Good Basis, Great Results

A course guide to Linear Algebra

Benjamin Kennedy
bkennedy@gettysburg.edu

Keir Lockridge
klockrid@gettysburg.edu

Department of Mathematics
Gettysburg College

v1.1 December 6, 2025

Contents

Preface

To students

Welcome to our course guide for *linear algebra*! Linear algebra is a beautiful and useful subject that crops up in most major subfields of mathematics, including geometry, algebra, analysis, statistics, and mathematical modeling. Across the sciences, it provides theoretical insights and sharp tools for problem solving. Linear algebra is not just a basket of computational algorithms for solving problems; it's a framework for thinking about them. We hope that by reading this book you will come to appreciate the unique perspective that linear algebra offers. Along the way, you'll meet a number of applications, including some to population dynamics, directed graphs, the Fibonacci sequence, fractal geometry, and data compression.

The material in this course is challenging, and you will have to work hard to master it. But it is worth learning, and we want you to succeed. You can do it—and we are here to help!

At the start of each chapter, we will list the key concepts covered and summarize the chapter's content. This list and summary will be good to review after you've read it. After you have worked through a chapter, you should be able to fluently discuss the key concepts listed; if not, then you have more work to do! Don't worry if you struggle with the reading at first. The skills involved can be learned, but they need to be cultivated deliberately. Reading mathematics is a methodical process that requires a great deal of attention and engagement. You'll need to read this book with pencil and paper in hand and a willingness to work through the some of the details. Having to re-read the material—and seek help from others to understand it—is totally normal. We recommend that you make flash cards for key theorems and definitions. You cannot be an effective problem solver if you have not internalized the core concepts in this book; it won't do to have to look up basic definitions and results every time you might need to use them.

You should do all the **Reading Questions** (RQs) as part of any reading assignment, but you do not need to do the **Exercises** unless you are specifically asked to do so (they may come up in class, for homework, or for extra practice). RQs will be marked in the margin with a circled "RQ". Attempting to answer these questions is an absolute must; you need to grapple with the ideas in the reading and prove to yourself that you're getting something out of it. *You won't be adequately prepared for class if you don't do the reading and try to work though some details on your own first.* RQs will not always be easy, but you must trust that you are capable of solving some problems without having someone tell you exactly what to do step-by-step. Our expectation is not that you always come to class with complete answers to every reading question, but you should be able to present evidence that you made a genuine effort to answer these questions and tried to do something specific.

Proofs of most theorems can be skipped on a first-pass reading, where your main concern is to follow the narrative thread, understand the definitions, and work though the examples. You should definitely come back to the proofs—especially when there's a reading question about one. If you are committed to doing well in this course (and achieving the learning goals), then you must be fully committed to moving beyond learning the rote execution of algorithmic computations. Please understand that this course is very much about *why* things are true and *why* they work. It can be demanding because it may be the first time you've really been asked to understand and produce mathematical arguments. *But this is something you can do!*

Is it time to get started? Go read Chapter 1!

To instructors

We hope we've written a book that students can read. We cover neither more nor less than what we actually expect students to study and learn during our one semester introduction to linear algebra course. Calculus II is a prerequisite—more for the mathematical experience than for the material covered there. Our pace is roughly one chapter per week, though Chapter 7 can be done in half a week and Chapters 9 and 11 might each take about a week and a half. In a 15 week semester, this leaves time for loose ends, exams, and orientation to computational tools (such as SageMath/Python in Cocalc).

Here is how the second author uses this book in his course.

- Students must read the book before class and be prepared to answer the Reading Questions. Some of the Exercises are used for group work and in-class activities.

- Routine homework exercises are assigned weekly through an online homework system (such as WeBWorK or MyOpenMath), though these systems certainly contain sophisticated problems (and some of those are assigned, too).
- Students complete three projects in CoCalc using SageMath/Python (on sports ranking, fractal geometry, and Google's PageRank algorithm).
- Most weeks (though not when projects are due), students are required to write up and turn in 1-2 problems (in LaTeX or in Markdown with LaTeX). The goal is to work on writing and reasoning; some of the book's exercises are appropriate for this purpose. These problems are highly scaffolded, especially in the beginning of the semester.
- There is an exam after Chapter 4 and Chapter 9, plus a final exam.

Feedback and other comments are welcome; please just send us an email!

Chapter 1 Vectors in Euclidean space

Key concepts

- Vectors in $\mathbf{R}^n$ and their geometric/algebraic properties
- Three situations where linear algebra comes up naturally
- Linear combinations of vectors
- The span of a set of vectors
- Matrices
- The matrix-vector product and linearity
- The matrix-vector equation $A\mathbf{x} = \mathbf{b}$

Summary. After we introduce vectors and discuss their properties, we will show you three real examples where linear algebra arises: population dynamics, economics, and geometry. Each example connects to an important idea in linear algebra: linear combinations, systems of linear equations, and special functions called linear transformations. We will learn about all of these concepts in great detail.

A central object of study moving forward is the matrix-vector equation $A\mathbf{x} = \mathbf{b}$, where A is an $n \times k$ matrix, $\mathbf{x}$ is a k-vector, and $\mathbf{b}$ is an n-vector. This equation may have no solutions, exactly one solution, or infinitely many solutions. It has a solution if and only if the vector $\mathbf{b}$ is a linear combination of the columns of A (or, put another way, if $\mathbf{b}$ lies in the span of A's columns).

Chapter 1

The first thing we need to do is introduce the most basic objects in linear algebra: *vectors.* Then we'll be able to give three quick examples where linear algebra arises in population dynamics, economics, and geometry.

Vectors §1.1

What's a vector?

An n-**dimensional vector** (or just n-vector) is an ordered list of n real numbers, organized in a column and enclosed in brackets. For example,

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

is a 3-vector. When typographically convenient, we may instead use the notation $(-2, 4, 1)$. We'll write $\mathbf{R}^n$ for the set of all n-vectors:

What's $\mathbf{R}^n$?

$$\mathbf{R}^n = \left\{ (x_1, x_2, \ldots, x_n) \;\middle|\; x_1, \ldots, x_n \in \mathbf{R} \right\}.$$

The notation above is called **set-builder notation**. In set-builder notation, the objects of the set are described to the left of the vertical bar, constraints on the objects are described to the right of the vertical bar, and the whole thing is enclosed in curly braces. $\mathbf{R}$ is the set of real numbers, and the symbol $\in$ means "is in" or "belongs to." So here is how we read the above set-builder notation for $\mathbf{R}^n$ out loud: $\mathbf{R}^n$ *is the set of all vectors* $(x_1, x_2, \ldots, x_n)$*, where* $x_1, x_2, \ldots, x_n$ *are real numbers.*

The set $\mathbf{R}^n$ is often called n-**dimensional Euclidean space.** In particular, $\mathbf{R}^2$ can be thought of as the plane and $\mathbf{R}^3$ can be thought of as three-dimensional space.

When we want to name a vector, we will do so with a boldface letter like this:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

What's the zero vector?

(Be aware that the use of boldface letters to denote vectors is not a universal convention; be alert when consulting other sources.) The **zero vector** is the special vector $\mathbf{0} = (0, \ldots, 0)$, where $x_i = 0$ for all i.

What's a scalar?

We can multiply vectors by real numbers (called **scalars** in this context) by scaling their entries: for example,

How do we scale vectors?

$$3 \cdot \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \cdot (-2) \\ 3 \cdot 4 \\ 3 \cdot 1 \end{bmatrix} = \begin{bmatrix} -6 \\ 12 \\ 3 \end{bmatrix}.$$

Vectors of the same dimension may be added by adding corresponding entries: for example,

How do we add vectors?

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 7 \\ -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \\ 7 \end{bmatrix}.$$

Vector addition has a nice geometric interpretation. Consider the vectors

$$\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}.$$

You can interpret these vectors as ordinary points in the plane with coordinates $(1, 2)$ and $(-4, 2)$. You can also interpret each vector as an arrow whose tail is at the origin and whose head is located at the corresponding point. One virtue of the latter interpretation is that it gives us a way to think about the sum $\mathbf{a} + \mathbf{b}$ geometrically: it is the fourth corner of the parallelogram determined by $\mathbf{a}$ and $\mathbf{b}$ (see Figure 1.1 below). We have

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

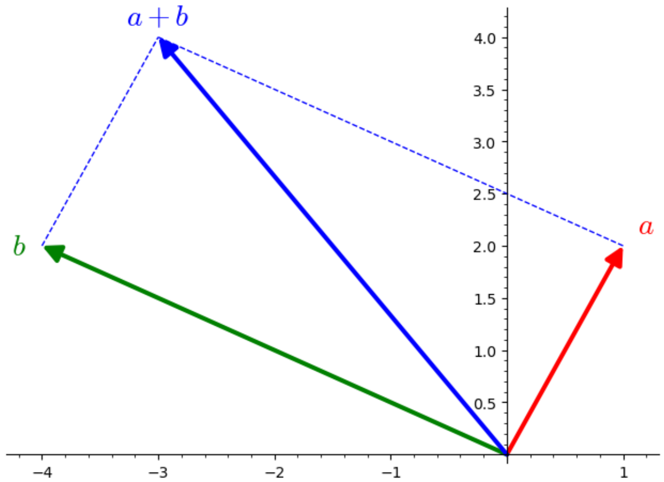


Figure 1.1: Adding two vectors

How do we interpret vector addition geometrically?

The next figure illustrates the effect of scaling the vector $\mathbf{a}$ to obtain $3\mathbf{a}$, $-\mathbf{a}$, and $-\frac{1}{2}\mathbf{a}$.

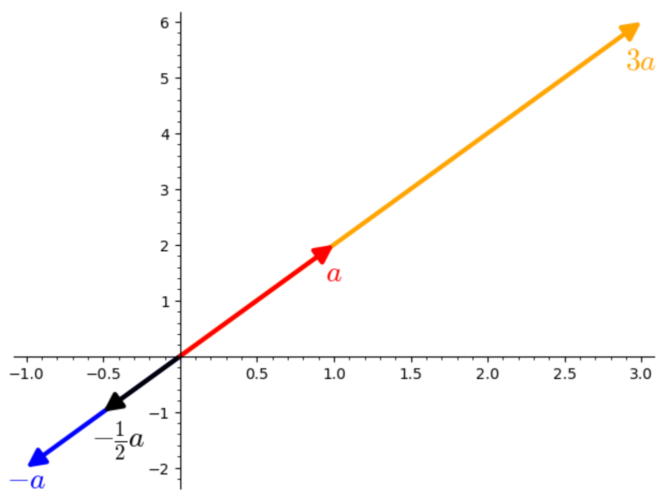


Figure 1.2: Scaling vectors

How do we interpret scalar multiplication geometrically?

Given two points $P, Q \in \mathbf{R}^n$, we can interpret the vector $P - Q$, sometimes written $\overrightarrow{QP}$, as an arrow that starts at Q and ends at P, as in Figure 1.3 below.

How can we interpret the vector $P - Q$ as the arrow from Q to P?

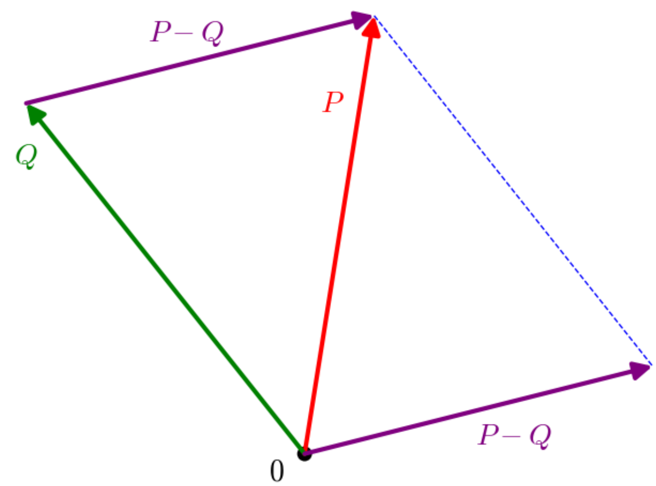


Figure 1.3: The vector from Q to P

To be clear, the vector $P - Q$, as a point, is the lower right corner of the parallelogram above; however, the top edge labeled $P - Q$ is parallel to this vector, has the same length, and points in the same direction. The meaning of this vector is perhaps best expressed by the equation

$$Q + \overrightarrow{QP} = P.$$

RQ **Reading Question 1A.** Viewing a vector as a point in $\mathbf{R}^n$, its length is defined to be the distance from this point to the origin $\mathbf{0}$. Viewing a vector as an arrow, its length is the length of the arrow. Find a formula for the length of a vector $(x_1, x_2) \in \mathbf{R}^2$ or $(x_1, x_2, x_3) \in \mathbf{R}^3$.

At some point, we will study a more general class of mathematical objects called vector spaces that includes $\mathbf{R}^n$. For now, we'll just summarize the relevant properties that make $\mathbf{R}^n$ a vector space. They are straightforward to verify if you write down coordinate versions of the vectors involved and then do the indicated operations.

ALGEBRAIC PROPERTIES OF $\mathbf{R}^n$

Theorem 1.1. *The following statements hold for all* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{R}^n$ *and all* $s, t \in \mathbf{R}$.

