns $v_1, v_2, \ldots, v_m$. Then the standard matrix for $T(x) = x_W$ is

$$A(A^TA)^{-1}A^T.$$

We can translate the above properties of orthogonal projections into properties of the associated standard matrix.

Properties of Projection Matrices. *Let W be a subspace of $\mathbf{R}^n$, define $T : \mathbf{R}^n \to \mathbf{R}^n$ by $T(x) = x_W$, and let B be the standard matrix for T. Then:*

1. $\mathrm{Col}(B) = W$.
2. $\mathrm{Nul}(B) = W^\perp$.
3. $B^2 = B$.
4. *If $W \neq \{0\}$, then 1 is an eigenvalue of B and the 1-eigenspace for B is W.*
5. *If $W \neq \mathbf{R}^n$, then 0 is an eigenvalue of B and the 0-eigenspace for B is $W^\perp$.*
6. *B is similar to the diagonal matrix with m ones and $n-m$ zeros on the diagonal, where $m = \dim(W)$.*

Proof. The first four assertions are translations of properties 5, 3, 4, and 2, respectively, using this important note in Section 3.1 and this theorem in Section 3.4. The fifth assertion is equivalent to the second, by this fact in Section 5.1.

For the final assertion, we showed in the proof of this theorem that there is a basis of $\mathbf{R}^n$ of the form $\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$, where $\{v_1, \ldots, v_m\}$ is a basis for W and $\{v_{m+1}, \ldots, v_n\}$ is a basis for $W^\perp$. Each v_i is an eigenvector of B: indeed, for $i \leq m$ we have

$$Bv_i = T(v_i) = v_i = 1 \cdot v_i$$

because v_i is in W, and for $i > m$ we have

$$Bv_i = T(v_i) = 0 = 0 \cdot v_i$$

because v_i is in $W^\perp$. Therefore, we have found a basis of eigenvectors, with associated eigenvalues $1, \ldots, 1, 0, \ldots, 0$ (m ones and $n - m$ zeros). Now we use the diagonalization theorem in Section 5.4. □

We emphasize that the properties of projection matrices would be very hard to prove in terms of matrices. By translating all of the statements into statements about linear transformations, they become much more transparent. For example, consider the projection matrix we found in this example. Just by looking at the matrix it is not at all obvious that when you square the matrix you get the same matrix back.

Example. Continuing with the above example, we showed that

$$B = \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}$$

is the standard matrix of the orthogonal projection onto

$$W = \text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

One can verify by hand that $B^2 = B$ (try it!). We compute $W^\perp$ as the null space of

$$\begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \end{pmatrix} \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}.$$

The free variable is x_3, and the parametric form is $x_1 = x_3$, $x_2 = -x_3$, so that

$$W^\perp = \text{Span}\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

It follows that B has eigenvectors

$$\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

with eigenvalues $1, 1, 0$, respectively, so that

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}^{-1}.$$

Remark. As we saw in this example, if you are willing to compute bases for W and $W^\perp$, then this provides a third way of finding the standard matrix B for projection onto W: indeed, if $\{v_1, v_2, \ldots, v_m\}$ is a basis for W and $\{v_{m+1}, v_{m+2}, \ldots, v_n\}$ is a basis for $W^\perp$, then

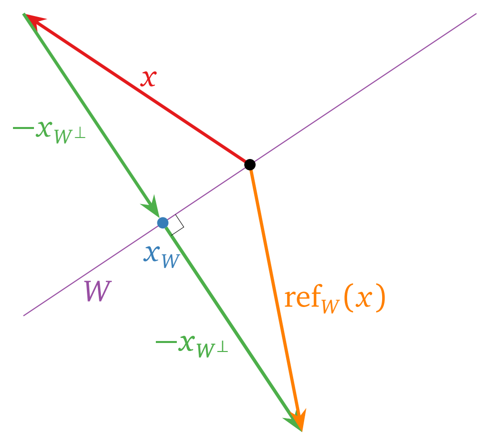
$$B = \begin{pmatrix} | & | & & | \\ v_1 & v_1 & \cdots & v_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} | & | & & | \\ v_1 & v_1 & \cdots & v_n \\ | & | & & | \end{pmatrix}^{-1},$$

where the middle matrix in the product is the diagonal matrix with m ones and $n - m$ zeros on the diagonal. However, since you already have a basis for W, it is faster to multiply out the expression $A(A^TA)^{-1}A^T$ as in the corollary.

Remark (Reflections)**.** Let W be a subspace of $\mathbf{R}^n$, and let x be a vector in $\mathbf{R}^n$. The *reflection* of x over W is defined to be the vector

$$\mathrm{ref}_W(x) = x - 2x_{W^\perp}.$$

In other words, to find $\mathrm{ref}_W(x)$ one starts at x, then moves to $x - x_{W^\perp} = x_W$, then continues in the same direction one more time, to end on the opposite side of W.



Since $x_{W^\perp} = x - x_W$, we also have

$$\mathrm{ref}_W(x) = x - 2(x - x_W) = 2x_W - x.$$

We leave it to the reader to check using the definition that:

1. $\mathrm{ref}_W \circ \mathrm{ref}_W = \mathrm{Id}_{\mathbf{R}^n}$.
2. The 1-eigenspace of ref_W is W, and the -1-eigenspace of ref_W is $W^\perp$.
3. ref_W is similar to the diagonal matrix with $m = \dim(W)$ ones on the diagonal and $n - m$ negative ones.

6.4 Orthogonal Sets

Objectives

1. Understand which is the best method to use to compute an orthogonal projection in a given situation.
2. *Recipes:* an orthonormal set from an orthogonal set, Projection Formula, $\mathcal{B}$-coordinates when $\mathcal{B}$ is an orthogonal set, Gram–Schmidt process.
3. *Vocabulary words:* **orthogonal set**, **orthonormal set**.

In this section, we give a formula for orthogonal projection that is considerably simpler than the one in Section 6.3, in that it does not require row reduction or matrix inversion. However, this formula, called the Projection Formula, only works in the presence of an *orthogonal* basis. We will also present the Gram–Schmidt process for turning an arbitrary basis into an orthogonal one.

6.4.1 Orthogonal Sets and the Projection Formula

Computations involving projections tend to be much easier in the presence of an *orthogonal* set of vectors.

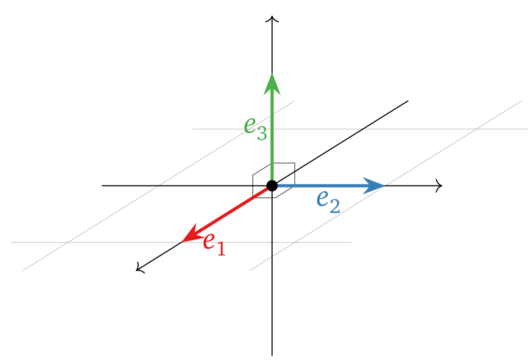
Definition. A set of *nonzero* vectors $\{u_1, u_2, \ldots, u_m\}$ is called **orthogonal** if $u_i \cdot u_j = 0$ whenever $i \neq j$. It is **orthonormal** if it is orthogonal, and in addition $u_i \cdot u_i = 1$ for all $i = 1, 2, \ldots, m$.

In other words, a set of vectors is orthogonal if different vectors in the set are perpendicular to each other. An orthonormal set is an orthogonal set of unit vectors.

Example. The standard coordinate vectors in $\mathbf{R}^n$ always form an orthonormal set. For instance, in $\mathbf{R}^3$ we check that

$$\begin{pmatrix}1\\0\\0\end{pmatrix} \cdot \begin{pmatrix}0\\1\\0\end{pmatrix} = 0 \qquad \begin{pmatrix}1\\0\\0\end{pmatrix} \cdot \begin{pmatrix}0\\0\\1\end{pmatrix} = 0 \qquad \begin{pmatrix}0\\1\\0\end{pmatrix} \cdot \begin{pmatrix}0\\0\\1\end{pmatrix} = 0.$$

Since $e_i \cdot e_i = 1$ for all $i = 1, 2, 3$, this shows that $\{e_1, e_2, e_3\}$ is orthonormal.



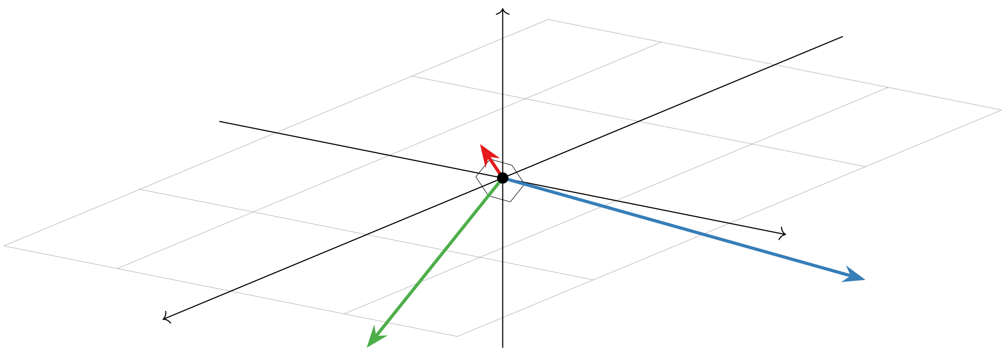
Example. Is this set orthogonal? Is it orthonormal?

$$\mathcal{B} = \left\{ \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\-2\\1\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix} \right\}$$

Solution. We check that

$$\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\-2\\1\end{pmatrix}=0 \qquad \begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\0\\-1\end{pmatrix}=0 \qquad \begin{pmatrix}1\\-2\\1\end{pmatrix}\cdot\begin{pmatrix}1\\0\\-1\end{pmatrix}=0.$$

Therefore, $\mathcal{B}$ is orthogonal.



The set $\mathcal{B}$ is not orthonormal because, for instance,

$$\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}=3\neq 1.$$

However, we can make it orthonormal by replacing each vector by the unit vector in the direction of each vector:

$$\left\{\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix},\ \frac{1}{\sqrt{6}}\begin{pmatrix}1\\-2\\1\end{pmatrix},\ \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix}\right\}$$

is orthonormal.

We saw in the previous example that it is easy to produce an orthonormal set of vectors from an orthogonal one by replacing each vector with the unit vector in the same direction.

Recipe: An orthonormal set from an orthogonal set. If $\{v_1, v_2, \ldots, v_m\}$ is an orthogonal set of vectors, then

$$\left\{\frac{v_1}{\|v_1\|}, \frac{v_2}{\|v_2\|}, \ldots, \frac{v_m}{\|v_m\|}\right\}$$

is an orthonormal set.

Example. Let a, b be scalars, and let

$$u_1 = \begin{pmatrix} a \\ b \end{pmatrix} \qquad u_2 = \begin{pmatrix} -b \\ a \end{pmatrix}.$$

Is $\mathcal{B} = \{u_1, u_2\}$ orthogonal?

Solution. We only have to check that

$$\begin{pmatrix} a \\ b \end{pmatrix} \cdot \begin{pmatrix} -b \\ a \end{pmatrix} = -ab + ab = 0.$$

Therefore, $\{u_1, u_2\}$ is orthogonal, *unless* $a = b = 0$.

Non-Example. Is this set orthogonal?

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \right\}$$

Solution. This set is not orthogonal because

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = 1 - 1 - 1 = -1 \neq 0.$$

We will see how to produce an orthogonal set from $\mathcal{B}$ in this subsection.

A nice property enjoyed by orthogonal sets is that they are automatically linearly independent.

Fact. *An orthogonal set is linearly independent. Therefore, it is a basis for its span.*

Proof. Suppose that $\{u_1, u_2, \ldots, u_m\}$ is orthogonal. We need to show that the equation

$$c_1 u_1 + c_2 u_2 + \cdots + c_m u_m = 0$$

has only the trivial solution $c_1 = c_2 = \cdots = c_m = 0$. Taking the dot product of both sides of this equation with u_1 gives

$$\begin{aligned} 0 = u_1 \cdot 0 = u_1 \cdot \big(c_1 u_1 + c_2 u_2 + \cdots + c_m u_m\big) \\ = c_1(u_1 \cdot u_1) + c_2(u_1 \cdot u_2) + \cdots + c_m(u_1 \cdot u_m) \\ = c_1(u_1 \cdot u_1) \end{aligned}$$

because $u_1 \cdot u_i = 0$ for $i > 1$. Since $u_1 \neq 0$ we have $u_1 \cdot u_1 \neq 0$, so $c_1 = 0$. Similarly, taking the dot product with u_i shows that each $c_i = 0$, as desired. □

One advantage of working with orthogonal sets is that it gives a simple formula for the orthogonal projection of a vector.