(1) *Vector addition is commutative:*

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}.$$

(2) *Vector addition is associative:*

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z}).$$

(3) *The zero vector is the additive identity:*

$$\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}.$$

(4) *Every vector* $\mathbf{x}$ *has additive inverse* $-\mathbf{x} = -1 \cdot \mathbf{x}$*:*

$$\mathbf{x} + (-\mathbf{x}) = -\mathbf{x} + \mathbf{x} = \mathbf{0}.$$

(5) *Scalar multiplication distributes over vector addition:*

$$s(\mathbf{x} + \mathbf{y}) = s\mathbf{x} + s\mathbf{y}.$$

(6) *Scalar multiplication distributes over real number addition:*

$$(s + t)\mathbf{x} = s\mathbf{x} + t\mathbf{x}.$$

(7) *The real number 1 acts as it should:*

$$1\mathbf{x} = \mathbf{x}.$$

(8) *Scalar multiplication is associative:*

$$s(t\mathbf{x}) = (st)\mathbf{x}.$$

Reading Question 1B. Pick three vectors $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$; pick two scalars s and t. Verify that the identities in Theorem 1.1 hold for your particular choices. Doing this won't prove that the theorem is true in general, but it will help you ensure that you understand what the theorem says. **RQ**

Exercise 1A. Carefully prove item Theorem 1.1 item (5).

Exercise 1B. Find a formula for the midpoint between two points in $\mathbf{R}^n$.

Exercise 1C. Vector addition and scalar multiplication satisfy the properties in Theorem 1.1 essentially because these properties hold in the ordinary real numbers, $\mathbf{R}$. But in $\mathbf{R}$, you can also multiply. Can you think of a multiplication operation for $\mathbf{R}^n$? Does it behave in a way that's surprising compared to multiplication for $\mathbf{R}$?

§1.2 Three motivating examples

Let's look at three situations where vectors naturally arise.

Example 1.2 (lionfish). Morris, Shertzer and Rice[1] consider a model for the growth of an invasive species of lionfish with a view toward population control. In their model, the female lionfish has three stages of development: larva, juvenile, and adult. Let L denote the current population of larvae, let J denote the current population of juveniles, and let A denote the current population of adults. Given these current population values, the model estimates the populations of larvae, juveniles, and adults (L', J', and A', respectively) one month later as follows:

$$\begin{aligned} L' &= & & 35315A \\ J' &= 0.00003L &+ 0.777J & \\ A' &= & 0.071J &+ 0.949A. \end{aligned}$$

This model says (for example) that for each adult this month, there will be 35315 larvae next month; for each larva this month, there will be 0.00003 juveniles next month; and so on.

We can use vectors to rewrite our system of equations as

$$\begin{bmatrix} L' \\ J' \\ A' \end{bmatrix} = L \begin{bmatrix} 0 \\ 0.00003 \\ 0 \end{bmatrix} + J \begin{bmatrix} 0 \\ 0.777 \\ 0.071 \end{bmatrix} + A \begin{bmatrix} 35315 \\ 0 \\ 0.949 \end{bmatrix}.$$

The right hand side of the equation above is called a **linear combination** of the three vectors shown; the scalars L, J, and A are called **weights**. Note that the entries in the vector with weight L are grouped conceptually in the sense that they tell you how the current population of larvae affects the subsequent populations of larvae, juveniles, and adults. A similar fact holds for the weights J and A.

[1] Morris, J.A., Shertzer, K.W. & Rice, J.A. A stage-based matrix population model of invasive lionfish with implications for control. *Biol Invasions* 13, 7–12 (2011).

RQ

Reading Question 1C. In Example 1.2 above, explain how to interpret the coefficients 0.777, 0.071 and 0.949 in the equation. Do juveniles take one month to mature, or longer? Why do the coefficients vary so much in size?

Exercise 1D. Suppose we have have two companies: Bender's Balloons and Fry's Falafels. If B and F are their stock prices in dollars per share on a given day, then the stock prices the following day, B' and F', are predicted by the rules $B' = 1.5F$ and $F' = .2B + .7F$. Write the vector (B', F') of future prices as a linear combination of vectors where the weights are B and F. If the initial stock prices are $B = \$2$ per share and $F = \$0.25$ per share, use a computer to predict the stock prices over a 25 day period. What do you notice?

Example 1.3 (robots and pies)**.** Consider a simple economy that consists of two productive sectors: the Robot sector (that produces robots) and the Pie sector (that produces pies). There is also a non-productive part of the economy that consumes robots (to play games with) and pies (for picnics). The Robot sector consumes inputs from both the Robot sector and the Pie sector: robots help make more robots and pie juice is an excellent mechanical lubricant. Similarly, the pie sector consumes inputs from both the Robot sector and the Pie sector: human workers eat pies for lunch, and the robots on the pie-making assembly line wear out and need to be replaced.

More concretely, consider the following table:

	inputs consumed per unit output	
	Robot Sector	**Pie Sector**
Robots	0.3	0.2
Pies	0.1	0.4

This table says, for example, that to produce 10 robots, the Robot sector consumes $10 \cdot 0.3 = 3$ robots and $10 \cdot 0.1 = 1$ pie. These assumptions form a special case of a *Leontief input-output model.*

Now, let d_1 and d_2 be the total number of robots and pies, respectively, consumed by the non-productive part of the economy, and let x_1 and x_2 denote the total number of robots and pies, respectively, produced by the entire economy. Then, according to the Leontief input-output model, we must have

$$x_1 = 0.3x_1 + 0.2x_2 + d_1$$
$$x_2 = 0.1x_1 + 0.4x_2 + d_2.$$

This system of equations may be rearranged into

$$\begin{aligned} 0.7x_1 - 0.2x_2 &= d_1 \\ -0.1x_1 + 0.6x_2 &= d_2. \end{aligned}$$

Finally, this may be expressed as a vector equation

$$x_1 \underbrace{\begin{bmatrix} 0.7 \\ -0.1 \end{bmatrix}}_{\mathbf{v}_1} + x_2 \underbrace{\begin{bmatrix} -0.2 \\ 0.6 \end{bmatrix}}_{\mathbf{v}_2} = \underbrace{\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}}_{\mathbf{d}}.$$

The vector $\mathbf{d}$ gives the outputs demanded by the non-productive sector of the economy, and it is a linear combination of the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ where the weights are the total amounts produced by the Robot and Pie sectors.

The equation above lets us figure out how much we need to produce to achieve any desired outputs for the non-productive sector. Suppose, for example, we want to produce $d_1 = 50$ robots and $d_2 = 102$ pies for the non-productive sector of the economy. Then, we can solve the above equation using substitution to find the total outputs x_1 and x_2.

We will soon learn a powerful method for solving vector equations like the one above that works well for any number of vectors, of any size.

RQ **Reading Question 1D.** Solve the vector equation at the end of Example 1.3 to find x_1 and x_2.

Example 1.4 (rotating the plane by $\pi/4$)**.** A simple task in computer graphics is to take the image on the screen and rotate it about a fixed point. For example, given a point $\mathbf{x} = (x_1, x_2)$ in the plane, suppose we'd like a formula for the coordinates of the point $R(x_1, x_2)$ obtained by rotating the point $\pi/4$ radians counter-clockwise about the origin.

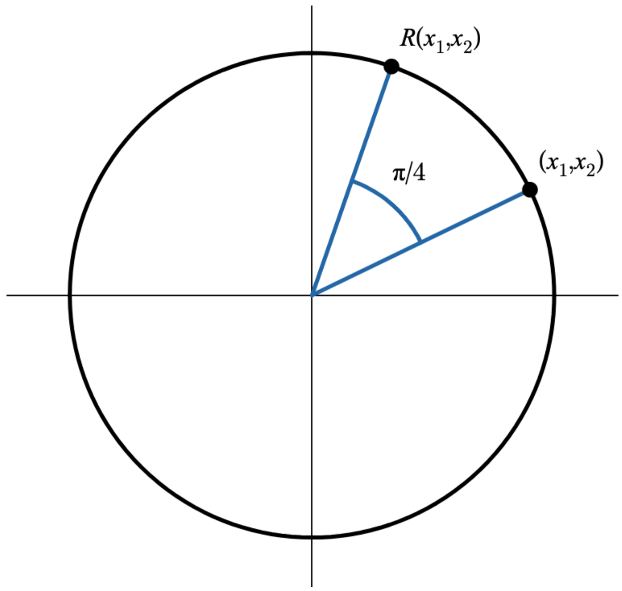


Figure 1.4: Rotating by $\pi/4$

Building on the ideas and notation in the previous example, first note that

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Now think carefully about how rotation about the origin works: if you take a vector, rotate it, and then scale it by some number, that's the same as first scaling the vector and then rotating it. Similarly, if you rotate a pair of vectors and then add the results, it's the same as first adding the vectors and then rotating the result. We will later see that this makes rotation an example of a **linear transformation**, and so the value of $R(x_1, x_2)$ is completely determined by $R(1, 0)$ and $R(0, 1)$ in the following way:

Rotation of the plane about the origin is a linear transformation from $\mathbf{R}^2$ to $\mathbf{R}^2$.

$$R\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = R\left(x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = x_1 R\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + x_2 R\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right).$$

Using trigonometry, you can check that $R(1, 0) = (1/\sqrt{2}, 1/\sqrt{2})$ and $R(0, 1) = (-1/\sqrt{2}, 1/\sqrt{2})$.

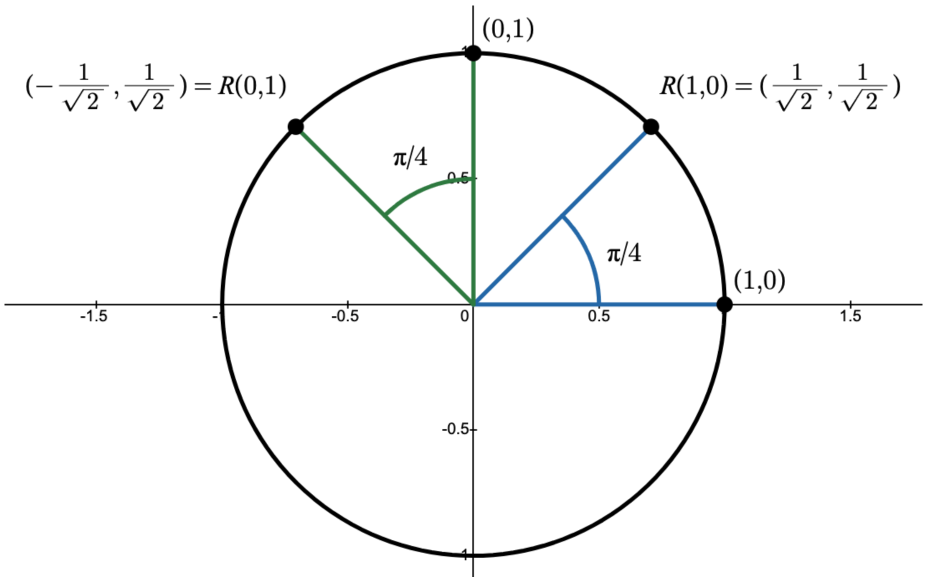


Figure 1.5: Rotating $(1, 0)$ and $(0, 1)$ by $\pi/4$

Thus,

$$R\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} + x_2 \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Our interesting geometric transformation takes (x_1, x_2) to a linear combination of two fixed vectors.

We'll record a general formula for rotation about the origin in Example 2.6.

RQ

Reading Question 1E. Find a formula similar to the one given in Example 1.4 for rotation of the plane about the origin counter-clockwise by $\pi/3$ radians.

Exercise 1E. Let's define a function that takes each input vector $(x, y) \in \mathbf{R}^2$ to an output vector in $\mathbf{R}^2$ according to the rule

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = x \begin{bmatrix} 1 \\ 2 \end{bmatrix} + y \begin{bmatrix} -2 \\ -4 \end{bmatrix}.$$

Pick five different vectors (x, y) and compute the output for each input you chose. Then, as best you can, describe the set of *all possible outputs* for this function.

Linear combinations and spans §1.3

Inspired by Example 1.3, let's consider another system of linear equations

$$\begin{aligned} 2\,x_1 - x_2 &= 1 \\ 3\,x_1 + x_2 &= 9 \end{aligned}.$$

We can write this system as a vector equation as follows:

$$x_1 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \end{bmatrix}. \tag{1.1}$$

Reading Question 1F. Rewrite the following system of equations as a vector equation: RQ

$$\begin{aligned} x_1 - x_2 + 4x_3 &= 0 \\ 2x_2 - x_3 &= -1 \\ -x_1 + 3x_2 + x_3 &= 3. \end{aligned}$$

Reading Question 1G. Rewrite the following vector equation as a system of equations: RQ

$$x_1 \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 7 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix}.$$

To see why we care about rewriting systems of equations as vector equations, look back at equation (1.1). A **solution** to this equation is a pair of real numbers x_1 and x_2 that makes the equation true. Writing our equation in vector form shows that the equation has a solution if and only if the vector $(1,\ 9)$ can be built from the vectors $(2,\ 3)$ and $(-1,\ 1)$ using vector addition and scalar multiplication. To help us talk about these ideas more carefully, we need a couple of definitions.

"A if and only if B" means that either both A and B are true or both A and B are false; they can't have opposite truth values.

What's a linear combination, and what's the span of a set of vectors?

Linear combinations and spans

Definition 1.5. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbf{R}^n$. For any real numbers $x_1, \ldots, x_k$ (called **weights**), the quantity

$$x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k$$

is called a **linear combination** of the vectors in S. The **span** of S is the set of *all* linear combinations of the vectors in S:

$$\operatorname{span} S = \{x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k \mid x_1, \ldots, x_k \in \mathbf{R}\}.$$

Study this vocabulary carefully: a linear combination is a single vector you can build from S, such as

$$3\mathbf{v}_1 + 0\mathbf{v}_2 - 2\mathbf{v}_3.$$

The *set* span S, by contrast, is the set of *all possible linear combinations of vectors in S*. Please also study this notation carefully: the objects $\mathbf{v}_1, \ldots, \mathbf{v}_k$ in Definition 1.5 are k *vectors* — they are *not* the entries of a single vector $\mathbf{v}$. Further, these vectors must all have the same number of entries (they must all lie in $\mathbf{R}^n$ for some fixed n). The objects $x_1, \ldots, x_k$ in Definition 1.5, on the other hand, are numbers. The importance of being alert to "object type" in linear algebra cannot be overstated.

Expressed in the language of Definition 1.5, the observation about the vector equation (1.1) that we made a moment ago is the following: *the equation has a solution if and only if*

$$\begin{bmatrix} 1 \\ 9 \end{bmatrix} \in \operatorname{span}\left\{\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}.$$

www.desmos.com /calculator/wbrt9ublfc

You can play with linear combinations using this Desmos illustration. Let $\mathbf{v} = (2, 3)$ and $\mathbf{w} = (-1, 1)$. In Figure 1.6 below, the point

$$P = \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \mathbf{v} - 2\mathbf{w}$$

is expressed as a linear combination of $\mathbf{v}$ and $\mathbf{w}$. We see that linear combinations give us a different way to describe locations in the plane: to get to P, you start at the origin and then move in the direction of $\mathbf{v}$ (up and to the right) once and then in the direction of $-\mathbf{w}$ (down and to the right) twice.

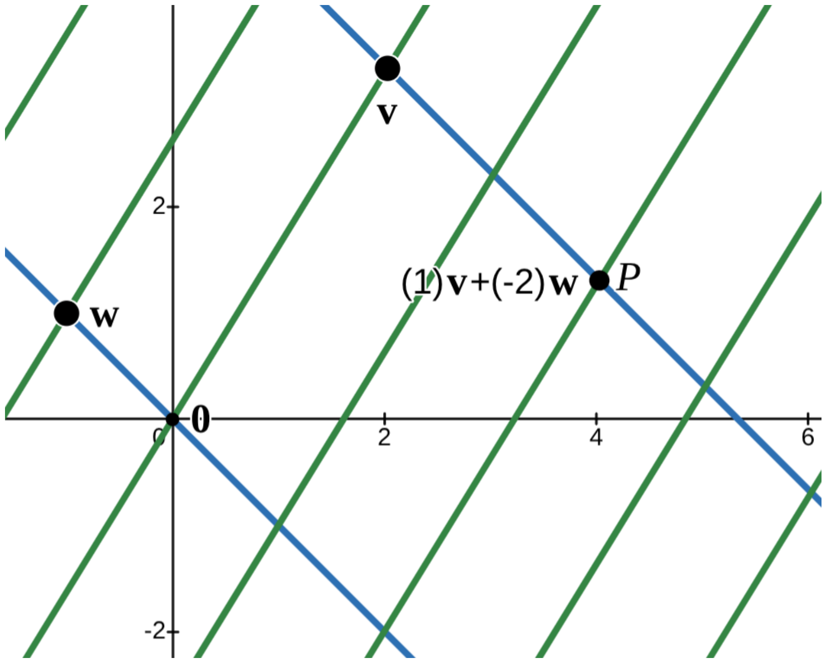


Figure 1.6: Linear combinations

Reading Question 1H. Show that each of the vectors

$$\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\-2\end{bmatrix}, \begin{bmatrix}2\\0\end{bmatrix}$$

lies in

$$\operatorname{span}\left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}\right\}.$$

[Hint: each vector is in the span if and only if a certain equation analogous to (1.1) has a solution.]

RQ

Exercise 1F. Show that the system of equations

$$\begin{aligned} x_1 + 2x_2 &= b_1 \\ x_1 - x_2 &= b_2 \end{aligned}$$

has a solution for all $b_1, b_2 \in \mathbf{R}$. Given this, what's the span of $\{(1,1), (2,-1)\}$?

Reading Question 1I. Find an example in each case or try to explain why it's not possible to do so.

RQ

1. A set S of two vectors with $\operatorname{span} S = \mathbf{R}^2$.
2. A set S of three vectors with $\operatorname{span} S = \mathbf{R}^2$.
3. A set S of one vector with $\operatorname{span} S = \mathbf{R}^2$.
4. A set S of one vector where $\operatorname{span} S$ is the line $y = 7x$.
5. A set S of two vectors where $\operatorname{span} S$ is the line $y = 7x$.
6. A set S of one vector where $\operatorname{span} S$ is the line $y = 7x + 1$.
7. A set S of vectors where $\operatorname{span} S = \{(x, y) \mid x \geq 0, y \geq 0\}$.

Exercise 1G. Determine the span of each set below and prove your answers. In each case, if there is a vector in $\mathbf{R}^2$ that is *not* in the span, then give an example of such a vector.

1. $\{\mathbf{a}\}$, for $\mathbf{a} = (3, 2)$
2. $\{\mathbf{b}\}$, for $\mathbf{b} = (0, 0)$
3. $\{\mathbf{a}, \mathbf{b}\}$
4. $\{(1, 0), (0, 1)\}$
5. $\{(5, 0), (1, 1)\}$
6. $\{(1, 3), (-2, -6)\}$

If S contains a nonzero vector, then $\operatorname{span} S$ is infinite.

Exercise 1H. Prove that if S is a set of vectors that contains a nonzero vector, then the span is infinite. Use this to find the one and only nonempty set of vectors in $\mathbf{R}^n$ whose span is finite.

Since we only have one vector in our set, we used t to denote the weight instead of x_1.

Example 1.6 (spans of one-vector sets)**.** Let $\mathbf{v}$ be any vector in $\mathbf{R}^n$. If $\mathbf{v} = \mathbf{0}$, then the span is

$$\operatorname{span}\{\mathbf{0}\} = \{t\mathbf{0} \mid t \in \mathbf{R}\} = \{\mathbf{0}\}.$$

If $\mathbf{v}$ is nonzero, then the set

$$\operatorname{span}\{\mathbf{v}\} = \{t\mathbf{v} \mid t \in \mathbf{R}\}$$

is a **line** in $\mathbf{R}^n$ through the origin with "direction vector" $\mathbf{v}$.

$\operatorname{span}\{\mathbf{v}, \mathbf{w}\}$ is $\{\mathbf{0}\}$, a line, or a plane.

Example 1.7 (spans of two-vector sets)**.** Consider a set of vectors $\{\mathbf{v}, \mathbf{w}\}$. If both vectors are zero, then $\operatorname{span} S = \{\mathbf{0}\}$. Next, suppose S contains a nonzero vector; let's say $\mathbf{w} \neq \mathbf{0}$. If $\mathbf{v}$ is a scalar multiple of $\mathbf{w}$, then $\mathbf{v} = c\mathbf{w}$ for some scalar c and we have

$$\begin{aligned} \operatorname{span} S &= \{s\mathbf{v} + t\mathbf{w} \mid s, t \in \mathbf{R}\} \\ &= \{s(c\mathbf{w}) + t\mathbf{w} \mid s, t \in \mathbf{R}\} \\ &= \{(sc + t)\mathbf{w} \mid s, t \in \mathbf{R}\} \\ &= \operatorname{span}\{\mathbf{w}\}. \end{aligned}$$

Make sure you understand the fourth equality!

In this case, span S is a line through the origin. If neither vector is a scalar multiple of the other (so they both must be nonzero!), then we call span S the **plane** through the origin determined by $\mathbf{v}$ and $\mathbf{w}$.

Exercise 1I. Show that the vectors $(0, 1, 1)$, $(0, -3, -3)$, and $(1, 0, 1)$ do not span all of $\mathbf{R}^3$ by finding a vector that is not in their span (and proving you're correct).

Exercise 1J. Suppose a vector $\mathbf{c}$ is **not** in the span of a set of vectors S. Is it possible that $t\mathbf{c}$ **is** in the span of S for some $t \neq 0$? Either give an example to show this is possible or explain why it's not possible. [To start, write down what it would mean, by definition of span, for $t\mathbf{c}$ to be in the span of $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.]

The strategy suggested here is called proof by contradiction.

Exercise 1K. Let S be a set of vectors. Prove that if $\mathbf{x} \in \operatorname{span} S$ and $t \in \mathbf{R}$, then $t\mathbf{x} \in \operatorname{span} S$. Prove that if $\mathbf{x}, \mathbf{y} \in \operatorname{span} S$, then $\mathbf{x} + \mathbf{y} \in \operatorname{span} S$.

Exercise 1L. Show that every vector $(a, b, c) \in \mathbf{R}^3$ is in the span of the set

$$\{(1, 2, -1), (0, 3, 4), (0, 0, -5)\}.$$

Matrices and the matrix-vector product §1.4

Look back at equation (1.1) in §1.3. The solution, if it exists, is a pair of real numbers, and their order matters in the sense that their roles are not interchangeable. In other words, a solution of this equation can be thought of as a 2-vector (x_1, x_2)! Similarly, a solution of a general vector equation

$$x_1\mathbf{v}_1 + \cdots + x_k\mathbf{v}_k = \mathbf{b}$$

can be thought of as a k-vector $\mathbf{x} = (x_1, \ldots, x_k)$. Our lives will be much easier in the long run if we can train ourselves to think of a solution of a vector equation as being a single thing (i.e. a vector) rather than as a list of things (i.e. the numbers $x_1, \ldots, x_k$). To facilitate this way of thinking, we need to define matrices and the matrix-vector product.

What's a matrix?

A **matrix** is just an $n \times k$ grid of numbers:

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,k} \end{bmatrix}.$$

This matrix has n rows and k columns. In situations where we do not name the entries of the matrix explicitly (as above), we can write $A_{i,j}$ for the entry of A that's in the ith row and the jth column. When we are referring to entries or locations in a matrix by row and column, we **always** name the row first: so, for example, the when we talk about the "$(3, 2)$ entry" of a matrix, we mean in the entry in the third row and the second column.

When working with computers, the first row/column index is usually 0.

Each column of an $n \times k$ matrix is an n-vector. If we give the columns of A the names $\mathbf{a}_1, \ldots, \mathbf{a}_k$, then we can also write this matrix as

$$A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_k].$$

Now, if

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix}$$

How's the matrix-vector product defined?

is a vector in $\mathbf{R}^k$, we can define the **matrix-vector product** by

$$A\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_k\mathbf{a}_k.$$

To form the product $A\mathbf{x}$, the dimension of $\mathbf{x}$ must equal the number of columns in A. The dimension of the vector $A\mathbf{x}$ is equal to the number of rows in A.

In other words, the vector $A\mathbf{x}$ is just the linear combination of the columns of A with weights $x_1, \ldots, x_k$. Observe that, in order for this definition to even make sense, *the number of columns in A must be equal to dimension of $\mathbf{x}$* so that there is exactly one weight per column. Further, since $A\mathbf{x}$ is a linear combination of the columns of A, which are vectors in $\mathbf{R}^n$, *the vector $A\mathbf{x}$ lies in $\mathbf{R}^n$.*

$$\underbrace{\overbrace{A}^{n\times k} \cdot \overbrace{\mathbf{x}}^{k-\text{vec.}}}_{n-\text{vec.}} \in \mathbf{R}^n$$

Example 1.8. We have

$$\begin{bmatrix} 1 & 2 & 0 & 3 \\ -1 & 1 & 5 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 3 \\ -1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 5 \end{bmatrix} + (-1) \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -1 \\ 15 \end{bmatrix}.$$

Reading Question 1J. Compute the matrix-vector product $A\mathbf{x}$ for

RQ

$$A = \begin{bmatrix} 1 & 2 & -3 \\ -1 & 0 & 2 \\ 0 & c & -4 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}.$$

Express your answer as a linear combination of three vectors (using the definition of the matrix-vector product) and as a single vector (simplify your answer).

Exercise 1M. Find a 3×2 matrix A such that

$$A \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_2 \\ x_1 \end{bmatrix}.$$

The next theorem tells us how to do arithmetic with our new product operation.