Projection Formula. *Let W be a subspace of $\mathbf{R}^n$, and let $\{u_1, u_2, \ldots, u_m\}$ be an* orthogonal *basis for W. Then for any vector x in $\mathbf{R}^n$, the orthogonal projection of x onto W is given by the formula*

$$x_W = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 + \cdots + \frac{x \cdot u_m}{u_m \cdot u_m} u_m.$$

Proof. Let

$$y = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 + \cdots + \frac{x \cdot u_m}{u_m \cdot u_m} u_m.$$

This vector is contained in W because it is a linear combination of $u_1, u_2, \ldots, u_m$. Hence we just need to show that $x - y$ is in $W^\perp$, i.e., that $u_i \cdot (x - y) = 0$ for each $i = 1, 2, \ldots, m$. For u_1, we have

$$\begin{aligned} u_1 \cdot (x - y) &= u_1 \cdot \left(x - \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 - \cdots - \frac{x \cdot u_m}{u_m \cdot u_m} u_m \right) \\ &= u_1 \cdot x - \frac{x \cdot u_1}{u_1 \cdot u_1}(u_1 \cdot u_1) - 0 - \cdots - 0 \\ &= 0. \end{aligned}$$

A similar calculation shows that $u_i \cdot (x - y) = 0$ for each i, so $x - y$ is in $W^\perp$, as desired. □

If $\{u_1, u_2, \ldots, u_m\}$ is an ortho*normal* basis for W, then the denominators $u_i \cdot u_i = 1$ go away, so the projection formula becomes even simpler:

$$x_W = (x \cdot u_1)u_1 + (x \cdot u_2)u_2 + \cdots + (x \cdot u_m)u_m.$$

Example. Suppose that $L = \mathrm{Span}\{u\}$ is a line. The set $\{u\}$ is an orthogonal basis for L, so the Projection Formula says that for any vector x, we have

$$x_L = \frac{x \cdot u}{u \cdot u} u,$$

as in this example in Section 6.3. See also this example in Section 6.3 and this example in Section 6.3.

Suppose that $\{u_1, u_2, \ldots, u_m\}$ is an orthogonal basis for a subspace W, and let $L_i = \mathrm{Span}\{u_i\}$ for each $i = 1, 2, \ldots, m$. Then we see that for any vector x, we have

$$\begin{aligned} x_W &= \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 + \cdots + \frac{x \cdot u_m}{u_m \cdot u_m} u_m \\ &= \quad x_{L_1} \quad + \quad x_{L_2} \quad + \cdots + \quad x_{L_m} \quad . \end{aligned}$$

In other words, for an *orthogonal* basis, the projection of x onto W is the sum of the projections onto the lines spanned by the basis vectors. In this sense, projection onto a line is the most important example of an orthogonal projection.

Example (Projection onto the xy-plane)**.** Continuing with this example in Section 6.3 and this example in Section 6.3, use the projection formula to compute the orthogonal projection of a vector onto the xy-plane in $\mathbf{R}^3$.

Solution. A basis for the xy-plane is given by the two standard coordinate vectors

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

The set $\{e_1, e_2\}$ is orthogonal, so for any vector $x = (x_1, x_2, x_3)$, we have

$$x_W = \frac{x \cdot e_1}{e_1 \cdot e_1} e_1 + \frac{x \cdot e_2}{e_2 \cdot e_2} e_2 = x_1 e_1 + x_2 e_2 = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}.$$

Use this link to view the online demo

Orthogonal projection of a vector onto the xy-plane in $\mathbf{R}^3$. Note that x_W is the sum of the projections of x onto the e_1- and e_2-coordinate axes (shown in orange and brown, respectively).

Example (Projection onto a plane in $\mathbf{R}^3$)**.** Let

$$W = \text{Span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \qquad x = \begin{pmatrix} 2 \\ 3 \\ -2 \end{pmatrix}.$$

Find x_W and $x_{W^\perp}$.

Solution. The vectors

$$u_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \qquad u_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

are orthogonal, so we can use the Projection Formula:

$$x_W = \frac{x \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{x \cdot u_2}{u_2 \cdot u_2} u_2 = \frac{4}{2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \frac{3}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}.$$

Then we have

$$x_{W^\perp} = x - x_W = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.$$

Use this link to view the online demo

Orthogonal projection of a vector onto the plane W. Note that x_W is the sum of the projections of x onto the lines spanned by u_1 and u_2 (shown in orange and brown, respectively).

Example (Projection onto a 3-space in $\mathbf{R}^4$)**.** Let

$$W = \operatorname{Span}\left\{\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\1\\1\end{pmatrix}\right\} \qquad x = \begin{pmatrix}0\\1\\3\\4\end{pmatrix}.$$

Compute x_W, and find the distance from x to W.

Solution. The vectors

$$u_1 = \begin{pmatrix}1\\0\\-1\\0\end{pmatrix} \qquad u_2 = \begin{pmatrix}0\\1\\0\\-1\end{pmatrix} \qquad u_3 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}$$

are orthogonal, so we can use the Projection Formula:

$$\begin{aligned} x_W &= \frac{x\cdot u_1}{u_1\cdot u_1}u_1 + \frac{x\cdot u_2}{u_2\cdot u_2}u_2 + \frac{x\cdot u_3}{u_3\cdot u_3}u_3 \\ &= \frac{-3}{2}\begin{pmatrix}1\\0\\-1\\0\end{pmatrix} + \frac{-3}{2}\begin{pmatrix}0\\1\\0\\-1\end{pmatrix} + \frac{8}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1\\1\\7\\7\end{pmatrix} \\ x_{W^\perp} &= x - x_W = \frac{1}{2}\begin{pmatrix}-1\\1\\-1\\1\end{pmatrix}. \end{aligned}$$

The distance from x to W is

$$\|x_{W^\perp}\| = \frac{1}{2}\sqrt{1+1+1+1} = 1.$$

Now let W be a subspace of $\mathbf{R}^n$ with orthogonal basis $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$, and let x be a vector in W. Then $x = x_W$, so by the projection formula, we have

$$x = x_W = \frac{x\cdot u_1}{u_1\cdot u_1}u_1 + \frac{x\cdot u_2}{u_2\cdot u_2}u_2 + \cdots + \frac{x\cdot u_m}{u_m\cdot u_m}u_m.$$

This gives us a way of expressing x as a linear combination of the basis vectors in $\mathcal{B}$: we have computed the $\mathcal{B}$-coordinates of x without row reducing!

Recipe: $\mathcal{B}$-coordinates when $\mathcal{B}$ is an orthogonal set. Let W be a subspace of $\mathbf{R}^n$ with orthogonal basis $\mathcal{B} = \{v_1, v_2, \ldots, v_m\}$ and let x be a vector in W. Then

$$[x]_{\mathcal{B}} = \left(\frac{x \cdot u_1}{u_1 \cdot u_1},\ \frac{x \cdot u_2}{u_2 \cdot u_2},\ \ldots,\ \frac{x \cdot u_m}{u_m \cdot u_m}\right).$$

As with orthogonal projections, if $\{u_1, u_2, \ldots, u_m\}$ is an ortho*normal* basis of W, then the formula is even simpler:

$$[x]_{\mathcal{B}} = \left(x \cdot u_1,\ x \cdot u_2,\ \ldots,\ x \cdot u_m\right).$$

Example (Computing coordinates with respect to an orthogonal basis)**.** Find the $\mathcal{B}$-coordinates of x, where

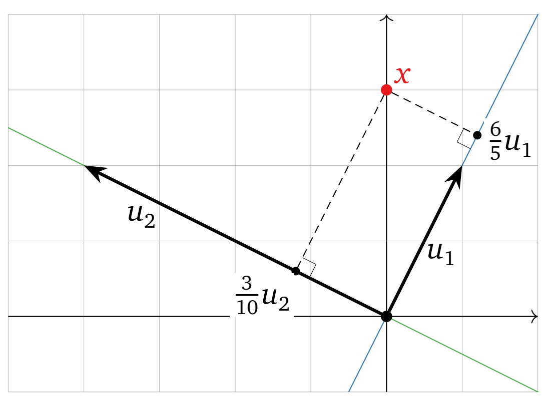
$$\mathcal{B} = \left\{\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}-4\\2\end{pmatrix}\right\} \qquad x = \begin{pmatrix}0\\3\end{pmatrix}.$$

Solution. Since

$$u_1 = \begin{pmatrix}1\\2\end{pmatrix} \qquad u_2 = \begin{pmatrix}-4\\2\end{pmatrix}$$

form an *orthogonal* basis of $\mathbf{R}^2$, we have

$$[x]_{\mathcal{B}} = \left(\frac{x \cdot u_1}{u_1 \cdot u_1},\ \frac{x \cdot u_2}{u_2 \cdot u_2}\right) = \left(\frac{3 \cdot 2}{1^2 + 2^2},\ \frac{3 \cdot 2}{(-4)^2 + 2^2}\right) = \left(\frac{6}{5},\ \frac{3}{10}\right).$$



Use this link to view the online demo

Computing $\mathcal{B}$-coordinates using the Projection Formula.

The following example shows that the Projection Formula does in fact require an orthogonal basis.

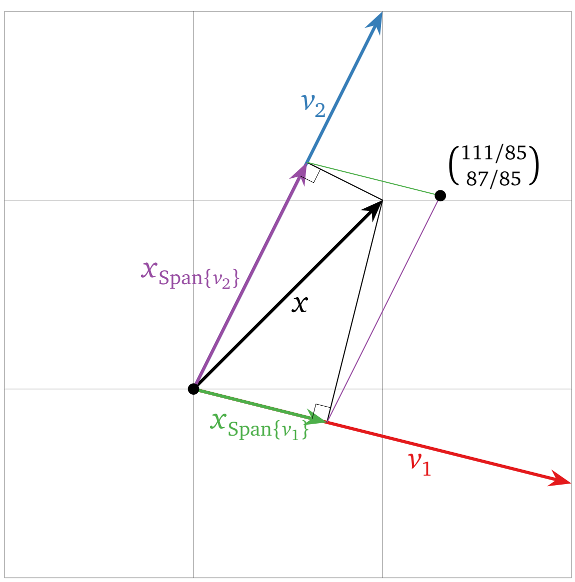
Non-Example (A non-orthogonal basis)**.** Consider the basis $\mathcal{B} = \{v_1, v_2\}$ of $\mathbf{R}^2$, where

$$v_1 = \begin{pmatrix} 2 \\ -1/2 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

This is not orthogonal because $v_1 \cdot v_2 = 2 - 1 = 1 \neq 0$. Let $x = \binom{1}{1}$. Let us try to compute $x = x_{\mathbf{R}^2}$ using the Projection Formula with respect to the basis $\mathcal{B}$:

$$x_{\mathbf{R}^2} = \frac{x \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{x \cdot v_2}{v_2 \cdot v_2} v_2 = \frac{3/2}{17/4} \begin{pmatrix} 2 \\ -1/2 \end{pmatrix} + \frac{3}{5} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 111/85 \\ 87/85 \end{pmatrix} \neq x.$$

Since $x = x_{\mathbf{R}^2}$, we see that the Projection Formula does not compute the orthogonal projection in this case. Geometrically, the projections of x onto the lines spanned by v_1 and v_2 do not sum to x, as we can see from the picture.



Use this link to view the online demo

When v_1 and v_2 are not orthogonal, then $x_{\mathbf{R}^2} = x$ is not necessarily equal to the sum (red) of the projections (orange and brown) of x onto the lines spanned by v_1 and v_2.

You need an orthogonal basis to use the Projection Formula.

6.4.2 The Gram–Schmidt Process

We saw in the previous subsection that orthogonal projections and $\mathcal{B}$-coordinates are much easier to compute in the presence of an *orthogonal* basis for a subspace. In this subsection, we give a method, called the *Gram–Schmidt Process*, for computing an orthogonal basis of a subspace.

The Gram–Schmidt Process. *Let $v_1, v_2, \ldots, v_m$ be a basis for a subspace W of $\mathbf{R}^n$. Define:*

1. $u_1 = v_1$
2. $u_2 = (v_2)_{\mathrm{Span}\{u_1\}^\perp} = v_2 - \dfrac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1$
3. $u_3 = (v_3)_{\mathrm{Span}\{u_1,u_2\}^\perp} = v_3 - \dfrac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \dfrac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2$

$\vdots$

m. $u_m = (v_m)_{\mathrm{Span}\{u_1,u_2,\ldots,u_{m-1}\}^\perp} = v_m - \displaystyle\sum_{i=1}^{m-1} \frac{v_m \cdot u_i}{u_i \cdot u_i} u_i.$

Then $\{u_1, u_2, \ldots, u_m\}$ is an orthogonal basis for the same subspace W.