Linearity of the matrix-vector product

Theorem 1.9. *Let A be an $n \times k$ matrix. For all scalars s and t and all vectors $\mathbf{x}, \mathbf{y} \in \mathbf{R}^k$,*

$$A(s\mathbf{x} + t\mathbf{y}) = sA\mathbf{x} + tA\mathbf{y}.$$

Proof. Let $A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_k]$, $\mathbf{x} = (x_1, \ldots, x_k)$, and $\mathbf{y} = (y_1, \ldots, y_k)$. We compute:

$$\begin{aligned}
A(s\mathbf{x}+t\mathbf{y}) &= A(s(x_1,\ldots,x_k)+t(y_1,\ldots,y_k)) \\
&= A((sx_1+ty_1,\ldots,sx_k+ty_k)) \\
&= (sx_1+ty_1)\mathbf{a}_1+\cdots+(sx_k+ty_k)\mathbf{a}_k \\
&= (sx_1\mathbf{a}_1+\cdots+sx_k\mathbf{a}_k)+(ty_1\mathbf{a}_1+\cdots+ty_k\mathbf{a}_k) \\
&= s(x_1\mathbf{a}_1+\cdots+x_k\mathbf{a}_k)+t(y_1\mathbf{a}_1+\cdots+y_k\mathbf{a}_k) \\
&= sA\mathbf{x}+tA\mathbf{y}.
\end{aligned}$$

■

RQ

Reading Question 1K. In the proof of Theorem 1.9, track exactly where we used scalar multipication, vector addition, and the definition of the matrix-vector product. What does the theorem tell you when $s = t = 0$? What if $s = t = 1$? What if $t = 0$ and $\mathbf{y} = \mathbf{0}$?

RQ

Reading Question 1L. Let A be a 5×3 matrix, and take $\mathbf{y} \in \mathbf{R}^3$ and $\mathbf{z} \in \mathbf{R}^5$. Suppose $A\mathbf{y} = \mathbf{z}$. Does $A\mathbf{x} = 5\mathbf{z}$ have a solution?

Exercise 1N. In each case below, either find a matrix A (with specific real number entries) so that the equation is satisfied for all x, y and z or explain why it's not possible to do so.

① $A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x+y+1 \end{bmatrix}$

② $A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x^2 \\ y^2 \\ z^2 \end{bmatrix}$

③ $A\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y+z \\ x+z \end{bmatrix}.$

Exercise 1O. Let A be a 2×1066 matrix and suppose $\mathbf{c}$ and $\mathbf{d}$ are vectors in $\mathbf{R}^{1066}$ such that $A(2\mathbf{c}+\mathbf{d}) = (1,2)$ and $A(3\mathbf{d}) = (-3,9)$. Find $A\mathbf{c}$ and $A\mathbf{d}$.

Exercise 1P. Find two different 2×2 matrices A and B such that $A\mathbf{x} = B\mathbf{x}$ for all $\mathbf{x} \in \mathbf{R}^2$ or explain why this is an impossible task. [Hint: consider $\mathbf{x} = (1,0)$ and $\mathbf{x} = (0,1)$.]

Exercise 1Q. Find two different 2 by 2 matrices A and B such that the span of the columns of A is equal to the span of the columns of B or explain why this is an impossible task.

We can rewrite equation (1.1) as a matrix-vector equation; it is equivalent to

$$\underbrace{\begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 1 \\ 9 \end{bmatrix}}_{\mathbf{b}}.$$

Now, let's make a connection between spans and the general matrix-vector equation $A\mathbf{x} = \mathbf{b}$.

SPAN AND THE MATRIX-VECTOR EQUATION $A\mathbf{x} = \mathbf{b}$

Theorem 1.10. *Let A be an $n \times k$ matrix and take $\mathbf{b} \in \mathbf{R}^n$. The following statements are equivalent.*

① *The equation $A\mathbf{x} = \mathbf{b}$ has a solution.*

② *The vector $\mathbf{b}$ is in the span of the columns of A.*

This theorem says that, in any specific situation, either both statements are true or both statements are false.

Proof. To prove that the two statement are equivalent, we must prove that if the first statement is true, then the second statement is true, and if the second statement is true, then so is the first. Write $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_k]$ and let $S = \{\mathbf{a}_1, \dots, \mathbf{a}_k\}$.

First, assume that the equation $A\mathbf{x} = \mathbf{b}$ has a solution. Then, there is a vector $\mathbf{c} = (c_1, \dots, c_k)$ such that $A\mathbf{c} = \mathbf{b}$. By definition of the matrix-vector product, this means that

Notice that we start the argument by clearly stating what we are assuming is true.

$$c_1\mathbf{a}_1 + \cdots + c_k\mathbf{a}_k = \mathbf{b}.$$

This proves that $\mathbf{b}$ is a linear combination of the columns of A, and hence $\mathbf{b} \in \operatorname{span} S$ by definition of span. We have established that ① implies ②.

You should try to prove that ② implies ①. ■

Reading Question 1M. Prove that ② implies ① in Theorem 1.10. The first sentence of your proof should be "Assume $\mathbf{b}$ is in the span of the columns of A." Then, you should unravel what this means using the definition of span. Have you by now proved the existence of a vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$?

RQ

As is probably clear to you by now, the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ is just a different way of writing down a system of linear equations. For example, the system of linear equations

$$\begin{aligned} x_1 + 3\,x_2 + 9\,x_3 - 8\,x_5 + 2\,x_6 &= -1 \\ 4\,x_2 + x_3 + 3\,x_4 - 2\,x_6 &= 1 \\ 6\,x_1 - x_2 - 3\,x_3 + 9\,x_4 + x_5 &= 5 \end{aligned}$$

is equivalent to the matrix-vector equation

$$\begin{bmatrix} 1 & 3 & 9 & 0 & -8 & 2 \\ 0 & 4 & 1 & 3 & 0 & -2 \\ 6 & -1 & -3 & 9 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 5 \end{bmatrix}.$$

What's a consistent system of linear equations?

If the equation $A\mathbf{x} = \mathbf{b}$ has a solution, then we'll say that it, and its associated system of linear equations, is **consistent** (otherwise, it is **inconsistent**).

Let $S = \{\mathbf{a}_1, \ldots, \mathbf{a}_k\}$ be a set of vectors in $\mathbf{R}^n$ and let $A = [\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_k]$.

RQ

Reading Question 1N. Explain the difference between S and A, just above.

Over the next two chapters, we will (among other things) answer the following questions, where S and A are as above.

- How can we algorithmically determine whether a vector $\mathbf{b}$ is in span S? Equivalently: when is $A\mathbf{x} = \mathbf{b}$ consistent?
- When $\mathbf{b}$ *does* lie in span S, it is by definition a linear combination of the vectors in S. How can we find the weights? Equivalently: when $A\mathbf{x} = \mathbf{b}$ is consistent, how can we actually find a solution $\mathbf{x}$?
- When $\mathbf{b}$ *does* lie in span S, is there more than one way to choose the weights to build $\mathbf{b}$ as a linear combination of the vectors in S? Equivalently: how many solutions does $A\mathbf{x} = \mathbf{b}$ have?

RQ

Reading Question 1O. Go ahead and find a solution for equation (1.1) on page 19. Is the solution you found the only one?

RQ

Reading Question 1P. You need to be sure that you understand the difference between types of mathematical objects, even when they seem to carry the same underlying information. The following mathematical objects are all different; what are they called in linear algebra?

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

② $\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}-1\\0\\9\end{bmatrix}\right\}$

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

④ $[7]$

⑤ 7

Reading Question 1Q. Which of the following statements make precise mathematical sense, and which are a bit off the mark? For each statement that is off the mark, describe what's wrong.

RQ

① The span of $\begin{bmatrix}1 & 3\\2 & 4\end{bmatrix}$ is $\begin{bmatrix}4\\6\end{bmatrix}$.

② The span of $\begin{bmatrix}1 & 3\\2 & 4\end{bmatrix}$ contains $\begin{bmatrix}4\\6\end{bmatrix}$.

③ The span of $\left\{\begin{bmatrix}1\\2\end{bmatrix} \begin{bmatrix}3\\4\end{bmatrix}\right\}$ contains $\begin{bmatrix}4\\6\end{bmatrix}$.

Exercise 1R. Describe the set of points (a, b) in the plane $\mathbf{R}^2$ such that

$$\begin{bmatrix}a\\b\end{bmatrix} \in \operatorname{span}\left\{\begin{bmatrix}-3\\4\end{bmatrix}, \begin{bmatrix}4\\-8\end{bmatrix}\right\}.$$

Rephrase this question to make it about whether (depending on a and b) a certain matrix-vector equation has a solution.

Exercise 1S. In each case, give an example (or explain why it isn't possible) of a 2×2 matrix A such that the equation $A\mathbf{x} = \mathbf{b}$ ...

① ... has a solution for every $\mathbf{b} \in \mathbf{R}^2$.

② ... has a solution for some $\mathbf{b}$ but not all.

③ ... never has a solution (no matter what $\mathbf{b}$ is).

④ ... has infinitely many solutions, for any $\mathbf{b}$.

Exercise 1T. Let A be an $n \times k$ matrix. You don't know what A is, but an oracle will tell you the vector $A\mathbf{x}$ for any specific vector $\mathbf{x}$. What is the smallest number of vectors $\mathbf{x}$ that you need to give the oracle so that you're guaranteed to be able figure out what A is?

Key concepts

- Functions: a domain, a codomain, and a rule for evaluation
- One-to-one and onto functions
- Linear transformations
- How to represent a linear transformation with a matrix
- Examples where linear transformations arise in the wild
- Linearly independent and dependent sets
- Homogeneous linear systems

Summary. This chapter covers linearly independent sets and linear transformations. Concerning linear (in)dependence, the two lists of equivalent statements in Theorem 2.13 summarize most of what we need to know; the lists are really the same!

The difference between a matrix and a linear transformation is a subtle one. Every matrix determines a linear transformation: given an $n \times k$ matrix A, there is a linear transformation $T\colon \mathbf{R}^k \to \mathbf{R}^n$ defined by $T(\mathbf{x}) = A\mathbf{x}$. The conceptual transition from A to T reflects a desire to view the matrix A *as a function.* On the other hand, if a function T is a linear mapping, then it is represented by the matrix

$$A = \begin{bmatrix} T(\mathbf{e}_1) & \dots & T(\mathbf{e}_k) \end{bmatrix}.$$

This means, among other things, that any rotation of the plane about the origin has a matrix representation.

Key connections are established here between the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ and concepts such as span, linear combination, and linearly independent set (see Theorems 2.9 and 2.17). The transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one if and only if the columns of A are linearly independent, if and only if the only solution to $A\mathbf{x} = \mathbf{0}$ is the zero vector. This linear transformation is onto if and only if the columns of A span the codomain, if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution *no matter what* $\mathbf{b}$ *is.*

Chapter 2

We have seen that linear algebra provides a framework for studying systems of linear equations. In Example 1.4 we also saw that there's a way to use linear algebra to represent certain geometric functions (also called maps or transformations). Let's dig into that example a bit further.

Example 2.1 (Example 1.4, continued)**.** At the end of Example 1.4, we found a formula for counter-clockwise rotation of the plane by $\pi/4$ radians about the origin:

$$R\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = x_1 \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} + x_2 \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

By definition of the matrix-vector product, we obtain

$$R\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Thus we see that applying the rotation R to the vector $\mathbf{x}$ is exactly the same as multiplying $\mathbf{x}$ by the 2×2 matrix

$$A = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

What's a function?

Functions are ubiquitous in mathematics. Recall that, if X and Y are any two nonempty sets, a **function** is a rule, written $f\colon X \to Y$, that assigns to each element $x \in X$ a unique element $f(x) \in Y$. In this context, the element x is called an **input** of the function and the element $f(x)$ is called the corresponding **output** or **value** of the function. When X and Y are clear from context, we often write $f\colon X \to Y$ simply as f. The set X of possible inputs to f is called the **domain** of f and the set Y where the outputs live is called the **codomain** of f.

What's the domain of a function? The codomain?

The function R in Example 2.1 has domain $\mathbf{R}^2$ and codomain $\mathbf{R}^2$ (since the inputs and outputs are both vectors in the plane.)

Reading Question 2A. Write down a function $f : \mathbf{R}^3 \to \mathbf{R}$ (that is, a function with domain $\mathbf{R}^3$ and codomain $\mathbf{R}$).

RQ

Definition and examples §2.1

Just as we did in Example 2.1, we can take any $n \times k$ matrix A and use it to define a function

$$T\colon \mathbf{R}^k \to \mathbf{R}^n$$

in the following way:

$$T(\mathbf{x}) = A\mathbf{x}.$$

Take care: T and A are different mathematical objects; T is a function and A is a matrix.

This definition makes sense: since A is $n \times k$ and $\mathbf{x}$ is a k-vector, the product $A\mathbf{x}$ is defined and is an n-vector. Thus our function T really does map from $\mathbf{R}^k$ to $\mathbf{R}^n$ (in other words, T really does have domain $\mathbf{R}^k$ and codomain $\mathbf{R}^n$). We will from time to time use the following alternative notation for this linear transformation:

$$\mathbf{x} \mapsto A\mathbf{x}.$$

What kind of object is $\mathbf{x} \mapsto A\mathbf{x}$?

By Theorem 1.9, this function T has a special property: if $\mathbf{v}$ and $\mathbf{w}$ are any two vectors and s and t are any two scalars, we have

$$\begin{aligned} T(s\mathbf{v} + t\mathbf{w}) &= A(s\mathbf{v} + t\mathbf{w}) && \text{[definition of } T\text{]} \\ &= sA\mathbf{v} + tA\mathbf{w} && \text{[Theorem 1.9]} \\ &= sT(\mathbf{v}) + tT(\mathbf{w}) && \text{[definition of } T \text{ again].} \end{aligned}$$

This property is very important and has a name.

LINEAR TRANSFORMATIONS BETWEEN EUCLIDEAN SPACES

Definition 2.2. A **linear transformation** is a function

$$T\colon \mathbf{R}^k \to \mathbf{R}^n$$

that satisfies

$$T(s\mathbf{v} + t\mathbf{w}) = sT(\mathbf{v}) + tT(\mathbf{w})$$

for all $s, t \in \mathbf{R}$ and all $\mathbf{v}, \mathbf{w} \in \mathbf{R}^k$.

What's a linear transformation?

The linearity property readily extends to more than two terms; if T is a linear transformation, then

$$T(s_1\mathbf{v}_1 + \cdots + s_p\mathbf{v}_p) = s_1T(\mathbf{v}_1) + \cdots + s_pT(\mathbf{v}_p)$$

for all $s_1, \ldots, s_p \in \mathbf{R}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_p \in \mathbf{R}^k$.

RQ

Reading Question 2B. Write down what the equation in Definition 2.2 says when $s = t = 1$. In English, what you wrote says "the T-value of the sum is the sum of the T-values." Next, write down what the equation says when $t = 0$. In English, what you wrote says "the T-value of a scalar multiple is the scalar multiple of the T-value." Can you think of any operations in Calculus that behave similarly?

Please understand that linearity is a very particular property; most functions are **not** linear.

RQ

Reading Question 2C. Find an example of function $f : \mathbf{R}^2 \to \mathbf{R}^2$ that is *not* linear, and explain how you know it's not linear.

RQ

Reading Question 2D. Suppose T is a linear transformation where $T(\mathbf{v}) = 3\mathbf{a} + \mathbf{b}$ and $T(\mathbf{w}) = -\mathbf{a} + 2\mathbf{b}$. Compute $T(7\mathbf{v} + 8\mathbf{w})$. Your answer will be a linear combination of $\mathbf{a}$ and $\mathbf{b}$.

Example 2.3 (scaling)**.** For any positive integer n and any fixed scalar α, the map

$$T\colon \mathbf{R}^n \to \mathbf{R}^n$$

defined by $\mathbf{x} \mapsto \alpha\mathbf{x}$ defines a linear transformation. To see this, we just choose arbitrary vectors $\mathbf{v}$ and $\mathbf{w}$ and scalars s and t and check that Definition 2.2 holds:

$$\begin{aligned} T(s\mathbf{v} + t\mathbf{w}) &= \alpha(s\mathbf{v} + t\mathbf{w}) \\ &= \alpha(s\mathbf{v}) + \alpha(t\mathbf{w}) \\ &= s(\alpha\mathbf{v}) + t(\alpha\mathbf{w}) \\ &= sT(\mathbf{v}) + tT(\mathbf{w}). \end{aligned}$$

Below, you can see the effect of scaling by $t = 1/2$ and $t = 2$ on the plane $\mathbf{R}^2$. Figure 2.1 shows what each scaling operation does to the unit square in the first quadrant with a vertex at the origin.

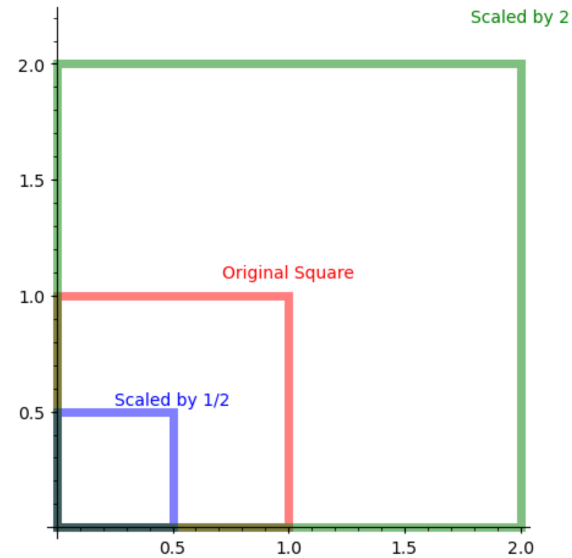


Figure 2.1: Scaling a square

Exercise 2A. If T is a linear transformation with

$$T\left(\begin{bmatrix}2\\5\end{bmatrix}\right)=\begin{bmatrix}1\\0\end{bmatrix} \text{ and } T\left(\begin{bmatrix}-1\\1\end{bmatrix}\right)=\begin{bmatrix}2\\2\end{bmatrix},$$

then find $T\left(\begin{bmatrix}0\\7\end{bmatrix}\right)$. [Hint: first, fill in the boxes with the correct scalars:

$$\begin{bmatrix}0\\7\end{bmatrix}=\square\begin{bmatrix}2\\5\end{bmatrix}+\square\begin{bmatrix}-1\\1\end{bmatrix}$$

Then, apply T to both sides of this equation; use the fact that T is linear, and the assumptions given in the problem, to finish the job.]

As we remarked before Definition 2.2, a function $T : \mathbf{R}^k \to \mathbf{R}^n$ defined by multiplication by an $n \times k$ matrix is always linear. Therefore, one way to prove that a given function T is linear is to show that it is defined by multiplication by a matrix! Here is an example of this type.

Example 2.4. Consider the function $T\colon \mathbf{R}^3 \to \mathbf{R}^3$ defined by

$$T(x,y,z) = (x-z,\ y+z,\ x-y+z).$$

We claim that T is linear. To prove this, we will find a matrix A such that $T(\mathbf{v}) = A\mathbf{v}$ for all $\mathbf{v} = (x, y, z) \in \mathbf{R}^3$. Look:

$$\begin{aligned} T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) &= \begin{bmatrix} x - z \\ y + z \\ x - y + z \end{bmatrix} \\ &= \begin{bmatrix} x \\ 0 \\ x \end{bmatrix} + \begin{bmatrix} 0 \\ y \\ -y \end{bmatrix} + \begin{bmatrix} -z \\ z \\ z \end{bmatrix} \\ &= x\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + z\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}}_{A} \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \end{aligned}$$

Look how we used vector addition to "separate" the variables.

We found the matrix A!

Examples 2.1 and 2.4 prompt the question of whether there are any linear transformations $T : \mathbf{R}^k \longrightarrow \mathbf{R}^n$ that are **not** defined by multiplication by some matrix A. It turns out that the answer is *no*. To see why, let's define a list of special vectors in $\mathbf{R}^k$:

$$\begin{aligned} \mathbf{e}_1 &= \overbrace{(1, 0, 0, \ldots, 0, 0)}^{k \text{ entries}} \\ \mathbf{e}_2 &= (0, 1, 0, \ldots, 0, 0) \\ \mathbf{e}_3 &= (0, 0, 1, \ldots, 0, 0) \\ &\vdots \\ \mathbf{e}_{k-1} &= (0, 0, 0, \ldots, 1, 0) \\ \mathbf{e}_k &= (0, 0, 0, \ldots, 0, 1). \end{aligned}$$

What are the vectors $\mathbf{e}_i$?

Simply put, $\mathbf{e}_i$ is the vector you get if you replace the ith entry in the zero vector with a 1. There are two such vectors in $\mathbf{R}^2$: $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$. There are three such vectors in $\mathbf{R}^3$: $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, and $\mathbf{e}_3 = (0, 0, 1)$. These vectors are important because every vector in $\mathbf{R}^k$ is a linear combination of them:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ x_k \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_k\mathbf{e}_k.$$

Now let $T\colon \mathbf{R}^k \longrightarrow \mathbf{R}^n$ be a linear transformation. Using linearity, we have

$$\begin{aligned} T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_k\mathbf{e}_k) &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_kT(\mathbf{e}_k) \\ &= \underbrace{[T(\mathbf{e}_1)\ T(\mathbf{e}_2)\ \cdots\ T(\mathbf{e}_k)]}_{A}\,\mathbf{x}. \end{aligned}$$

We therefore see that *doing T to $\mathbf{x}$ is the same as multiplying $\mathbf{x}$ by the matrix A whose columns are $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_k)$*. We have proven the following theorem.

The matrix of a linear transformation

Theorem 2.5. *Let*

$$T\colon \mathbf{R}^k \longrightarrow \mathbf{R}^n$$

be a linear transformation. The matrix

$$A = [T(\mathbf{e}_1)\ T(\mathbf{e}_2)\ \cdots\ T(\mathbf{e}_k)]$$

is called the **(standard) matrix representation** *of T and it is the unique matrix A satisfying $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbf{R}^k$.*

How do we find the matrix of a linear transformation?

Later in the book, we'll look at other ways to represent a linear transformation using a matrix. Usually, we will just refer to the above matrix as the matrix representation of T, but if we are worried about confusion we'll call it the *standard* matrix representation.

Reading Question 2E. Find the matrix representation for scaling in Example 2.3 when $n = 3$. Verify that the matrix in Example 2.4 has columns given by $T(\mathbf{e}_1), T(\mathbf{e}_2)$, and $T(\mathbf{e}_3)$ by computing these three quantities directly using the definition of T.

RQ

Example 2.6 (rotating the plane)**.** Let T denote rotation of the plane counterclockwise by θ radians about the origin in $\mathbf{R}^2$. Let's find the matrix representation A for T. The vector $T(\mathbf{e}_1)$ in Figure 2.2 has coordinates $(\cos\theta, \sin\theta)$ by definition of sine and cosine. The vector $T(\mathbf{e}_2)$ is $\pi/2$ radians away (counterclockwise) from $T(\mathbf{e}_1)$.

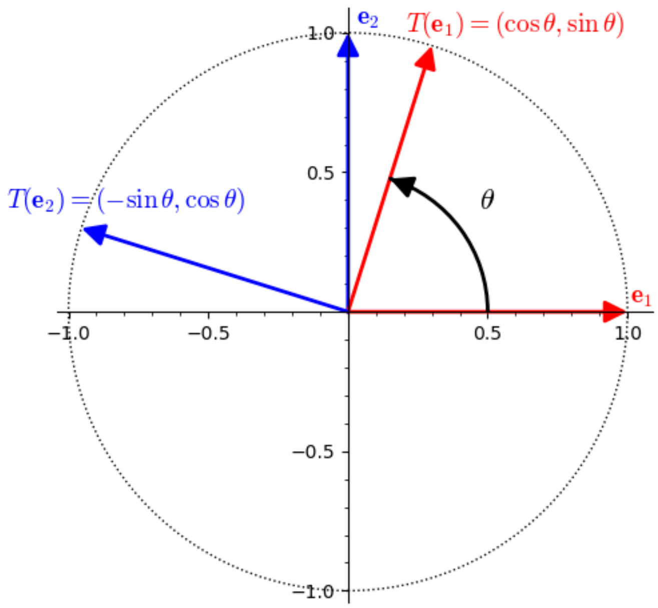


Figure 2.2: Rotation by θ CCW

So, we have

$$T(\mathbf{e}_1) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}$$
$$T(\mathbf{e}_2) = \begin{bmatrix} \cos(\theta + \pi/2) \\ \sin(\theta + \pi/2) \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}.$$

From this we obtain

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

by Theorem 2.5.

Exercise 2B. Let T be be the linear transformation from $\mathbf{R}^2$ to itself that first reflects the plane across the line $y = x$ and then rotates the plane counter-clockwise $\pi/4$ radians about the origin. Find the matrix representation for T. [Hint: draw pictures to figure out where T sends $\mathbf{e}_1$ and $\mathbf{e}_2$.] What if you do the two aforementioned operations in the other order?

Example 2.7. Look back at the model for lionfish populations in Example 1.2. Recall that the numbers L, J, and A are the current populations of larvae, juveniles, and adults; and that L', J', and A' are the predicted populations one month in the future. With matrix-by-vector multiplication in hand, we can now write

$$\begin{bmatrix} L' \\ J' \\ A' \end{bmatrix} = \begin{bmatrix} 0 & 0 & 35315 \\ 0.0003 & 0.777 & 0 \\ 0 & 0.071 & 0.949 \end{bmatrix} \begin{bmatrix} L \\ J \\ A \end{bmatrix}.$$

In other words, the predicted future population vector (L', J', A') is a linear transformation of the current population vector (L, J, A). By applying the same linear transformtion to (L', J', A') we obtain the predicted population *two* months in the future, and so on.

Exercise 2C. Imagine an insect that has two life stages: adult and juvenile. Suppose that, if J and A are the populations of adults and juveniles this year, then the predicted populations J' and A' in one year are given by

$$\begin{bmatrix} J' \\ A' \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ .5 & .5 \end{bmatrix} \begin{bmatrix} J \\ A \end{bmatrix}.$$

Write down the interpretation of the four entries in the above matrix. Then figure out, if there are 100 adults and no juveniles this year, what the predicted numbers of junveniles and adults are ten years from now.

Exercise 2D. Find the matrix representation for the linear mapping $T(x_1, x_2, x_3, x_4) = (x_1 + 2x_2, 0, 2x_2 + x_4, x_2 - x_4)$.

Onto linear transformations §2.2

For any function $f\colon X \to Y$, the **image** of f is the set of outputs of f, defined by

What's the image of a function?

$$\operatorname{im} f = \{f(x) \mid x \in X\}.$$

The image of a function is also often called its **range**.

The distinction between a function's image and its codomain is important. The codomain is the set where the outputs live; the image is the set of outputs that actually occur. For example, consider the function $f\colon \mathbf{R} \to \mathbf{R}$ defined by $f(x) = \sin x$. The codomain of f is $\mathbf{R}$ — all outputs are real. It is not the case, however, that every real number *actually occurs* as an output of f. In fact, $\sin x$

takes on every value in between −1 and 1 (inclusive) and no others, so $\operatorname{im} f = [-1, 1]$.