Proof. First we claim that each u_i is in W, and in fact that u_i is in $\mathrm{Span}\{v_1, v_2, \ldots, v_i\}$. Clearly $u_1 = v_1$ is in $\mathrm{Span}\{v_1\}$. Then u_2 is a linear combination of u_1 and v_2, which are both in $\mathrm{Span}\{v_1, v_2\}$, so u_2 is in $\mathrm{Span}\{v_1, v_2\}$ as well. Similarly, u_3 is a linear combination of u_1, u_2, and v_3, which are all in $\mathrm{Span}\{v_1, v_2, v_3\}$, so u_3 is in $\mathrm{Span}\{v_1, v_2, v_3\}$. Continuing in this way, we see that each u_i is in $\mathrm{Span}\{v_1, v_2, \ldots, v_i\}$.

Now we claim that $\{u_1, u_2, \ldots, u_m\}$ is an orthogonal set. Let $1 \le i < j \le m$. Then $u_j = (v_j)_{\mathrm{Span}\{u_1,u_2,\ldots,u_{j-1}\}^\perp}$, so by definition u_j is orthogonal to every vector in $\mathrm{Span}\{u_1, u_2, \ldots, u_{j-1}\}$. In particular, u_j is orthogonal to u_i.

We still have to prove that each u_i is nonzero. Clearly $u_1 = v_1 \neq 0$. Suppose that $u_i = 0$. Then $(v_i)_{\mathrm{Span}\{u_1,u_2,\ldots,u_{i-1}\}^\perp} = 0$, which means that v_i is in $\mathrm{Span}\{u_1, u_2, \ldots, u_{i-1}\}$. But each $u_1, u_2, \ldots, u_{i-1}$ is in $\mathrm{Span}\{v_1, v_2, \ldots, v_{i-1}\}$ by the first paragraph, so v_i is in $\mathrm{Span}\{v_1, v_2, \ldots, v_{i-1}\}$. This contradicts the increasing span criterion in Section 2.5; therefore, u_i must be nonzero.

The previous two paragraphs justify the use of the projection formula in the equalities

$$(v_i)_{\mathrm{Span}\{u_1,u_2,\ldots,u_{i-1}\}^\perp} = v_i - (v_i)_{\mathrm{Span}\{u_1,u_2,\ldots,u_{i-1}\}} = v_i - \sum_{j=1}^{i-1} \frac{v_i \cdot u_j}{u_j \cdot u_j} u_j$$

in the statement of the theorem.

Since $\{u_1, u_2, \ldots, u_m\}$ is an orthogonal set, it is linearly independent. Thus it is a set of m linearly independent vectors in W, so it is a basis for W by the basis theorem in Section 2.7. Similarly, for every i, we saw that the set $\{u_1, u_2, \ldots, u_i\}$ is contained in the i-dimensional subspace $\mathrm{Span}\{v_1, v_2, \ldots, v_i\}$, so $\{u_1, u_2, \ldots, u_i\}$ is an orthogonal basis for $\mathrm{Span}\{v_1, v_2, \ldots, v_i\}$. □

Example (Two vectors)**.** Find an orthogonal basis $\{u_1, u_2\}$ for $W = \mathrm{Span}\{v_1, v_2\}$, where

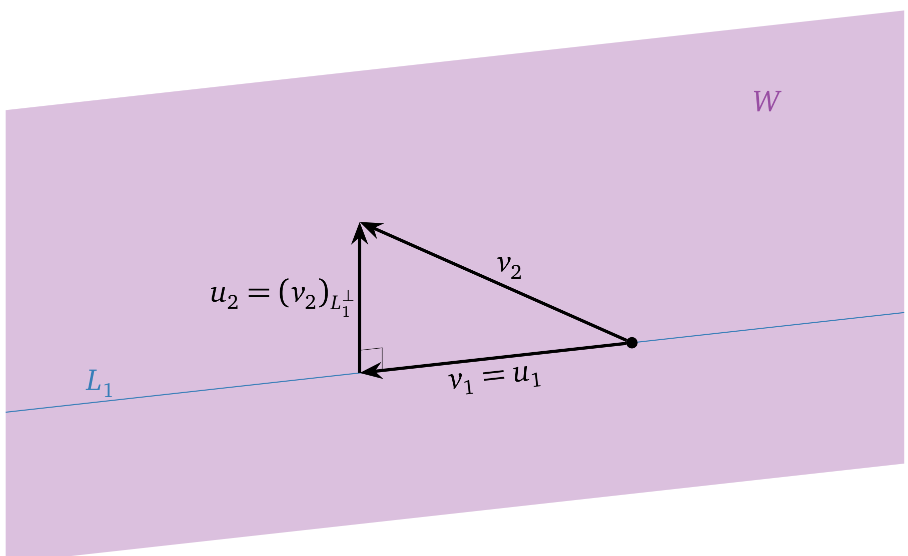
$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Solution. We run Gram–Schmidt: first take $u_1 = v_1$, then

$$u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2}{2}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Then $\{u_1, u_2\}$ is an orthogonal basis for W: indeed, it is clear that $u_1 \cdot u_2 = 0$.

Geometrically, we are simply replacing v_2 with the part of v_2 that is perpendicular to the line $L_1 = \mathrm{Span}\{v_1\}$:



Example (Three vectors)**.** Find an orthogonal basis $\{u_1, u_2, u_3\}$ for $W = \mathrm{Span}\{v_1, v_2, v_3\} = \mathbf{R}^3$, where

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad v_3 = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}.$$

Solution. We run Gram–Schmidt:

1. $u_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

2. $u_2 = v_2 - \dfrac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \dfrac{2}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

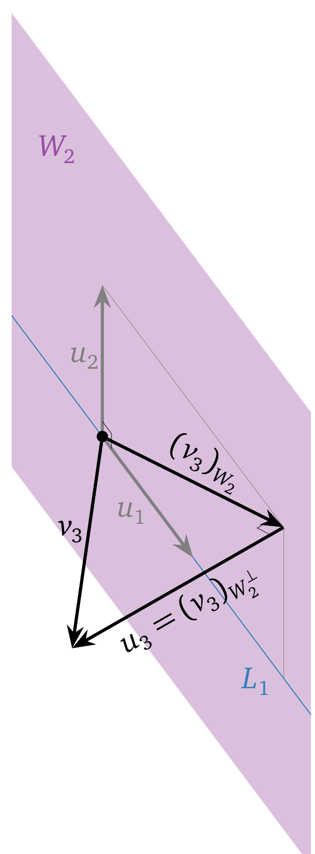
3. $u_3 = v_3 - \dfrac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \dfrac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2$

$$= \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} - \frac{4}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{1} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Then $\{u_1, u_2, u_3\}$ is an orthogonal basis for W: indeed, we have

$$u_1 \cdot u_2 = 0 \qquad u_1 \cdot u_3 = 0 \qquad u_2 \cdot u_3 = 0.$$

Geometrically, once we have u_1 and u_2, we replace v_3 by the part that is orthogonal to $W_2 = \text{Span}\{u_1, u_2\} = \text{Span}\{v_1, v_2\}$:



Example (Three vectors in $\mathbf{R}^4$)**.** Find an orthogonal basis $\{u_1, u_2, u_3\}$ for $W = \text{Span}\{v_1, v_2, v_3\}$, where

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad v_2 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix} \qquad v_3 = \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix}.$$

Solution. We run Gram–Schmidt:

1. $u_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

2. $u_2 = v_2 - \dfrac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix} - \dfrac{6}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix}$

3. $u_3 = v_3 - \dfrac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \dfrac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2$

$$= \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix} - \frac{0}{24} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{-20}{25} \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix}.$$

Then $\{u_1, u_2, u_3\}$ is an orthogonal basis for W.

We saw in the proof of the Gram–Schmidt Process that for every i between 1 and m, the set $\{u_1, u_2, \ldots, u_i\}$ is a an orthogonal basis for $\text{Span}\{v_1, v_2, \ldots, v_i\}$.

If we had started with a spanning set $\{v_1, v_2, \ldots, v_m\}$ which is linearly dependent, then for some i, the vector v_i is in $\text{Span}\{v_1, v_2, \ldots, v_{i-1}\}$ by the increasing span criterion in Section 2.5. Hence

$$0 = (v_i)_{\text{Span}\{v_1, v_2, \ldots, v_{i-1}\}^\perp} = (v_i)_{\text{Span}\{u_1, u_2, \ldots, u_{i-1}\}^\perp} = u_i.$$

You can use the Gram–Schmidt Process to produce an orthogonal basis from any spanning set: if some $u_i = 0$, just throw away u_i and v_i, and continue.

6.4.3 Two Methods to Compute the Projection

We have now presented *two* methods for computing the orthogonal projection of a vector: this theorem in Section 6.3 involves row reduction, and the projection formula requires an orthogonal basis. Here are some guidelines for which to use in a given situation.

1. If you already have an orthogonal basis, it is almost always easier to use the projection formula. This often happens in the sciences.

2. If you are going to have to compute the projections of many vectors onto the same subspace, it is worth your time to run Gram–Schmidt to produce an orthogonal basis, so that you can use the projection formula.

3. If you only have to project one or a few vectors onto a subspace, it is faster to use the theorem in Section 6.3. This is the method we will follow in Section 6.5.

6.5 The Method of Least Squares

Objectives

1. Learn examples of best-fit problems.

2. Learn to turn a best-fit problem into a least-squares problem.

3. *Recipe:* find a least-squares solution (two ways).

4. *Picture:* geometry of a least-squares solution.

5. *Vocabulary words:* **least-squares solution**.

In this section, we answer the following important question:

> Suppose that $Ax = b$ does not have a solution. What is the best approximate solution?

For our purposes, the best approximate solution is called the *least-squares solution*. We will present two methods for finding least-squares solutions, and we will give several applications to best-fit problems.

6.5.1 Least-Squares Solutions

We begin by clarifying exactly what we will mean by a "best approximate solution" to an inconsistent matrix equation $Ax = b$.

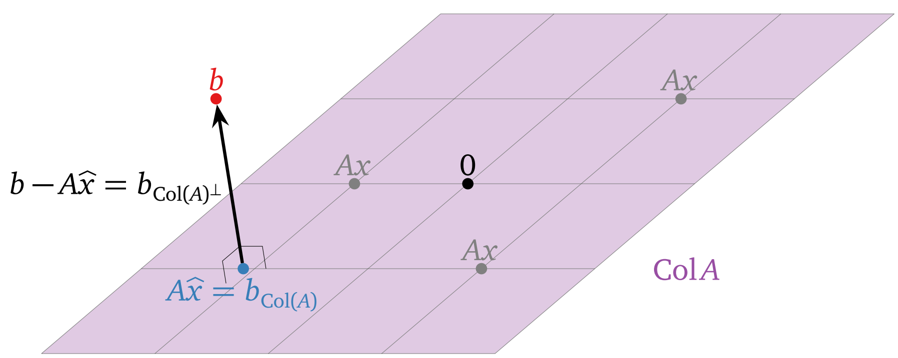
Definition. Let A be an $m \times n$ matrix and let b be a vector in $\mathbf{R}^m$. A **least-squares solution** of the matrix equation $Ax = b$ is a vector $\widehat{x}$ in $\mathbf{R}^n$ such that

$$\text{dist}(b, A\widehat{x}) \leq \text{dist}(b, Ax)$$

for all other vectors x in $\mathbf{R}^n$.

Recall that $\text{dist}(v, w) = \|v - w\|$ is the distance between the vectors v and w. The term "least squares" comes from the fact that $\text{dist}(b, Ax) = \|b - A\widehat{x}\|$ is the square root of the sum of the squares of the entries of the vector $b - A\widehat{x}$. So a least-squares solution minimizes the sum of the squares of the differences between the entries of $A\widehat{x}$ and b. In other words, a least-squares solution solves the equation $Ax = b$ as closely as possible, in the sense that the sum of the squares of the difference $b - Ax$ is minimized.

Least Squares: Picture Suppose that the equation $Ax = b$ is inconsistent. Recall from this note in Section 2.3 that the column space of A is the set of all other vectors c such that $Ax = c$ is consistent. In other words, $\text{Col}(A)$ is the set of all vectors of the form Ax. Hence, the closest vector of the form Ax to b is the orthogonal projection of b onto $\text{Col}(A)$. This is denoted $b_{\text{Col}(A)}$, following this notation in Section 6.3.