For linear transformations, images are spans.

The image of $\mathbf{x} \mapsto A\mathbf{x}$ is the span of the columns of A.

For linear transformations, images are spans

Theorem 2.8. *Let $T: \mathbf{R}^k \to \mathbf{R}^n$ be a linear transformation with matrix representation $A = [\mathbf{a}_1 \ \dots \ \mathbf{a}_k]$. Then,*

$$\operatorname{im} T = \operatorname{span}\{\mathbf{a}_1, \dots, \mathbf{a}_k\}.$$

Proof. This follows from Theorem 1.10 and the definition of image: $\mathbf{b} \in \operatorname{im} T$ if and only if $A\mathbf{x} = \mathbf{b}$ has a solution, if and only if $\mathbf{b}$ is in the span of the columns of A. ■

RQ

Reading Question 2F. For each matrix A below, find the domain, codomain, and image of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$.

① $\begin{bmatrix} 5 & 1 \\ 10 & 2 \end{bmatrix}$

② $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

③ $\begin{bmatrix} 1 & 1 \end{bmatrix}$

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

⑤ $\begin{bmatrix} 9 & 0 & 1 & 1 \\ 2 & 3 & 1 & 1 \end{bmatrix}$

What's an onto function?

The situation where the image and codomain of f are the same is special, so we give it a name. A function f is **onto** if its codomain is equal to its image. Since the image is always contained in the codomain, to prove that a function is onto it suffices to do the following.

How do I prove that a funtion is onto?

How to show that a function is onto

Let $f: X \to Y$ be a function. To show f is onto, take an arbitrary $y \in Y$ and show that there exists some $x \in X$ where $f(x) = y$. This proves that every element in the codomain is in the image.

Reading Question 2G. Prove that the function $f: \mathbf{R} \to \mathbf{R}$ defined by $f(x) = 3x - 2$ is onto.

RQ

The following theorem connects onto functions to what we've learned so far. If S is a set of vectors in $\mathbf{R}^n$ and $\operatorname{span} S = \mathbf{R}^n$, then we'll say that the vectors in S span $\mathbf{R}^n$.

Onto linear transformations

Theorem 2.9. *Let A be an $n \times k$ matrix. The following statements are equivalent.*

1. *The equation $A\mathbf{x} = \mathbf{b}$ has a solution for all $\mathbf{b} \in \mathbf{R}^n$.*
2. *The columns of A span $\mathbf{R}^n$.*
3. *The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is onto.*

How does "onto" relate to other concepts in linear algebra so far?

Reading Question 2H. Let

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

Find a nonzero vector $\mathbf{b} \in \mathbf{R}^3$ so that $A\mathbf{x} = \mathbf{b}$ has a solution, and then find another vector $\mathbf{b}$ where $A\mathbf{x} = \mathbf{b}$ doesn't have a solution. Is $\mathbf{x} \mapsto A\mathbf{x}$ onto? Describe the image geometrically.

RQ

Exercise 2E. Prove Theorem 2.9 by proving the following implications:

$$(1) \implies (2) \implies (3) \implies (1).$$

You can do it – unravel the definitions of the concepts involved.

Exercise 2F. Can a linear transformation $\mathbf{R} \to \mathbf{R}^3$ be onto? Why or why not? How many linear transformations $\mathbf{R}^3 \to \mathbf{R}$ are *not* onto?

Exercise 2G. Each item below gives a function $\mathbf{R}^k \to \mathbf{R}^n$. Which functions are onto? Which functions are linear?

1. $(x, y, z) \mapsto (x^3, y^3, z^3)$
2. $(x, y, z) \mapsto (x + y, 2y, x)$

③ $(x, y) \mapsto (x + y, x - y)$
④ $(x, y, z) \mapsto (x - y, x + z, 1)$

---

§2.3 Linear independence

Suppose the equation $A\mathbf{x} = \mathbf{b}$ is consistent. Then, there is either one and only one solution to this equation or more than one solution to this equation. What must be true about the columns of A so that if there *is* a solution, it's unique? Let's explore a bit. Suppose we have two distinct solutions $\mathbf{c}$ and $\mathbf{d}$. Then,

A solution to $A\mathbf{x} = \mathbf{b}$ doesn't need to be called $\mathbf{x}$. This is especially clear if you need to talk about more than one solution!

$$A\mathbf{c} = \mathbf{b} \text{ and } A\mathbf{d} = \mathbf{b}.$$

This implies that $A\mathbf{c} = A\mathbf{d}$, and using the fact that the matrix-vector product is linear we obtain:

$$\begin{aligned} A\mathbf{c} &= A\mathbf{d} \\ A\mathbf{c} - A\mathbf{d} &= \mathbf{0} \\ A(\mathbf{c} - \mathbf{d}) &= \mathbf{0}. \end{aligned}$$

Since $\mathbf{c} - \mathbf{d} \neq \mathbf{0}$, we have a nonzero solution to $A\mathbf{x} = \mathbf{0}$. This observation goes both ways. Suppose that $A\mathbf{x} = \mathbf{b}$ has a solution (let's call it $\mathbf{c}$) and that $A\mathbf{x} = \mathbf{0}$ has a nonzero solution (let's call it $\mathbf{z}$). Then

A nonzero vector is any vector that is not precisely equal to the zero vector $\mathbf{0}$.

$$\begin{aligned} A(\mathbf{c} + \mathbf{z}) &= A\mathbf{c} + A\mathbf{z} \\ &= \mathbf{b} + \mathbf{0} \\ &= \mathbf{b}, \end{aligned}$$

and so we have found a solution to $A\mathbf{x} = \mathbf{b}$ (namely, $\mathbf{c} + \mathbf{z}$) that is different from $\mathbf{c}$. Let's sum up. Assuming $A\mathbf{x} = \mathbf{b}$ is consistent:

- $A\mathbf{x} = \mathbf{b}$ *has more than one solution if and only if* $A\mathbf{x} = \mathbf{0}$ *has a nonzero solution;* or, equivalently,
- $A\mathbf{x} = \mathbf{b}$ *has a unique solution if and only if the only solution to* $A\mathbf{x} = \mathbf{0}$ *is* $\mathbf{x} = \mathbf{0}$.

Let's focus our attention on the **homogeneous equation** $A\mathbf{x} = \mathbf{0}$, also called a **homogeneous linear system**. If we write $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_k]$ and $\mathbf{x} = (x_1, \ldots, x_k)$, the equation $A\mathbf{x} = \mathbf{0}$ can be written

What's a homogeneous equation?

$$x_1\mathbf{a}_1 + \cdots + x_k\mathbf{a}_k = \mathbf{0}.$$

Certainly, the zero vector $\mathbf{x} = \mathbf{0} = (0, \ldots, 0)$ is a solution for any homogeneous equation, so it is sometimes referred to as the **trivial solution**. To help us further analyze the solution set for this equation, we need three definitions.

What's the trivial solution to a homogeneous equation?

Reading Question 2I. Suppose the homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a nonzero solution $\mathbf{c}$. Explain why it must have infinitely many solutions.

RQ

Linearly dependent and linearly independent sets

What's a linearly dependent set? What's a dependence relation? What's a linearly independent set?

Definition 2.10. Let $S = \{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ be a set of vectors in $\mathbf{R}^n$. Consider the equation

$$x_1\mathbf{a}_1 + \cdots + x_k\mathbf{a}_k = \mathbf{0}. \qquad (\star)$$

① If equation $(\star)$ has a nonzero solution, then S is called a **linearly dependent set**.

A nonzero solution is one where $x_i \neq 0$ for at least one index i.

② If $\mathbf{c} = (c_1, \dots, c_k)$ is a specific nonzero solution to equation $(\star)$, then

$$c_1\mathbf{a}_1 + \cdots + c_k\mathbf{a}_k = \mathbf{0}$$

is called a **dependence relation** for S.

③ If the only solution to equation $(\star)$ is $\mathbf{x} = \mathbf{0}$, then S is called a **linearly independent set**.

Example 2.11. The set

$$S = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ -6 \end{bmatrix} \right\}$$

is linearly dependent because

$$3\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 1\begin{bmatrix} -3 \\ -6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

is a dependence relation for S.

Example 2.12. Let's show that the set

$$S = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} \right\}$$

is linearly independent. We need to solve the equation

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

This equation can be written as the linear system

$$\begin{aligned} x_1 + 2\,x_3 &= 0 \\ 2\,x_1 + x_2 - x_3 &= 0 \\ x_1 + 2\,x_2 &= 0 \end{aligned}$$

From the third line, $x_1 = -2x_2$. Plugging this into the second line and solving for x_3, we obtain $x_3 = -3x_2$. From the first line it now folows that $0 = -2x_2 + 2(-3x_2) = -8x_2$, which forces $x_2 = 0$, and this in turn implies $x_1 = x_3 = 0$. We have shown that the only solution to our equation is the zero vector, which means the set S is linearly independent.

RQ

Reading Question 2J. Prove that $\{\mathbf{0}\}$ is a linearly dependent set. (In fact, any set containing $\mathbf{0}$ is linearly dependent.) You need to study the equation $x_1\mathbf{0} = \mathbf{0}$.

Any set containing $\mathbf{0}$ *is linearly dependent.*

Exercise 2H. Convert each sentence into mathematical symbols and then convert what you wrote into a dependence relation.

1. "One vector is a scalar multiple of another vector."
2. "One vector is in the span of two other vectors."

Exercise 2I. Determine whether each set below is linearly independent.

1. $\{(5, 0, 1), (2, 3, -1), (-1, 0, -1)\}$
2. $\{(1, 2, 1), (3, 0, 2), (5, 4, 4)\}$

Exercise 2J. Suppose $\{\mathbf{a}, \mathbf{b}\}$ is a set of linearly independent vectors in $\mathbf{R}^k$. What about $\{\mathbf{a} - \mathbf{b}, \mathbf{b} - \mathbf{a}\}$? What about $\{\mathbf{a} - \mathbf{b}, \mathbf{b} + \mathbf{a}\}$? In each case, either give a dependence relation or prove that the set is linearly independent using the definition.

Let's next look at four equivalent statements that characterize what it means for a set to be linearly independent. By negating each of these statements, we obtain a list of statements that are equivalent to a set being linearly dependent.

LINEAR INDEPENDENCE VS. LINEAR DEPENDENCE

What are the different ways to think about linear independence and dependence?

Theorem 2.13. *Let $S = \{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ be a set of vectors in $\mathbf{R}^n$ and let $A = [\mathbf{a}_1 \ \dots \ \mathbf{a}_k]$.*

Linear Independence

The following are equivalent:

1. *S is a linearly independent set.*
2. *The only solution to*
$$x_1\mathbf{a}_1 + \cdots + x_k\mathbf{a}_k = \mathbf{0}$$
is the zero vector.
3. *The homogeneous equation*
$$A\mathbf{x} = \mathbf{0}$$
has a unique solution.
4. *No $\mathbf{a}_i$ lies in the span of the remaining vectors in S.*

$\{\mathbf{v}, \mathbf{w}\}$ is linearly independent iff $\mathbf{v}, \mathbf{w}$ span a plane.

Linear Dependence

The following are equivalent:

1. *S is a linearly dependent set.*
2. *The equation*
$$x_1\mathbf{a}_1 + \cdots + x_k\mathbf{a}_k = \mathbf{0}$$
has a nonzero solution.
3. *The homogeneous equation*
$$A\mathbf{x} = \mathbf{0}$$
has ∞-many solutions.
4. *Some $\mathbf{a}_i$ lies in the span of the remaining vectors in S.*

$\{\mathbf{v}, \mathbf{w}\}$ is linearly dependent iff $\mathbf{v}, \mathbf{w}$ span a line or $\{\mathbf{0}\}$.

We hope you agree that (1), (2), and (3) are three ways of saying the same thing (and so are equivalent). The equivalence of (1) and (4) is less obvious, and is left to the following exercise.

Exercise 2K. Show that (1) and (4) are equivalent in Theorem 2.13.

Exercise 2L. For each case below, find a set of distinct, nonzero vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ in $\mathbf{R}^3$ that satisfies the given conditions.

1. S is linearly dependent and $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly independent.
2. S is linearly dependent but no two of the vectors form a linearly dependent set.
3. S is linearly independent and every entry in every vector is nonzero.

Exercise 2M. True or false: If $S = \{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is a linearly dependent set, then some vector in S is a scalar multiple of another vector in S. If it's true, prove it. If it's false, give a specific example that shows it's false. Is it true for some values of p but not for others (that is, does the truth of the statement depend on the size of S)?

Suppose S is a finite set of vectors that isn't equal to $\{\mathbf{0}\}$. Since items (1) and (4) above are equivalent, if S contains a vector that lies in the span of the remaining vectors, then we can throw it out without changing the span. Repeating this process, we end up with a set S' of linearly independent vectors that have the same span as S. We give more details in the proof of the next theorem.

The set $\{\mathbf{0}\}$ is linearly dependent, and if we remove $\mathbf{0}$ we are left with the empty set $\emptyset = \{\}$ (the set with no elements). So, we'll declare that the empty set is linearly independent and that the span of the empty set is $\{\mathbf{0}\}$. This convention makes the statement of the next theorem cleaner than it would be otherwise.