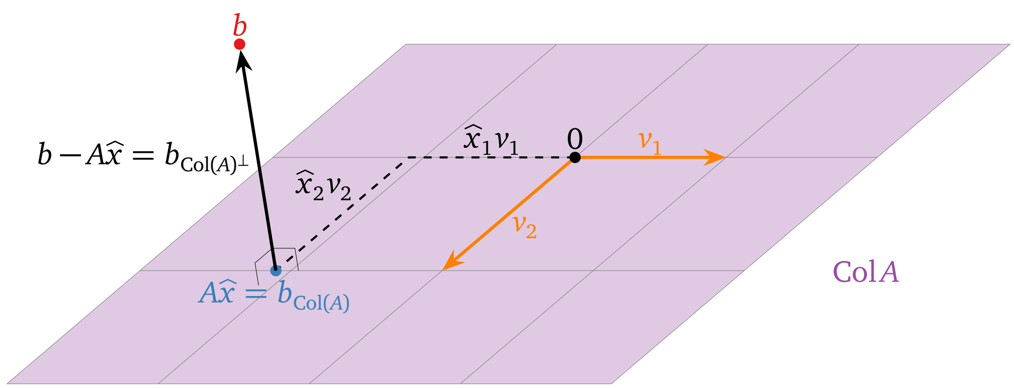
A least-squares solution of $Ax = b$ is a solution $\widehat{x}$ of the consistent equation $Ax = b_{\text{Col}(A)}$

Note. If $Ax = b$ is consistent, then $b_{\text{Col}(A)} = b$, so that a least-squares solution is the same as a usual solution.

Where is $\widehat{x}$ in this picture? If $v_1, v_2, \ldots, v_n$ are the columns of A, then

$$A\widehat{x} = A\begin{pmatrix} \widehat{x}_1 \\ \widehat{x}_2 \\ \vdots \\ \widehat{x}_n \end{pmatrix} = \widehat{x}_1 v_1 + \widehat{x}_2 v_2 + \cdots + \widehat{x}_n v_n.$$

Hence the entries of $\widehat{x}$ are the "coordinates" of $b_{\mathrm{Col}(A)}$ with respect to the spanning set $\{v_1, v_2, \ldots, v_m\}$ of $\mathrm{Col}(A)$. (They are honest $\mathcal{B}$-coordinates if the columns of A are linearly independent.)



Use this link to view the online demo

The violet plane is $\mathrm{Col}(A)$. *The closest that Ax can get to b is the closest vector on* $\mathrm{Col}(A)$ *to b, which is the orthogonal projection $b_{\mathrm{Col}(A)}$ (in blue). The vectors v_1, v_2 are the columns of A, and the coefficients of $\widehat{x}$ are the lengths of the green lines. Click and drag b to move it.*

We learned to solve this kind of orthogonal projection problem in Section 6.3.

Theorem. *Let A be an $m \times n$ matrix and let b be a vector in $\mathbf{R}^m$. The least-squares solutions of $Ax = b$ are the solutions of the matrix equation*

$$A^T A x = A^T b$$

Proof. By this theorem in Section 6.3, if $\widehat{x}$ is a solution of the matrix equation $A^T A x = A^T b$, then $A\widehat{x}$ is equal to $b_{\mathrm{Col}(A)}$. We argued above that a least-squares solution of $Ax = b$ is a solution of $Ax = b_{\mathrm{Col}(A)}$. □

In particular, finding a least-squares solution means solving a consistent system of linear equations. We can translate the above theorem into a recipe:

Recipe 1: Compute a least-squares solution. Let A be an $m \times n$ matrix and let b be a vector in $\mathbf{R}^n$. Here is a method for computing a least-squares solution of $Ax = b$:

1. Compute the matrix $A^T A$ and the vector $A^T b$.
2. Form the augmented matrix for the matrix equation $A^T Ax = A^T b$, and row reduce.
3. This equation is always consistent, and any solution $\widehat{x}$ is a least-squares solution.

To reiterate: once you have found a least-squares solution $\widehat{x}$ of $Ax = b$, then $b_{\text{Col}(A)}$ is equal to $A\widehat{x}$.

Example. Find the least-squares solutions of $Ax = b$ where:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}.$$

What quantity is being minimized?

Solution. We have

$$A^T A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix}$$

and

$$A^T b = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}.$$

We form an augmented matrix and row reduce:

$$\left(\begin{array}{cc|c} 5 & 3 & 0 \\ 3 & 3 & 6 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 5 \end{array}\right).$$

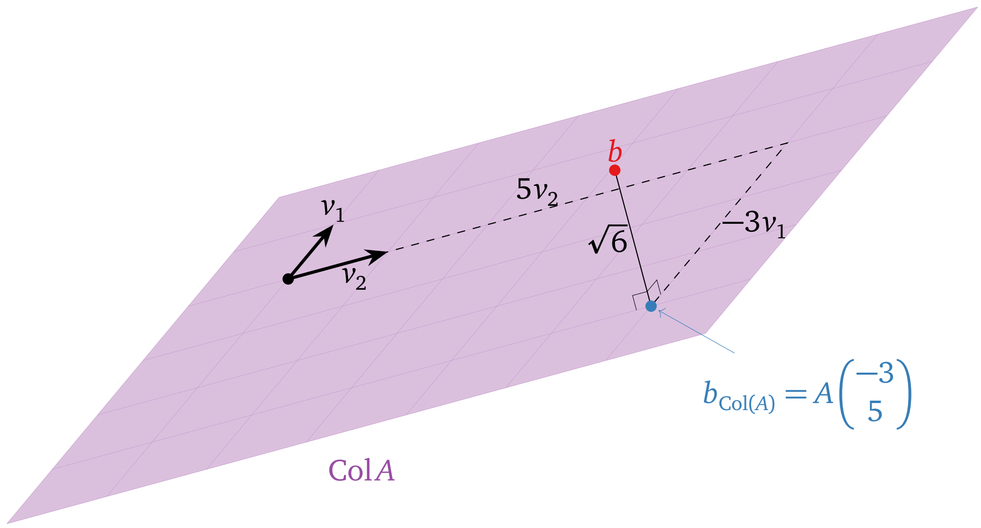
Therefore, the only least-squares solution is $\widehat{x} = \binom{-3}{5}$.

This solution minimizes the distance from $A\widehat{x}$ to b, i.e., the sum of the squares of the entries of $b - A\widehat{x} = b - b_{\text{Col}(A)} = b_{\text{Col}(A)^\perp}$. In this case, we have

$$\|b - A\widehat{x}\| = \left\| \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 5 \\ 2 \\ -1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \right\| = \sqrt{1^2 + (-2)^2 + 1^2} = \sqrt{6}.$$

Therefore, $b_{\mathrm{Col}(A)} = A\widehat{x}$ is $\sqrt{6}$ units from b.

In the following picture, v_1, v_2 are the columns of A:



Use this link to view the online demo

The violet plane is Col(A). *The closest that Ax can get to b is the closest vector on* Col(A) *to b, which is the orthogonal projection $b_{\mathrm{Col}(A)}$ (in blue). The vectors v_1, v_2 are the columns of A, and the coefficients of $\widehat{x}$ are the $\mathcal{B}$-coordinates of $b_{\mathrm{Col}(A)}$, where $\mathcal{B} = \{v_1, v_2\}$.*

Example. Find the least-squares solutions of $Ax = b$ where:

$$A = \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} \qquad b = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Solution. We have

$$A^T A = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$$

and

$$A^T b = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

We form an augmented matrix and row reduce:

$$\left(\begin{array}{cc|c} 5 & -1 & 2 \\ -1 & 5 & -2 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & -1/3 \end{array}\right).$$

Therefore, the only least-squares solution is $\widehat{x} = \frac{1}{3}\binom{1}{-1}$.

Use this link to view the online demo

The red plane is Col(A). *The closest that Ax can get to b is the closest vector on* Col(A) *to b, which is the orthogonal projection $b_{\mathrm{Col}(A)}$ (in blue). The vectors v_1, v_2 are the columns of A, and the coefficients of $\widehat{x}$ are the $\mathcal{B}$-coordinates of $b_{\mathrm{Col}(A)}$, where $\mathcal{B} = \{v_1, v_2\}$.*

The reader may have noticed that we have been careful to say "the least-squares solutions" in the plural, and "a least-squares solution" using the indefinite article. This is because a least-squares solution need not be unique: indeed, if the columns of A are linearly dependent, then $Ax = b_{\mathrm{Col}(A)}$ has infinitely many solutions. The following theorem, which gives equivalent criteria for uniqueness, is an analogue of this corollary in Section 6.3.

Theorem. *Let A be an $m \times n$ matrix and let b be a vector in $\mathbf{R}^m$. The following are equivalent:*

1. *$Ax = b$ has a unique least-squares solution.*
2. *The columns of A are linearly independent.*
3. *$A^T A$ is invertible.*

In this case, the least-squares solution is

$$\widehat{x} = (A^T A)^{-1} A^T b.$$

Proof. The set of least-squares solutions of $Ax = b$ is the solution set of the consistent equation $A^T Ax = A^T b$, which is a translate of the solution set of the homogeneous equation $A^T Ax = 0$. Since $A^T A$ is a square matrix, the equivalence of 1 and 3 follows from the invertible matrix theorem in Section 5.1. The set of least squares-solutions is also the solution set of the consistent equation $Ax = b_{\mathrm{Col}(A)}$, which has a unique solution if and only if the columns of A are linearly independent by this important note in Section 2.5. □

Example (Infinitely many least-squares solutions)**.** Find the least-squares solutions of $Ax = b$ where:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 2 & -3 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}.$$

Solution. We have

$$A^T A = \begin{pmatrix} 3 & 3 & -3 \\ 3 & 5 & -7 \\ -3 & -7 & 11 \end{pmatrix} \qquad A^T b = \begin{pmatrix} 6 \\ 0 \\ 6 \end{pmatrix}.$$

We form an augmented matrix and row reduce:

$$\left(\begin{array}{ccc|c} 3 & 3 & -3 & 6 \\ 3 & 5 & -7 & 0 \\ -3 & -7 & 11 & 6 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 5 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

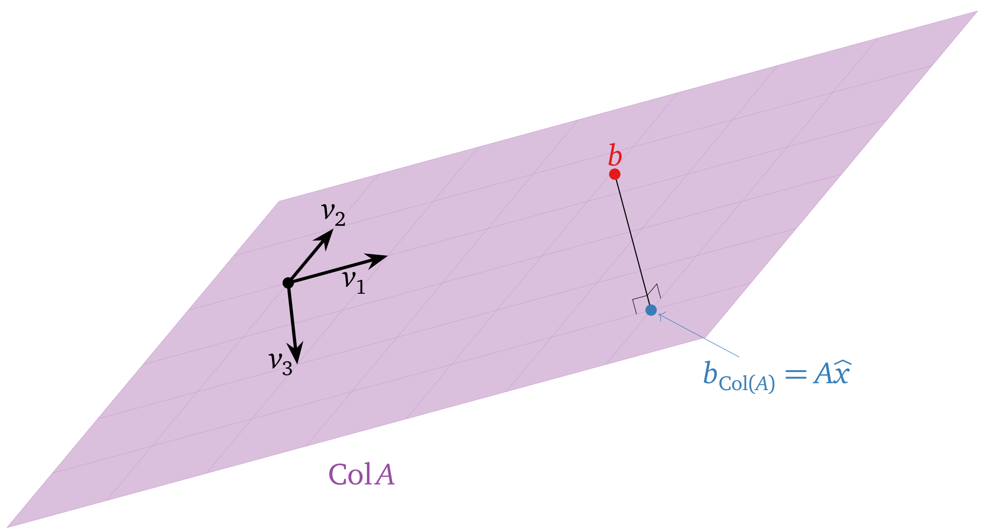
The free variable is x_3, so the solution set is

$$\begin{cases} x_1 = -x_3 + 5 \\ x_2 = 2x_3 - 3 \\ x_3 = x_3 \end{cases} \xrightarrow[\text{vector form}]{\text{parametric}} \quad \widehat{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 5 \\ -3 \\ 0 \end{pmatrix}.$$

For example, taking $x_3 = 0$ and $x_3 = 1$ gives the least-squares solutions

$$\widehat{x} = \begin{pmatrix} 5 \\ -3 \\ 0 \end{pmatrix} \quad \text{and} \quad \widehat{x} = \begin{pmatrix} 4 \\ -1 \\ 1 \end{pmatrix}.$$

Geometrically, we see that the columns v_1, v_2, v_3 of A are coplanar:



Therefore, there are many ways of writing $b_{\text{Col}(A)}$ as a linear combination of v_1, v_2, v_3.

Use this link to view the online demo

The three columns of A are coplanar, so there are many least-squares solutions. (The demo picks one solution when you move b.)

As usual, calculations involving projections become easier in the presence of an orthogonal set. Indeed, if A is an $m\times n$ matrix with *orthogonal* columns $u_1, u_2, \ldots, u_m$, then we can use the projection formula in Section 6.4 to write

$$b_{\text{Col}(A)} = \frac{b\cdot u_1}{u_1\cdot u_1}u_1 + \frac{b\cdot u_2}{u_2\cdot u_2}u_2 + \cdots + \frac{b\cdot u_m}{u_m\cdot u_m}u_m = A\begin{pmatrix}(b\cdot u_1)/(u_1\cdot u_1)\\(b\cdot u_2)/(u_2\cdot u_2)\\\vdots\\(b\cdot u_m)/(u_m\cdot u_m)\end{pmatrix}.$$

Note that the least-squares solution is unique in this case, since an orthogonal set is linearly independent.

Recipe 2: Compute a least-squares solution. Let A be an $m\times n$ matrix with *orthogonal* columns $u_1, u_2, \ldots, u_m$, and let b be a vector in $\mathbf{R}^n$. Then the least-squares solution of $Ax = b$ is the vector

$$\widehat{x} = \left(\frac{b\cdot u_1}{u_1\cdot u_1}, \frac{b\cdot u_2}{u_2\cdot u_2}, \ldots, \frac{b\cdot u_m}{u_m\cdot u_m}\right).$$

This formula is particularly useful in the sciences, as matrices with orthogonal columns often arise in nature.

Example. Find the least-squares solution of $Ax = b$ where:

$$A = \begin{pmatrix}1&0&1\\0&1&1\\-1&0&1\\0&-1&1\end{pmatrix} \qquad b = \begin{pmatrix}0\\1\\3\\4\end{pmatrix}.$$

Solution. Let u_1, u_2, u_3 be the columns of A. These form an orthogonal set, so

$$\widehat{x} = \left(\frac{b\cdot u_1}{u_1\cdot u_1}, \frac{b\cdot u_2}{u_2\cdot u_2}, \frac{b\cdot u_3}{u_3\cdot u_3}\right) = \left(\frac{-3}{2}, \frac{-3}{2}, \frac{8}{4}\right) = \left(-\frac{3}{2}, -\frac{3}{2}, 2\right).$$

Compare this example in Section 6.4.

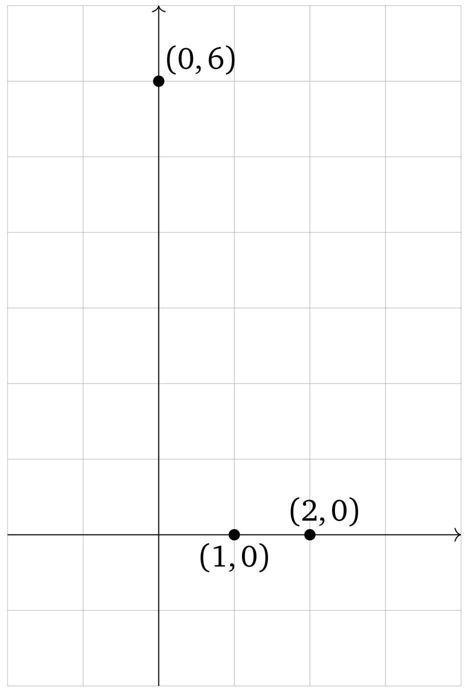
6.5.2 Best-Fit Problems

In this subsection we give an application of the method of least squares to data modeling. We begin with a basic example.

Example (Best-fit line)**.** Suppose that we have measured three data points

$$(0,6), \quad (1,0), \quad (2,0),$$

and that our model for these data asserts that the points should lie on a line. Of course, these three points do not actually lie on a single line, but this could be due to errors in our measurement. How do we predict which line they are supposed to lie on?



The general equation for a (non-vertical) line is

$$y = Mx + B.$$

If our three data points were to lie on this line, then the following equations would be satisfied:

$$\begin{aligned} 6 &= M \cdot 0 + B \\ 0 &= M \cdot 1 + B \\ 0 &= M \cdot 2 + B. \end{aligned} \tag{6.5.1}$$

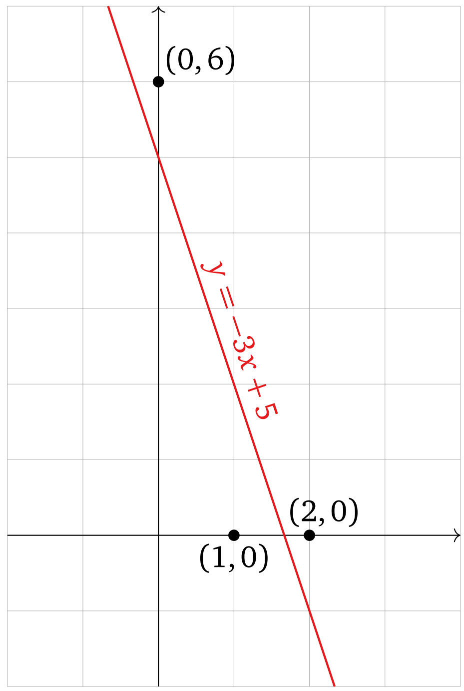
In order to find the best-fit line, we try to solve the above equations in the unknowns M and B. As the three points do not actually lie on a line, there is no actual solution, so instead we compute a least-squares solution.

Putting our linear equations into matrix form, we are trying to solve $Ax = b$ for

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \qquad x = \begin{pmatrix} M \\ B \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}.$$

We solved this least-squares problem in this example: the only least-squares solution to $Ax = b$ is $\widehat{x} = \binom{M}{B} = \binom{-3}{5}$, so the best-fit line is

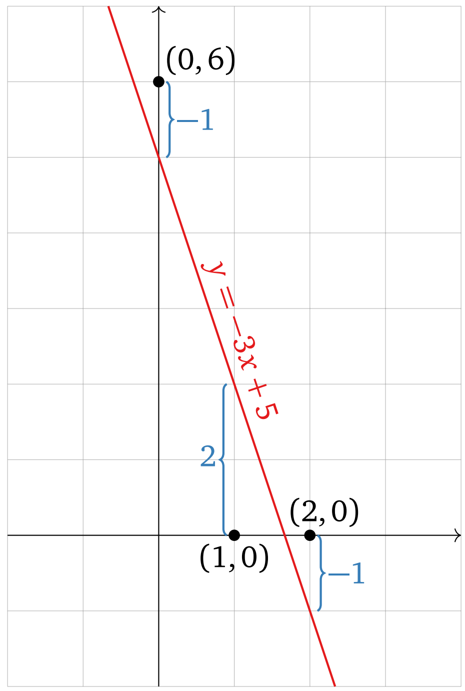
$$y = -3x + 5.$$



What exactly is the line $y = f(x) = -3x + 5$ minimizing? The least-squares solution $\widehat{x}$ minimizes the sum of the squares of the entries of the vector $b - A\widehat{x}$. The vector b is the left-hand side of (6.5.1), and

$$A\begin{pmatrix}-3\\5\end{pmatrix} = \begin{pmatrix}-3(0)+5\\-3(1)+5\\-3(2)+5\end{pmatrix} = \begin{pmatrix}f(0)\\f(1)\\f(2)\end{pmatrix}.$$

In other words, $A\widehat{x}$ is the vector whose entries are the y-coordinates of the graph of the line at the values of x we specified in our data points, and b is the vector whose entries are the y-coordinates of those data points. The difference $b - A\widehat{x}$ is the vertical distance of the graph from the data points:



$$b - A\widehat{x} = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix} - A\begin{pmatrix} -3 \\ 5 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}$$

The best-fit line minimizes the sum of the squares of these vertical distances.

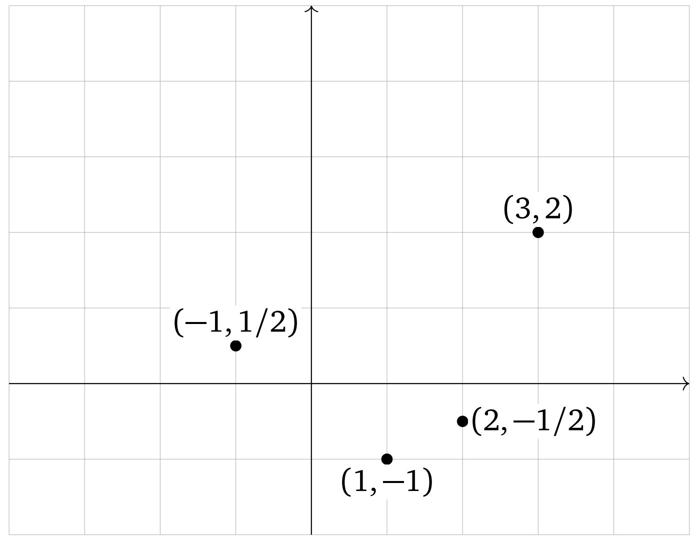
Interactive: Best-fit line.

Use this link to view the online demo

The best-fit line minimizes the sum of the squares of the vertical distances (violet). Click and drag the points to see how the best-fit line changes.

Example (Best-fit parabola)**.** Find the parabola that best approximates the data points

$$(-1,\ 1/2), \quad (1,\ -1), \quad (2,\ -1/2), \quad (3,\ 2).$$



What quantity is being minimized?

Solution. The general equation for a parabola is

$$y = Bx^2 + Cx + D.$$

If the four points were to lie on this parabola, then the following equations would be satisfied:

$$\begin{aligned} \frac{1}{2} &= B(-1)^2 + C(-1) + D \\ -1 &= B(1)^2 + C(1) + D \\ -\frac{1}{2} &= B(2)^2 + C(2) + D \\ 2 &= B(3)^2 + C(3) + D. \end{aligned} \tag{6.5.2}$$

We treat this as a system of equations in the unknowns B, C, D. In matrix form, we can write this as $Ax = b$ for

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{pmatrix} \qquad x = \begin{pmatrix} B \\ C \\ D \end{pmatrix} \qquad b = \begin{pmatrix} 1/2 \\ -1 \\ -1/2 \\ 2 \end{pmatrix}.$$

We find a least-squares solution by multiplying both sides by the transpose:

$$A^T A = \begin{pmatrix} 99 & 35 & 15 \\ 35 & 15 & 5 \\ 15 & 5 & 4 \end{pmatrix} \qquad A^T b = \begin{pmatrix} 31/2 \\ 7/2 \\ 1 \end{pmatrix},$$

then forming an augmented matrix and row reducing:

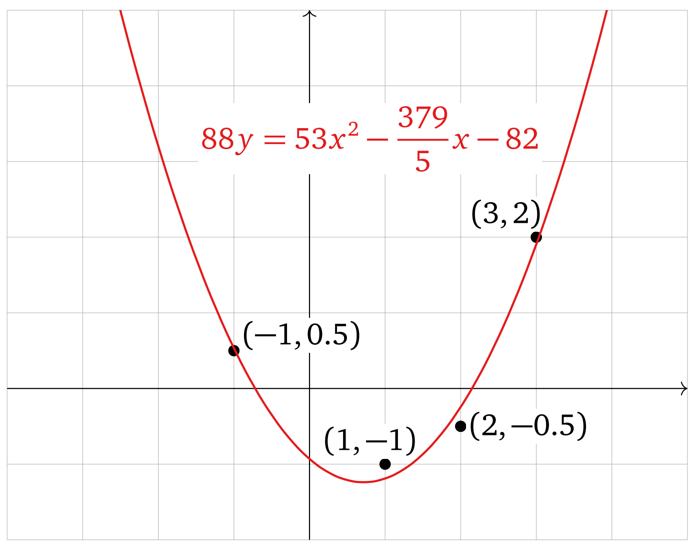
$$\left(\begin{array}{ccc|c} 99 & 35 & 15 & 31/2 \\ 35 & 15 & 5 & 7/2 \\ 15 & 5 & 4 & 1 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 53/88 \\ 0 & 1 & 0 & -379/440 \\ 0 & 0 & 1 & -41/44 \end{array}\right) \implies \widehat{x} = \begin{pmatrix} 53/88 \\ -379/440 \\ -41/44 \end{pmatrix}.$$