YOU CAN ALWAYS REPLACE A SPANNING SET WITH A LI SET

Every span is actually the span of a LI set!

Theorem 2.14. *For any finite set of vectors S, there is a linearly independent subset S' of S such that* $\operatorname{span} S' = \operatorname{span} S$.

Proof. If $S = \{\mathbf{0}\}$, then we can take $S' = \emptyset$. Otherwise, S contains at least one nonzero vector. If S is linearly independent, then we can take $S' = S$. Otherwise, there is some vector $\mathbf{v} \in S$ that lies in the span of the remaining vectors. Denote the remaining vectors in S by $\mathbf{a}_1, \dots \mathbf{a}_k$:

$$S = \{\mathbf{v}, \mathbf{a}_1, \dots, \mathbf{a}_k\}.$$

We can write $\mathbf{v}$ as a linear combination of the vectors $\mathbf{a}_i$:

$$\mathbf{v} = c_1\mathbf{a}_1 + \cdots + c_k\mathbf{a}_k.$$

Choose any vector in $\operatorname{span} S$:

$$\mathbf{w} = t\mathbf{v} + t_1\mathbf{a}_1 + \cdots + t_k\mathbf{a}_k.$$

Then,

$$\begin{aligned}\mathbf{w} &= t\mathbf{v} + t_1\mathbf{a}_1 + \cdots + t_k\mathbf{a}_k \\ &= t(c_1\mathbf{a}_1 + \cdots + c_k\mathbf{a}_k) + t_1\mathbf{a}_1 + \cdots + t_k\mathbf{a}_k \\ &= (tc_1 + t_1)\mathbf{a}_1 + \cdots + (tc_k + t_k)\mathbf{a}_k.\end{aligned}$$

We have shown that we can *remove the vector* $\mathbf{v}$ *from S without changing* $\operatorname{span} S$.

We can repeat this process until we arrive at a linearly independent set S' whose span is span S, or a set that contains only the zero vector. ■

Example 2.15. Consider the set of vectors in $\mathbf{R}^3$:

$$S = \left\{ \mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \right\}.$$

Since $\mathbf{v} = \mathbf{u} + \mathbf{w}$, $\mathbf{v}$ is a linear combination of the remaining two vectors, so we can throw it out of S without changing the span:

$$\operatorname{span} S = \operatorname{span}\{\mathbf{u}, \mathbf{w}\}.$$

This set of remaining vectors is linearly independent: if it weren't, then one of the two remaining vectors would have to be a linear combination of the other. This is the same as saying it would need to be a scalar multiple of the other. But this is not the case (you should check this). So, we've found a linearly independent subset of S that has the same span as S.

Reading Question 2K. Let $S = \{(1,-1), (2,-2), (7,26), (3,9)\}$. Find a linearly independent subset of S with the same span as S. Show that every vector you throw out is a linear combination of the ones that remain. RQ

Exercise 2N. Let $\mathbf{v}$ be a vector in $\mathbf{R}^n$. When is $\{\mathbf{v}\}$ a linearly independent set? Let $\mathbf{v}, \mathbf{w}$ be non-zero vectors in $\mathbf{R}^n$. If $S = \{\mathbf{v}, \mathbf{w}\}$ is linearly independent, then what is span S? If S is linearly dependent, then what is span S?

One-to-one linear transformations §2.4

Suppose $f : X \to Y$ is a function and suppose y is in the image of f. This means there is an $x \in X$ such that $f(x) = y$. But is there more than one such x? If it's never the case that two distinct inputs can map to the same output, then we call the function one-to-one. Put another way, f is **one-to-one** if, for every $y \in Y$, there is *at most* one $x \in X$ with $f(x) = y$ (there will be no such x if y is not in the image of f).

What's a one-to-one function?

Example 2.16. Define $f\colon \mathbf{R} \to \mathbf{R}$ by $f(x) = x^2$. Then, f is *not* one-to-one because $f(1) = 1^2 = (-1)^2 = f(-1)$. The distinct inputs 1 and -1 map to the same output.

Define $g\colon [0,\infty) \to \mathbf{R}$ by $g(x) = x^2$. The function g *is* one-to-one because, for any real number y, the equation $y = x^2$ has no solutions when $y < 0$ and it has exactly one nonnegative solution when $y \geq 0$.

"A function **is** *its graph."*

Note that the functions f and g have the same rule for evaluation, but they are NOT the same function because they have different domains! The graphs of these functions are different:

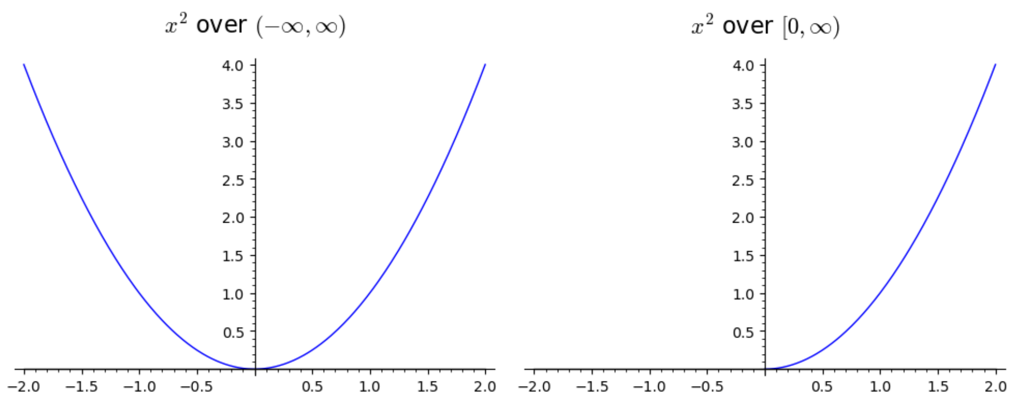


Figure 2.3: $x \mapsto x^2$ over two different domains

How do I prove that a function is one-to-one?

How to show that a functon is one-to-one

Let $f\colon X \to Y$ be a function. To show f is one-to-one, take a pair of arbitrary elements $a, b \in X$ and suppose $f(a) = f(b)$. Use this to prove that $a = b$. This then implies that you can't have two distinct inputs mapping to the same output.

RQ

Reading Question 2L. Prove that the function in Reading Question 2G is one-to-one.

RQ

Reading Question 2M. Students sometimes say that "a function being one-to-one means every input has a unique output." You would never say that, though, because it's wrong. Explain.

Exercise 2O. Prove that if T is a linear transformation, then $T(\mathbf{0}) = \mathbf{0}$. If $T: \mathbf{R}^k \to \mathbf{R}^n$ is an arbitrary function with $T(\mathbf{0}) = \mathbf{0}$, then does it necessarily follow that T is linear?

A linear transformation must send $\mathbf{0}$ *to* $\mathbf{0}$.

The next theorem brings one-to-one functions into our circle of ideas.

ONE-TO-ONE LINEAR TRANSFORMATIONS

Theorem 2.17. *Let A be an $n \times k$ matrix. The following statements are equivalent.*

1. *The equation $A\mathbf{x} = \mathbf{0}$ has a unique solution.*
2. *The columns of A are linearly independent.*
3. *The linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.*

How does "one-to-one" relate to other concepts in linear algebra so far?

Proof. Items (1) and (2) are equivalent by definition of linearly independent. Statements (3) and (1) are equivalent by Exercise 2P below. ■

Exercise 2P. Show that if T is a linear transformation, then T is one-to-one if and only if the zero vector is the only solution to $T(\mathbf{x}) = \mathbf{0}$. [Hint: first, suppose T is one-to-one. You want to prove that the zero vector is the only solution to $T(\mathbf{x}) = \mathbf{0}$. To that end, suppose $\mathbf{c}$ is a vector with $T(\mathbf{c}) = \mathbf{0}$. Now use the definition of one-to-one and Exercise 2O to prove that $\mathbf{c} = \mathbf{0}$. For the other direction, use the suggested strategy for proving a function is one-to-one, given above. Crucially, you'll need to use the fact that T is linear!]

Exercise 2Q. Can a linear transformation $\mathbf{R}^3 \to \mathbf{R}$ be one-to-one? Why or why not?

Example 2.18. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

and let $T(\mathbf{x}) = A\mathbf{x}$. Note that $T : \mathbf{R}^2 \longrightarrow \mathbf{R}^3$ and

$$T(x_1, x_2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix} = (x_1,\ x_2,\ x_2)\,.$$

We have

$$\operatorname{im} T = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\} = \text{ a plane in } \mathbf{R}^3.$$

This plane is pictured in Figure 2.4.

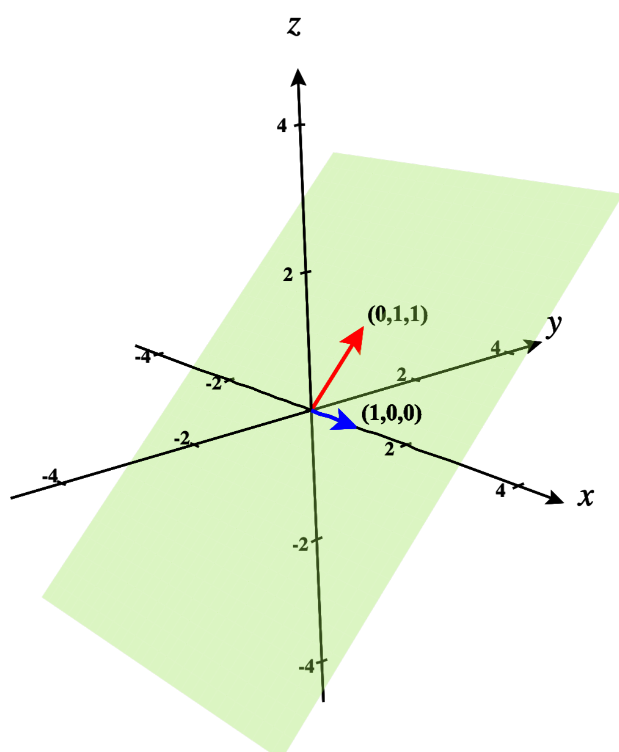


Figure 2.4: $\operatorname{im} T$ in Example 2.18

T is NOT onto because the range of T is not all of $\mathbf{R}^3$. T is one-to-one since $T(\mathbf{x}) = \mathbf{0}$ has ONLY the trivial solution (check!).

Exercise 2R. Tobias is interested in a particular function $f\colon \mathbf{R}^2 \to \mathbf{R}^2$. The function f fixes the closed lower half plane (when $y \le 0$, $f(x, y) = (x, y)$), and it reflects the open upper half plane across the y axis (when $y > 0$, $f(x, y) = (-x, y)$). He wants to represent this function using a matrix. Why is he doomed to failure?

Exercise 2S. Let S be a set of vectors in $\mathbf{R}^k$, and let $T : \mathbf{R}^k \to \mathbf{R}^n$ be a linear transformation.

(1) If S is linearly independent, must $T(S)$ be linearly independent?
(2) If S is linearly dependent, must $T(S)$ be linearly dependent?
(3) If $T(S)$ is linearly independent, must S be linearly independent?
(4) If $T(S)$ is linearly dependent, must S be linearly dependent?

Exercise 2T. Suppose the set of vectors $S = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is linearly independent. Decide whether each set below is linearly independent. Justify your answers.

(1) $S_1 = \{\mathbf{a} - 2\mathbf{b}, -\mathbf{c} + 4\mathbf{b}, \mathbf{c} - 2\mathbf{a}\}$
(2) $S_2 = \{\mathbf{a} - 2\mathbf{b}, -\mathbf{c} + 4\mathbf{b}, \mathbf{c} + 2\mathbf{a}\}$

Exercise 2U. Suppose $T : \mathbf{R}^3 \to \mathbf{R}^3$ is a linear transformation where

$$T(\mathbf{e}_1 - \mathbf{e}_2) = (-2, 4, -1)$$
$$T(\mathbf{e}_2 + \mathbf{e}_3) = (7, -2, 2)$$
$$T(2\mathbf{e}_1 + \mathbf{e}_3) = (6, 4, 1).$$

Find the matrix representation for T and cite the appropriate theorem from Chapter 2. Is T one-to-one? Justify your answer using a key theorem from Chapter 2. If the answer is no, give a specific nonzero vector such $\mathbf{x}$ such that $T(\mathbf{x}) = \mathbf{0}$.

Chapter 3 How to solve linear equations (and related problems)

Key concepts

- Row operations
- Row echelon forms of a matrix (REF)
- The reduced row echelon form of a matrix (RREF)
- The row reduction algorthim (to find a REF and the RREF)
- Pivot positions, pivot variables, and free variables
- The parametric vector form (PVF) of the solution set to $A\mathbf{x} = \mathbf{b}$
- Homogeneous vs. inhomogeneous solution sets
- The kernel of a matrix or linear transformation
- The PVF of the solution set for $A\mathbf{x} = \mathbf{0}$ *is* the kernel

Summary. Section 3.4 in this chapter is a pretty good summary of the main ideas; read The Concept Connector's Theorem 3.13 (the Onto Theorem) and Theorem 3.14 (the One-to-one Theorem). Row reduction and pivot positions provide a concrete, computational way to answer a lot of questions about linear systems and linear transformations.