The best-fit parabola is

$$y = \frac{53}{88}x^2 - \frac{379}{440}x - \frac{41}{44}.$$

Multiplying through by 88, we can write this as

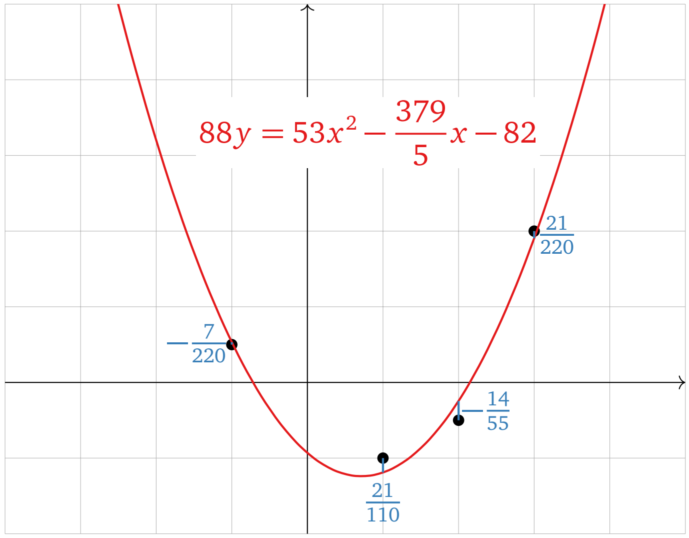
$$88y = 53x^2 - \frac{379}{5}x - 82.$$



Now we consider what exactly the parabola $y = f(x)$ is minimizing. The least-squares solution $\widehat{x}$ minimizes the sum of the squares of the entries of the vector $b - A\widehat{x}$. The vector b is the left-hand side of (6.5.2), and

$$A\widehat{x} = \begin{pmatrix} \frac{53}{88}(-1)^2 - \frac{379}{440}(-1) - \frac{41}{44} \\ \frac{53}{88}(1)^2 - \frac{379}{440}(1) - \frac{41}{44} \\ \frac{53}{88}(2)^2 - \frac{379}{440}(2) - \frac{41}{44} \\ \frac{53}{88}(3)^2 - \frac{379}{440}(3) - \frac{41}{44} \end{pmatrix} = \begin{pmatrix} f(-1) \\ f(1) \\ f(2) \\ f(3) \end{pmatrix}.$$

In other words, $A\widehat{x}$ is the vector whose entries are the y-coordinates of the graph of the parabola at the values of x we specified in our data points, and b is the vector whose entries are the y-coordinates of those data points. The difference $b - A\widehat{x}$ is the vertical distance of the graph from the data points:



$$b - A\widehat{x} = \begin{pmatrix} 1/2 \\ -1 \\ -1/2 \\ 2 \end{pmatrix} - A\begin{pmatrix} 53/88 \\ -379/440 \\ -41/44 \end{pmatrix} = \begin{pmatrix} -7/220 \\ 21/110 \\ -14/55 \\ 21/220 \end{pmatrix}$$

The best-fit parabola minimizes the sum of the squares of these vertical distances.

Use this link to view the online demo

The best-fit parabola minimizes the sum of the squares of the vertical distances (violet). Click and drag the points to see how the best-fit parabola changes.

Example (Best-fit linear function)**.** Find the linear function $f(x,y)$ that best approximates the following data:

x	y	$f(x,y)$
1	0	0
0	1	1
−1	0	3
0	−1	4

What quantity is being minimized?

Solution. The general equation for a linear function in two variables is

$$f(x,y) = Bx + Cy + D.$$

We want to solve the following system of equations in the unknowns B, C, D:

$$\begin{aligned} B(1) + \quad C(0) + D &= 0 \\ B(0) + \quad C(1) + D &= 1 \\ B(-1) + \quad C(0) + D &= 3 \\ B(0) + C(-1) + D &= 4. \end{aligned} \tag{6.5.3}$$

In matrix form, we can write this as $Ax = b$ for

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix} \qquad x = \begin{pmatrix} B \\ C \\ D \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 4 \end{pmatrix}.$$

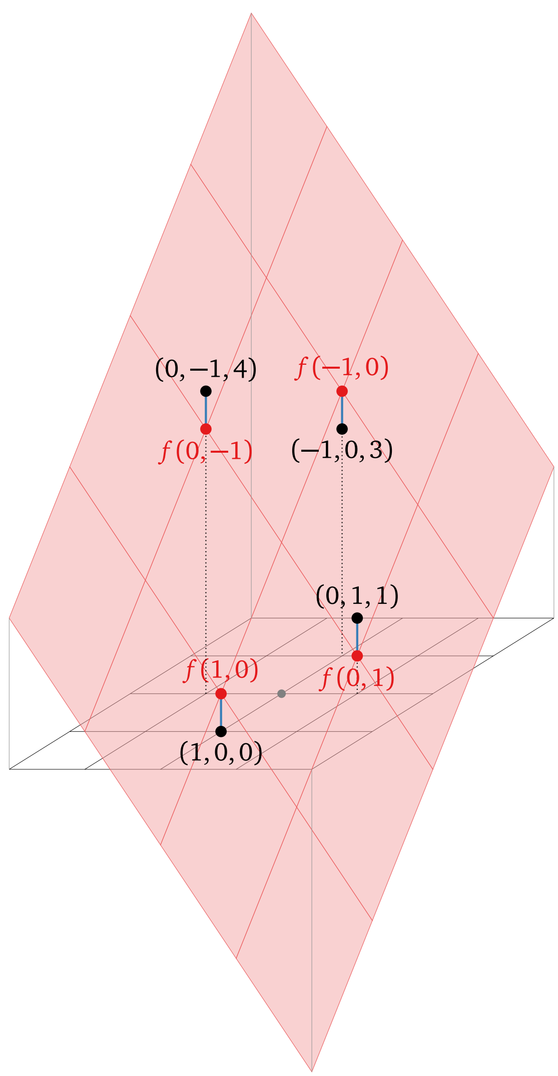
We observe that the columns u_1, u_2, u_3 of A are *orthogonal*, so we can use recipe 2:

$$\widehat{x} = \left(\frac{b \cdot u_1}{u_1 \cdot u_1}, \frac{b \cdot u_2}{u_2 \cdot u_2}, \frac{b \cdot u_3}{u_3 \cdot u_3}\right) = \left(\frac{-3}{2}, \frac{-3}{2}, \frac{8}{4}\right) = \left(-\frac{3}{2}, -\frac{3}{2}, 2\right).$$

Therefore, the best-fit linear equation is

$$f(x,y) = -\frac{3}{2}x - \frac{3}{2}y + 2.$$

Here is a picture of the graph of $f(x,y)$:



Now we consider what quantity is being minimized by the function $f(x,y)$. The least-squares solution $\widehat{x}$ minimizes the sum of the squares of the entries of the vector $b - A\widehat{x}$. The vector b is the right-hand side of (6.5.3), and

$$A\widehat{x} = \begin{pmatrix} -\frac{3}{2}(1) - \frac{3}{2}(0) + 2 \\ -\frac{3}{2}(0) - \frac{3}{2}(1) + 2 \\ -\frac{3}{2}(-1) - \frac{3}{2}(0) + 2 \\ -\frac{3}{2}(0) - \frac{3}{2}(-1) + 2 \end{pmatrix} = \begin{pmatrix} f(1,0) \\ f(0,1) \\ f(-1,0) \\ f(0,-1) \end{pmatrix}.$$

In other words, $A\widehat{x}$ is the vector whose entries are the values of f evaluated on the points (x,y) we specified in our data table, and b is the vector whose entries are the desired values of f evaluated at those points. The difference $b - A\widehat{x}$ is the vertical distance of the graph from the data points, as indicated in the above picture. The best-fit linear function minimizes the sum of these vertical distances.

Use this link to view the online demo

The best-fit linear function minimizes the sum of the squares of the vertical distances (violet). Click and drag the points to see how the best-fit linear function changes.

All of the above examples have the following form: some number of data points (x,y) are specified, and we want to find a function

$$y = B_1 g_1(x) + B_2 g_2(x) + \cdots + B_m g_m(x)$$

that best approximates these points, where $g_1, g_2, \ldots, g_m$ are fixed functions of x. Indeed, in the best-fit line example we had $g_1(x) = x$ and $g_2(x) = 1$; in the best-fit parabola example we had $g_1(x) = x^2$, $g_2(x) = x$, and $g_3(x) = 1$; and in the best-fit linear function example we had $g_1(x_1,x_2) = x_1$, $g_2(x_1,x_2) = x_2$, and $g_3(x_1,x_2) = 1$ (in this example we take x to be a vector with two entries). We evaluate the above equation on the given data points to obtain a system of linear equations in the unknowns $B_1, B_2, \ldots, B_m$—once we evaluate the g_i, they just become numbers, so it does not matter what they are—and we find the least-squares solution. The resulting best-fit function minimizes the sum of the squares of the vertical distances from the graph of $y = f(x)$ to our original data points.

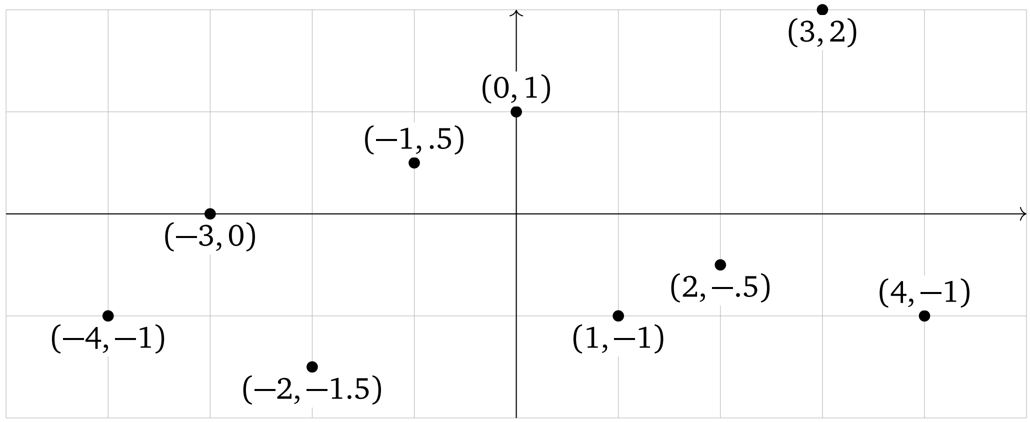
To emphasize that the nature of the functions g_i really is irrelevant, consider the following example.

Example (Best-fit trigonometric function)**.** What is the best-fit function of the form

$$y = B + C\cos(x) + D\sin(x) + E\cos(2x) + F\sin(2x) + G\cos(3x) + H\sin(3x)$$

passing through the points

$$\begin{pmatrix} -4 \\ -1 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ -1.5 \end{pmatrix}, \begin{pmatrix} -1 \\ .5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -.5 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 4 \\ -1 \end{pmatrix}?$$



Solution. We want to solve the system of equations

$$\begin{aligned}
-1 &= B + C\cos(-4) + D\sin(-4) + E\cos(-8) + F\sin(-8) + G\cos(-12) + H\sin(-12)\\
0 &= B + C\cos(-3) + D\sin(-3) + E\cos(-6) + F\sin(-6) + G\cos(-9) + H\sin(-9)\\
-1.5 &= B + C\cos(-2) + D\sin(-2) + E\cos(-4) + F\sin(-4) + G\cos(-6) + H\sin(-6)\\
0.5 &= B + C\cos(-1) + D\sin(-1) + E\cos(-2) + F\sin(-2) + G\cos(-3) + H\sin(-3)\\
1 &= B + C\cos(0) + D\sin(0) + E\cos(0) + F\sin(0) + G\cos(0) + H\sin(0)\\
-1 &= B + C\cos(1) + D\sin(1) + E\cos(2) + F\sin(2) + G\cos(3) + H\sin(3)\\
-0.5 &= B + C\cos(2) + D\sin(2) + E\cos(4) + F\sin(4) + G\cos(6) + H\sin(6)\\
2 &= B + C\cos(3) + D\sin(3) + E\cos(6) + F\sin(6) + G\cos(9) + H\sin(9)\\
-1 &= B + C\cos(4) + D\sin(4) + E\cos(8) + F\sin(8) + G\cos(12) + H\sin(12).
\end{aligned}$$

All of the terms in these equations are *numbers*, except for the unknowns B, C, D, E, F, G, H:

$$\begin{aligned}
-1 &= B - 0.6536C + 0.7568D - 0.1455E - 0.9894F + 0.8439G + 0.5366H\\
0 &= B - 0.9900C - 0.1411D + 0.9602E + 0.2794F - 0.9111G - 0.4121H\\
-1.5 &= B - 0.4161C - 0.9093D - 0.6536E + 0.7568F + 0.9602G + 0.2794H\\
0.5 &= B + 0.5403C - 0.8415D - 0.4161E - 0.9093F - 0.9900G - 0.1411H\\
1 &= B + C + E + G\\
-1 &= B + 0.5403C + 0.8415D - 0.4161E + 0.9093F - 0.9900G + 0.1411H\\
-0.5 &= B - 0.4161C + 0.9093D - 0.6536E - 0.7568F + 0.9602G - 0.2794H\\
2 &= B - 0.9900C + 0.1411D + 0.9602E - 0.2794F - 0.9111G + 0.4121H\\
-1 &= B - 0.6536C - 0.7568D - 0.1455E + 0.9894F + 0.8439G - 0.5366H.
\end{aligned}$$

Hence we want to solve the least-squares problem

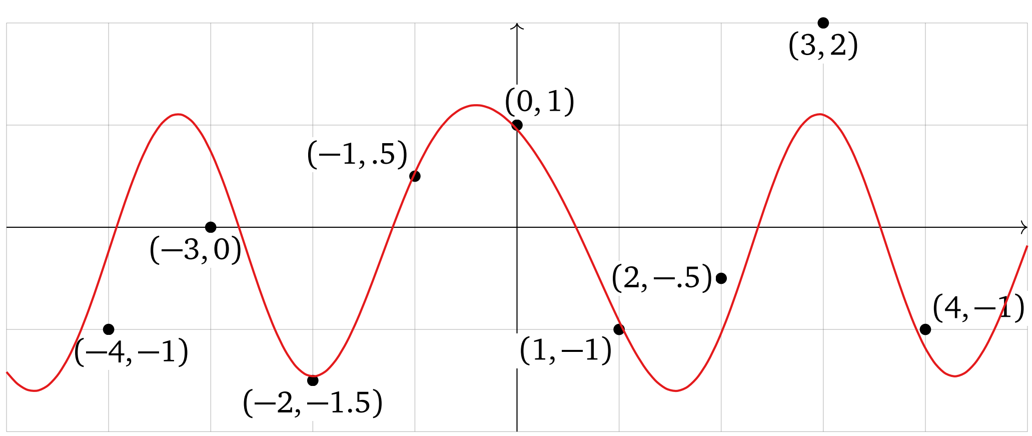
$$\begin{pmatrix} 1 & -0.6536 & 0.7568 & -0.1455 & -0.9894 & 0.8439 & 0.5366 \\ 1 & -0.9900 & -0.1411 & 0.9602 & 0.2794 & -0.9111 & -0.4121 \\ 1 & -0.4161 & -0.9093 & -0.6536 & 0.7568 & 0.9602 & 0.2794 \\ 1 & 0.5403 & -0.8415 & -0.4161 & -0.9093 & -0.9900 & -0.1411 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0.5403 & 0.8415 & -0.4161 & 0.9093 & -0.9900 & 0.1411 \\ 1 & -0.4161 & 0.9093 & -0.6536 & -0.7568 & 0.9602 & -0.2794 \\ 1 & -0.9900 & 0.1411 & 0.9602 & -0.2794 & -0.9111 & 0.4121 \\ 1 & -0.6536 & -0.7568 & -0.1455 & 0.9894 & 0.8439 & -0.5366 \end{pmatrix} \begin{pmatrix} B \\ C \\ D \\ E \\ F \\ G \\ H \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -1.5 \\ 0.5 \\ 1 \\ -1 \\ -0.5 \\ 2 \\ -1 \end{pmatrix}.$$

We find the least-squares solution with the aid of a computer:

$$\widehat{x} \approx \begin{pmatrix} -0.1435 \\ 0.2611 \\ -0.2337 \\ 1.116 \\ -0.5997 \\ -0.2767 \\ 0.1076 \end{pmatrix}.$$

Therefore, the best-fit function is

$$y \approx -0.1435 + 0.2611\cos(x) - 0.2337\sin(x) + 1.116\cos(2x) - 0.5997\sin(2x) \\ - 0.2767\cos(3x) + 0.1076\sin(3x).$$



$y \approx -0.14 + 0.26\cos(x) - 0.23\sin(x) + 1.11\cos(2x) - 0.60\sin(2x) - 0.28\cos(3x) + 0.11\sin(3x)$

As in the previous examples, the best-fit function minimizes the sum of the squares of the vertical distances from the graph of $y = f(x)$ to the data points.

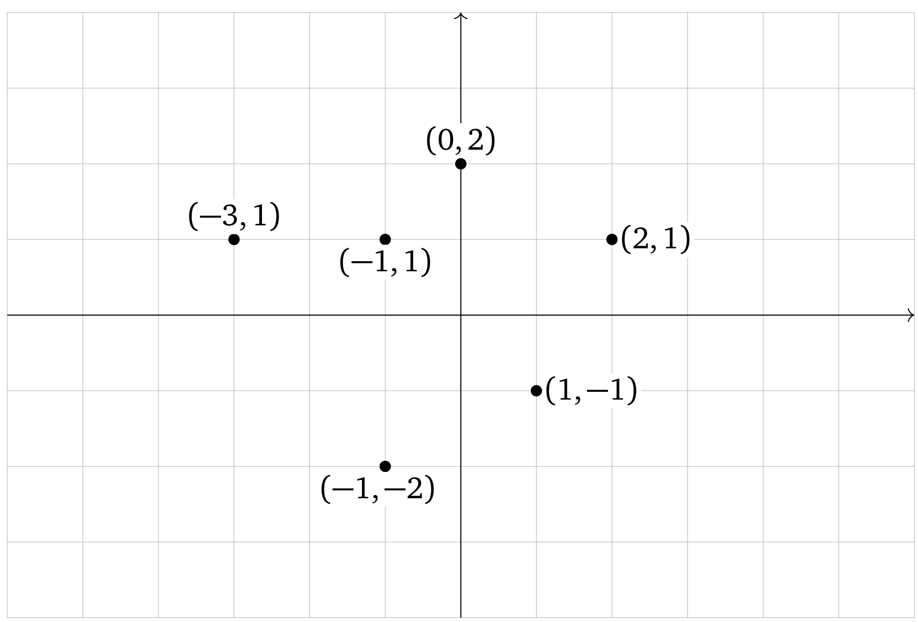
Use this link to view the online demo

The best-fit function minimizes the sum of the squares of the vertical distances (violet). Click and drag the points to see how the best-fit function changes.

The next example has a somewhat different flavor from the previous ones.

Example (Best-fit ellipse)**.** Find the best-fit ellipse through the points

$$(0,2),\ (2,1),\ (1,-1),\ (-1,-2),\ (-3,1),\ (-1,-1).$$



What quantity is being minimized?

Solution. The general equation for an ellipse (actually, for a nondegenerate conic section) is

$$x^2 + By^2 + Cxy + Dx + Ey + F = 0.$$

This is an *implicit equation*: the ellipse is the set of all solutions of the equation, just like the unit circle is the set of solutions of $x^2 + y^2 = 1$. To say that our data points lie on the ellipse means that the above equation is satisfied for the given values of x and y:

$$\begin{aligned}
(0)^2 + B(2)^2 + C(0)(2) + D(0) + E(2) + F &= 0\\
(2)^2 + B(1)^2 + C(2)(1) + D(2) + E(1) + F &= 0\\
(1)^2 + B(-1)^2 + C(1)(-1) + D(1) + E(-1) + F &= 0\\
(-1)^2 + B(-2)^2 + C(-1)(-2) + D(-1) + E(-2) + F &= 0\\
(-3)^2 + B(1)^2 + C(-3)(1) + D(-3) + E(1) + F &= 0\\
(-1)^2 + B(-1)^2 + C(-1)(-1) + D(-1) + E(-1) + F &= 0.
\end{aligned} \tag{6.5.4}$$

To put this in matrix form, we move the constant terms to the right-hand side of the equals sign; then we can write this as $Ax = b$ for

$$A = \begin{pmatrix} 4 & 0 & 0 & 2 & 1 \\ 1 & 2 & 2 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 4 & 2 & -1 & -2 & 1 \\ 1 & -3 & -3 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 \end{pmatrix} \qquad x = \begin{pmatrix} B \\ C \\ D \\ E \\ F \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ -4 \\ -1 \\ -1 \\ -9 \\ -1 \end{pmatrix}.$$

We compute

$$A^T A = \begin{pmatrix} 36 & 7 & -5 & 0 & 12 \\ 7 & 19 & 9 & -5 & 1 \\ -5 & 9 & 16 & 1 & -2 \\ 0 & -5 & 1 & 12 & 0 \\ 12 & 1 & -2 & 0 & 6 \end{pmatrix} \qquad A^T b = \begin{pmatrix} -19 \\ 17 \\ 20 \\ -9 \\ -16 \end{pmatrix}.$$

We form an augmented matrix and row reduce:

$$\left(\begin{array}{ccccc|c} 36 & 7 & -5 & 0 & 12 & -19 \\ 7 & 19 & 9 & -5 & 1 & 17 \\ -5 & 9 & 16 & 1 & -2 & 20 \\ 0 & -5 & 1 & 12 & 0 & -9 \\ 12 & 1 & -2 & 0 & 6 & -16 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 405/266 \\ 0 & 1 & 0 & 0 & 0 & -89/133 \\ 0 & 0 & 1 & 0 & 0 & 201/133 \\ 0 & 0 & 0 & 1 & 0 & -123/266 \\ 0 & 0 & 0 & 0 & 1 & -687/133 \end{array}\right).$$

The least-squares solution is

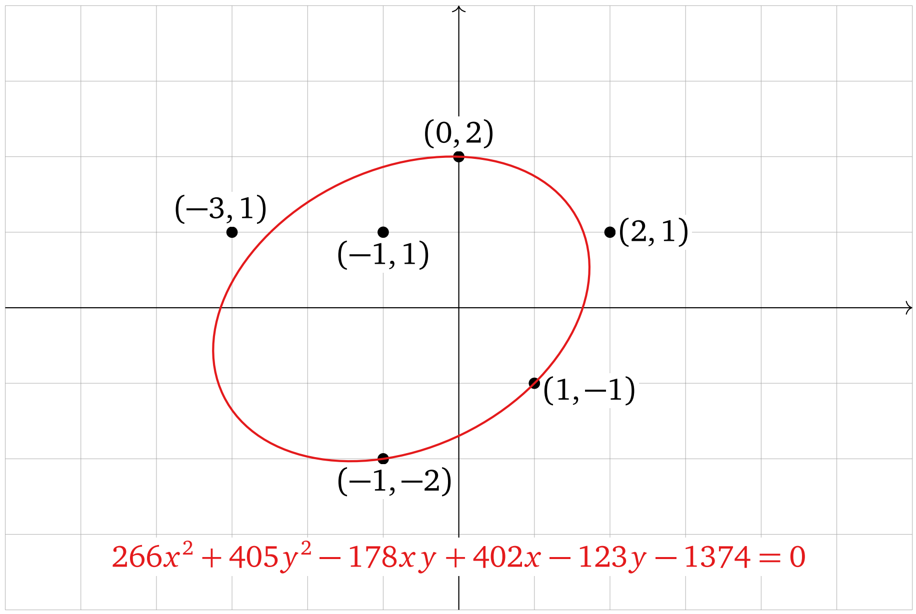
$$\widehat{x} = \begin{pmatrix} 405/266 \\ -89/133 \\ 201/133 \\ -123/266 \\ -687/133 \end{pmatrix},$$

so the best-fit ellipse is

$$x^2 + \frac{405}{266}y^2 - \frac{89}{133}xy + \frac{201}{133}x - \frac{123}{266}y - \frac{687}{133} = 0.$$

Multiplying through by 266, we can write this as

$$266x^2 + 405y^2 - 178xy + 402x - 123y - 1374 = 0.$$



Now we consider the question of what quantity is minimized by this ellipse. The least-squares solution $\widehat{x}$ minimizes the sum of the squares of the entries of the vector $b - A\widehat{x}$, or equivalently, of $A\widehat{x} - b$. The vector $-b$ contains the constant terms of the left-hand sides of (6.5.4), and

$$A\widehat{x} = \begin{pmatrix} \frac{405}{266}(2)^2 - \frac{89}{133}(0)(2) + \frac{201}{133}(0) - \frac{123}{266}(2) - \frac{687}{133} \\ \frac{405}{266}(1)^2 - \frac{89}{133}(2)(1) + \frac{201}{133}(2) - \frac{123}{266}(1) - \frac{687}{133} \\ \frac{405}{266}(-1)^2 - \frac{89}{133}(1)(-1) + \frac{201}{133}(1) - \frac{123}{266}(-1) - \frac{687}{133} \\ \frac{405}{266}(-2)^2 - \frac{89}{133}(-1)(-2) + \frac{201}{133}(-1) - \frac{123}{266}(-2) - \frac{687}{133} \\ \frac{405}{266}(1)^2 - \frac{89}{133}(-3)(1) + \frac{201}{133}(-3) - \frac{123}{266}(1) - \frac{687}{133} \\ \frac{405}{266}(-1)^2 - \frac{89}{133}(-1)(-1) + \frac{201}{133}(-1) - \frac{123}{266}(-1) - \frac{687}{133} \end{pmatrix}$$

contains the rest of the terms on the left-hand side of (6.5.4). Therefore, the entries of $A\widehat{x} - b$ are the quantities obtained by evaluating the function

$$f(x,y) = x^2 + \frac{405}{266}y^2 - \frac{89}{133}xy + \frac{201}{133}x - \frac{123}{266}y - \frac{687}{133}$$

on the given data points.

If our data points actually lay on the ellipse defined by $f(x,y) = 0$, then evaluating $f(x,y)$ on our data points would always yield zero, so $A\widehat{x} - b$ would be the zero vector. This is not the case; instead, $A\widehat{x} - b$ contains the *actual* values of $f(x,y)$ when evaluated on our data points. The quantity being minimized is the sum of the squares of these values:

$$\text{minimized} = f(0,2)^2 + f(2,1)^2 + f(1,-1)^2 + f(-1,-2)^2 + f(-3,1)^2 + f(-1,-1)^2.$$

One way to visualize this is as follows. We can put this best-fit problem into the framework of this example by asking to find an equation of the form

$$f(x,y) = x^2 + By^2 + Cxy + Dx + Ey + F$$

which best approximates the data table

x	y	$f(x,y)$
0	2	0
2	1	0
1	−1	0
−1	−2	0
−3	1	0
−1	−1	0.

The resulting function minimizes the sum of the squares of the vertical distances from these data points $(0,2,0)$, $(2,1,0)$, ..., which lie on the xy-plane, to the graph of $f(x,y)$.

Use this link to view the online demo

The best-fit ellipse minimizes the sum of the squares of the vertical distances (violet) from the points $(x,y,0)$ to the graph of $f(x,y)$ on the left. The ellipse itself is the zero set of $f(x,y)$, on the right. Click and drag the points on the right to see how the best-fit ellipse changes. Can you arrange the points so that the best-fit conic section is actually a hyperbola?

Note. Gauss invented the method of least squares to find a best-fit ellipse: he correctly predicted the (elliptical) orbit of the asteroid Ceres as it passed behind the sun in 1801.

Appendix A

Complex Numbers

In this Appendix we give a brief review of the arithmetic and basic properties of the complex numbers.

As motivation, notice that the rotation matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

has characteristic polynomial $f(\lambda) = \lambda^2 + 1$. A zero of this function is a square root of -1. If we want this polynomial to have a root, then we have to use a larger number system: we need to declare by *fiat* that there exists a square root of -1.

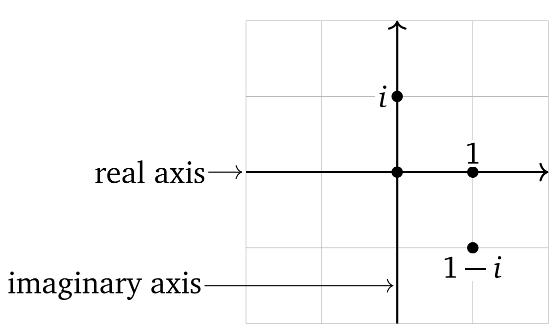
Definition.

1. The **imaginary number** i is defined to satisfy the equation $i^2 = -1$.
2. A **complex number** is a number of the form $a + bi$, where a, b are real numbers.

The set of all complex numbers is denoted **C**.

The real numbers are just the complex numbers of the form $a + 0i$, so that **R** is contained in **C**.

We can identify **C** with $\mathbf{R}^2$ by $a + bi \longleftrightarrow \binom{a}{b}$. So when we draw a picture of **C**, we draw the plane:



Arithmetic of Complex Numbers. We can perform all of the usual arithmetic operations on complex numbers: add, subtract, multiply, divide, absolute value. There is also an important new operation called complex conjugation.

- *Addition* is performed componentwise:

$$(a+bi)+(c+di)=(a+c)+(b+d)i.$$

- *Multiplication* is performed using distributivity and $i^2=-1$:

$$(a+bi)(c+di)=ac+adi+bci+bdi^2=(ac-bd)+(ad+bc)i.$$

- *Complex conjugation* replaces i with $-i$, and is denoted with a bar:

$$\overline{a+bi}=a-bi.$$

  The number $\overline{a+bi}$ is called the **complex conjugate** of $a+bi$. One checks that for any two complex numbers z,w, we have

$$\overline{z+w}=\overline{z}+\overline{w} \quad \text{and} \quad \overline{zw}=\overline{z}\cdot\overline{w}.$$

  Also, $(a+bi)(a-bi)=a^2+b^2$, so $z\overline{z}$ is a nonnegative *real* number for any complex number z.

- The *absolute value* of a complex number z is the real number $|z|=\sqrt{z\overline{z}}$:

$$|a+bi|=\sqrt{a^2+b^2}.$$

  One chacks that $|zw|=|z|\cdot|w|$.

- *Division* by a nonzero real number proceeds componentwise:

$$\frac{a+bi}{c}=\frac{a}{c}+\frac{b}{c}i.$$

- *Division* by a nonzero complex number requires multiplying the numerator and denominator by the complex conjugate of the denominator:

$$\frac{z}{w}=\frac{z\overline{w}}{w\overline{w}}=\frac{z\overline{w}}{|w|^2}.$$

  For example,

$$\frac{1+i}{1-i}=\frac{(1+i)^2}{1^2+(-1)^2}=\frac{1+2i+i^2}{2}=i.$$

- The *real* and *imaginary* parts of a complex number are

$$\mathrm{Re}(a+bi)=a \qquad \mathrm{Im}(a+bi)=b.$$

The point of introducing complex numbers is to find roots of polynomials. It turns out that introducing i is sufficent to find the roots of any polynomial.

Fundamental Theorem of Algebra. *Every polynomial of degree n has exactly n (real and) complex roots, counted with multiplicity.*

Equivalently, if $f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ is a polynomial of degree n, then f factors as

$$f(x) = (x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n)$$

for (not necessarily distinct) complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Degree-2 Polynomials. The quadratic formula gives the roots of a degree-2 polynomial, real or complex:

$$f(x) = x^2 + bx + c \implies x = \frac{-b \pm \sqrt{b^2 - 4c}}{2}.$$

For example, if $f(x) = x^2 - \sqrt{2}x + 1$, then

$$x = \frac{\sqrt{2} \pm \sqrt{-2}}{2} = \frac{\sqrt{2}}{2}(1 \pm i) = \frac{1 \pm i}{\sqrt{2}}.$$

Note that if b, c are real numbers, then the two roots are complex conjugates.

A complex number z is real if and only if $z = \overline{z}$. This leads to the following observation.

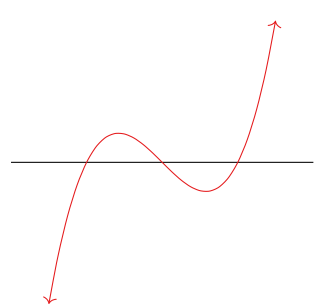
If f is a polynomial with real coefficients, and if λ is a complex root of f, then so is $\overline{\lambda}$:

$$\begin{aligned} 0 = \overline{f(\lambda)} &= \overline{\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0} \\ &= \overline{\lambda}^n + a_{n-1}\overline{\lambda}^{n-1} + \cdots + a_1\overline{\lambda} + a_0 = f\big(\overline{\lambda}\big). \end{aligned}$$

Therefore, complex roots of real polynomials come in *conjugate pairs*.

Degree-3 Polynomials. A real cubic polynomial has either three real roots, or one real root and a conjugate pair of complex roots.

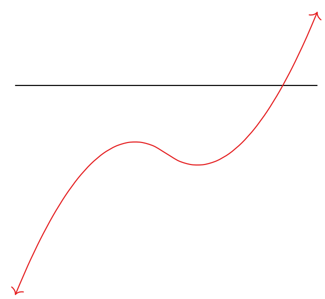
For example, $f(x) = x^3 - x = x(x-1)(x+1)$ has three real roots; its graph looks like this:



On the other hand, the polynomial

$$g(x) = x^3 - 5x^2 + x - 5 = (x-5)(x^2+1) = (x-5)(x+i)(x-i)$$

has one real root at 5 and a conjugate pair of complex roots $\pm i$. Its graph looks like this:



Appendix B

Notation

The following table defines the notation used in this book. Page numbers or references refer to the first appearance of each symbol.

Symbol	Description	Page
0	The number zero	3
$\mathbf{R}$	The real numbers	3
$\mathbf{R}^n$	Real n-space	3
R_i	Row i of a matrix	13
$\binom{1}{2}$	A vector	31
0	The zero vector	31
$\text{Span}\{v_1, v_2, \ldots, v_k\}$	Span of vectors	40
$\{x \mid \text{condition}\}$	Set builder notation	41
$m \times n$ matrix	Size of a matrix	44
$\text{Col}(A)$	Column space	88
$\text{Nul}(A)$	Null space	88
$\dim V$	Dimension of a subspace	93
$\text{rank}(A)$	The rank of a matrix	110
$\text{nullity}(A)$	The nullity of a matrix	110
$T: \mathbf{R}^n \to \mathbf{R}^m$	transformation with domain $\mathbf{R}^n$ and codomain $\mathbf{R}^m$	120
$\text{Id}_{\mathbf{R}^n}$	Identity transformation	122
$e_1, e_2, \ldots$	Standard coordinate vectors	147
I_n	$n \times n$ identity matrix	148
a_{ij}	The i,j entry of a matrix	157
0	The zero transformation	166
0	The zero matrix	167
A^{-1}	Inverse of a matrix	169
T^{-1}	Inverse of a transformation	176
$\det(A)$	The determinant of a matrix	188
A^T	Transpose of a matrix	199
A_{ij}	Minor of a matrix	207
C_{ij}	Cofactor of a matrix	207

(Continued on next page)

Symbol	**Description**	**Page**
$\mathrm{adj}(A)$	Adjugate matrix	218
$\mathrm{vol}(P)$	Volume of a region	224
$\mathrm{vol}(A)$	Volume of the parallelepiped of a matrix	225
$T(S)$	The image of a region under a transformation	232
$\mathrm{Tr}(A)$	Trace of a matrix	256
$\mathrm{Re}(v)$	Real part of a complex vector	311
$\mathrm{Im}(v)$	Imaginary part of a complex vector	311
$x \cdot y$	Dot product of two vectors	341
$x \perp y$	x is orthogonal to y	345
$W^\perp$	Orthogonal complement of a subspace	348
$\mathrm{Row}(A)$	Row space of a matrix	354
x_W	Orthogonal projection of x onto W	358
$x_{W^\perp}$	Orthogonal part of x with respect to W	358
$\mathbf{C}$	The complex numbers	409
$\overline{z}$	Complex conjugate	410
$\mathrm{Re}(z)$	Real part of a complex number	410
$\mathrm{Im}(z)$	Imaginary part of a complex number	410

Appendix C

Hints and Solutions to Selected Exercises

Appendix D

GNU Free Documentation License

0. PREAMBLE The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

The "publisher" means any person or entity that distributes copies of the Document to the public.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties — for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled “History” in the various original documents, forming one section Entitled “History”; likewise combine any sections Entitled “Acknowledgements”, and any sections Entitled “Dedications”. You must delete all sections Entitled “Endorsements”.

6. COLLECTIONS OF DOCUMENTS You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an “aggregate” if the copyright resulting from the compilation is not used to limit the legal rights of the compilation’s users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document’s Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled “Acknowledgements”, “Dedications”, or “History”, the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION You may not copy, modify, sublicense, or distribute the Document except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense, or distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, receipt of a copy of some or all of the same material does not give you any rights to use it.

10. FUTURE REVISIONS OF THIS LICENSE The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See `http://www.gnu.org/copyleft/`.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License “or any later version” applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. RELICENSING "Massive Multiauthor Collaboration Site" (or "MMC Site") means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A "Massive Multiauthor Collaboration" (or "MMC") contained in the site means any set of copyrightable works thus published on the MMC site.

"CC-BY-SA" means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

"Incorporate" means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is "eligible for relicensing" if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.

ADDENDUM: How to use this License for your documents To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

```
Copyright (C)  YEAR  YOUR NAME.
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.3
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts.
A copy of the license is included in the section entitled "GNU
Free Documentation License".
```

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with… Texts." line with this:

```
with the Invariant Sections being LIST THEIR TITLES, with the
Front-Cover Texts being LIST, and with the Back-Cover Texts being LIST.
```

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.

Index

This book was authored in MathBook XML, with modifications by Joe Rabinoff.