Given a linear system $A\mathbf{x} = \mathbf{b}$, a REF of the augmented matrix $M = [A\ \mathbf{b}]$ gives you qualitative information about the solution set. It's consistent if and only if M has no pivot in the last column. When it's consistent, pivot columns in A correspond to pivot variables and non-pivot columns correspond to free variables. A consistent linear system has a unique solution exactly when there are no free variables (a pivot in every column of A). You can use the RREF of M (when the system is consistent) to write down a nice description of the solution set in "parametric vector form". When you do so, you are expressing the solution set as the span of a set of linearly independent vectors. Concerning the related linear transformation $T(\mathbf{x}) = A\mathbf{x}$, T is one-to-one if and only if A has a pivot in every column, and T is onto if and only if T has a pivot in every row.

Chapter 3

Up to this point, most of our examples have been relatively simple, usually involving vectors in $\mathbf{R}^2$ or $\mathbf{R}^3$. Whenever we've had to solve equations, we have been able to do so with high-school-type substitution. In this chapter, we will introduce the row reduction algorithm, a rubber-meets-the-road method for solving linear systems of any size. Along the way, we will answer a lot of questions about spans, linear independence, linear transformations, and vector equations.

Warm up: two equations, two variables §3.1

There are a lot of definitions to learn in this paragraph!

Let's return to the matrix-vector equation

$$A\mathbf{x} = \mathbf{b}.$$

We will sometimes refer to this equation as a **system of linear equations** (SLE) or just a **linear system** for short. It is **consistent** if it has a solution, and **inconsistent** otherwise. The matrix A is called the **coefficient matrix** and the vector $\mathbf{b}$ is called the **augmentation vector**. The **augmented matrix** associated to this equation is the matrix $[A\ \mathbf{b}]$, and the last column of this matrix is called the **augmentation column**. The set (perhaps empty) of all solutions $\mathbf{x}$ is called the **solution set**.

Reading Question 3A. Write down the system of linear equations that corresponds to the augmented matrix

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

as a matrix-vector equation and as a system of individual linear equations.

RQ

We know that $A\mathbf{x} = \mathbf{b}$ might be inconsistent (when $\mathbf{b}$ is not in the span of the columns of A), might have a unique solution (when the equation is consistent and the columns of A are linearly independent), or might have infinitely many solutions (when the equation is consistent and the columns of A are linearly dependent). Let's look at these three possibilities when A is 2×2. For such a matrix, the corresponding linear system will have two equations and two variables. Each of these two equations will look like this:

$$ax_1 + bx_2 = c;$$

the solution set of this single equation is a line in the plane (let's assume a and b are not both zero). For our system of two equations, there are three possibilities:

I. the lines might intersect in a single point;

II. the lines might not intersect at all; or,

III. the lines might be the same.

Let's look at an example that illustrates each case. In all the examples that follow, we write our linear systems on the left and the corresponding augmented matrices (which contain all the same information) on the right.

I.

$$\begin{aligned} x_1 - 2\,x_2 &= 1 \\ 2\,x_1 + x_2 &= 1 \end{aligned} \qquad \left[\begin{array}{ccc} 1 & -2 & 1 \\ 2 & 1 & 1 \end{array}\right]$$

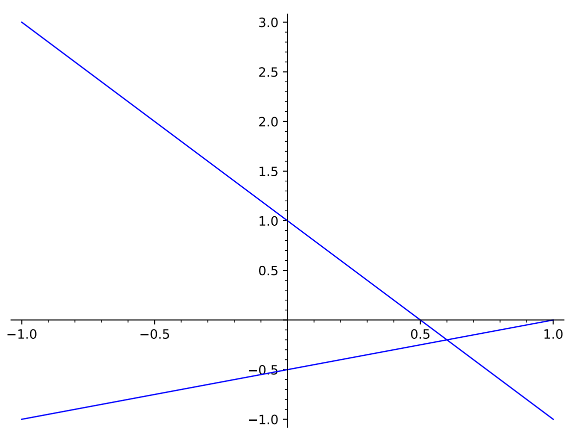


Figure 3.1: Lines intersecting in a single point

Our objective is to alter our system in such a way that we *don't* change the solution set, but *do* make the solution set easier to identify. There are a lot of things we can do to our system without changing the solution set. For example, we could scale either equation (let's call the top equation R_1 for "row 1" and the bottom equation R_2 for "row 2") by a nonzero value; we denote the process of scaling the ith equation by $c \neq 0$ by

$$R_i \longmapsto cR_i.$$

We could also add a multiple of one equation to another: we write

$$R_1 \longmapsto R_1 + cR_2$$

to denote what we get if we add cR_2 to R_1 and leave R_2 alone. We will discuss in more detail later why these operations don't change the solution set; but roughly speaking the reason is that they are "reversible" so no information is lost when we perform them.

Let's use these operations to solve the above linear system. Our goal is to isolate the variables. The first time you track though what's done below, focus on the equations. Then read it again and notice that the matrices on the right keep track of all the moves we made, ignoring superfluous information (variable names, equals signs, etc).

$$\begin{aligned} x_1 - 2x_2 &= 1 \\ 2x_1 + x_2 &= 1 \end{aligned} \qquad \qquad \begin{bmatrix} 1 & -2 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

$$R_1 \longmapsto -2R_1$$

$$\begin{aligned} -2x_1 + 4x_2 &= -2 \\ 2x_1 + x_2 &= 1 \end{aligned} \qquad \qquad \begin{bmatrix} -2 & 4 & -2 \\ 2 & 1 & 1 \end{bmatrix}$$

$$R_2 \longmapsto R_2 + R_1$$

$$\begin{aligned} -2x_1 + 4x_2 &= -2 \\ 5x_2 &= -1 \end{aligned} \qquad \qquad \begin{bmatrix} -2 & 4 & -2 \\ 0 & 5 & -1 \end{bmatrix} \text{(REF)}$$

$$R_2 \longmapsto (1/5)R_2$$

$$\begin{aligned} -2x_1 + 4x_2 &= -2 \\ x_2 &= -\tfrac{1}{5} \end{aligned} \qquad \qquad \begin{bmatrix} -2 & 4 & -2 \\ 0 & 1 & -\frac{1}{5} \end{bmatrix}$$

$$R_1 \longmapsto R_1 - 4R_2$$

$$\begin{aligned} -2x_1 &= -\tfrac{6}{5} \\ x_2 &= -\tfrac{1}{5} \end{aligned} \qquad \qquad \begin{bmatrix} -2 & 0 & -\frac{6}{5} \\ 0 & 1 & -\frac{1}{5} \end{bmatrix}$$

$$R_1 \longmapsto (-1/2)R_1$$

$$\begin{aligned} x_1 &= \tfrac{3}{5} \\ x_2 &= -\tfrac{1}{5} \end{aligned} \qquad \qquad \begin{bmatrix} 1 & 0 & \frac{3}{5} \\ 0 & 1 & -\frac{1}{5} \end{bmatrix} \text{(RREF)}$$

Above, we made a note of when the augmented matrix is first in "row echelon form" (REF) and "reduced row echelon form" (RREF); you'll learn more about these forms soon. The solution set for our linear system contains exactly one point: $\{(3/5, -1/5)\}$.

Let's use the same strategy to solve two more linear systems, illustrating cases II and III.

II.
$$\begin{aligned} x_1 - 2\,x_2 &= 1 \\ 2\,x_1 - 4\,x_2 &= 0 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 \\ 2 & -4 & 0 \end{bmatrix}$$

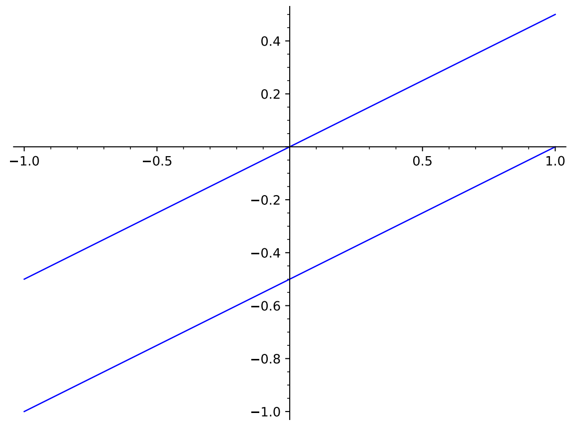


Figure 3.2: Non-intersecting lines

$$\begin{aligned} x_1 - 2\,x_2 &= 1 \\ 2\,x_1 - 4\,x_2 &= 0 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 \\ 2 & -4 & 0 \end{bmatrix}$$
$$R_2 \longmapsto (-1/2)R_2$$
$$\begin{aligned} x_1 - 2\,x_2 &= 1 \\ -x_1 + 2\,x_2 &= 0 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 \\ -1 & 2 & 0 \end{bmatrix}$$
$$R_2 \longmapsto R_2 + R_1$$
$$\begin{aligned} x_1 - 2\,x_2 &= 1 \\ 0 &= 1 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Look at that bottom row — there is *no* choice of x_1 and x_2 that makes $0 = 1$ true! We therefore conclude that, in this case, the solution set is the empty set $\emptyset$. Geometrically, the fact that the solution set is empty reflects the fact that the two lines determined by the two individual equations in our system have no points in common.

III.
$$\begin{aligned} x_1 - 2\,x_2 &= 1 \\ -7\,x_1 + 14\,x_2 &= -7 \end{aligned} \qquad \begin{bmatrix} 1 & -2 & 1 \\ -7 & 14 & -7 \end{bmatrix}$$

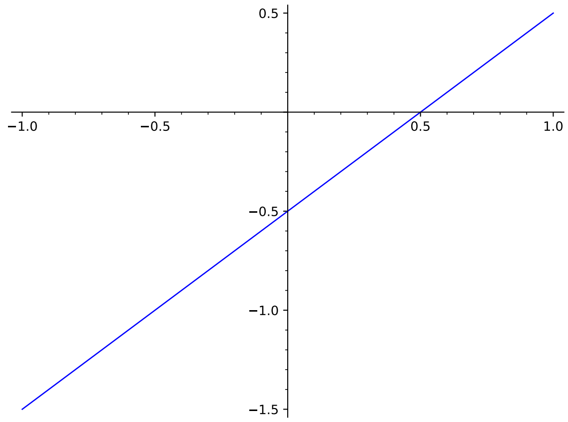


Figure 3.3: Coincident lines

$$\begin{aligned} x_1 - 2\,x_2 &= 1 \\ -7\,x_1 + 14\,x_2 &= -7 \end{aligned} \qquad \left[\begin{array}{ccc} 1 & -2 & 1 \\ -7 & 14 & -7 \end{array}\right]$$

$$R_2 \longmapsto (1/7)R_2$$

$$\begin{aligned} x_1 - 2\,x_2 &= 1 \\ -x_1 + 2\,x_2 &= -1 \end{aligned} \qquad \left[\begin{array}{ccc} 1 & -2 & 1 \\ -1 & 2 & -1 \end{array}\right]$$

$$R_2 \longmapsto R_2 + R_1$$

$$\begin{aligned} x_1 - 2\,x_2 &= 1 \\ 0 &= 0 \end{aligned} \qquad \left[\begin{array}{ccc} 1 & -2 & 1 \\ 0 & 0 & 0 \end{array}\right]$$

Look again at that bottom row: $0 = 0$ is true no matter what x_1 and x_2 are! In other words, that bottom row yields no information about the solution set. In System III, in fact, the second equation is really the first equation in disguise: the two individual equations determine *the same line*. Every point on this line is a solution to the system of equations. To find solutions, we may "freely" specify one of the variables — x_2, say — and then solve for the other. We may write the solution set like so: $\{(1 + 2x_2, x_2) \mid x_2 \in \mathbb{R}\}$. Using vector notation, the solution set is the set of vectors $\mathbf{x}$ of the form

$$\mathbf{x} = \left[\begin{array}{c} 1 \\ 0 \end{array}\right] + x_2 \left[\begin{array}{c} 2 \\ 1 \end{array}\right]$$

where x_2 is any real number. This is exactly the line $x_1 - 2x_2 = 1$. This solution set is infinite, but it takes only one parameter to describe it.

The above process is called row reduction. There is a row reduction algorithm for solving a linear system of any size.

Reading Question 3B. Write down the line $3x - 5y = 2$ in parametric vector form. Flip a

RQ

coin: if it comes up heads, solve for x and use y as the free variable; it it comes up tails, solve for y and use x as the free variable.

§3.2 Echelon forms and the row reduction algorithm

Let's keep thinking about the two row operations we used in the previous section and add one more, where we just swap a pair of rows (or a pair of equations).

What're the three row operations?

Row operations

Definition 3.1. Let A be an $n \times k$ matrix with rows $R_1, \dots, R_n$. Each of the following **row operations** modifies the matrix A to create a new matrix.

Scale. For any nonzero $c \in \mathbf{R}$ and row R_i, scale R_i by c.

Swap. Swap any pair of rows R_i and R_j.

Replace. For any $c \in \mathbf{R}$ and pair R_i, R_j, replace R_i with $R_i + cR_j$.

What're row equivalent matrices?

If a matrix B is obtained from a matrix A via any sequence of row operations, then we say A and B are **row equivalent** and write $A \sim B$.

RQ

Reading Question 3C. Row operations are reversible. For example, to undo $R_i \mapsto cR_i$, just scale the new row cR_i by $1/c$. How would you reverse $R_1 \mapsto R_1 + 7R_2$?

If A is row equivalent to B, then B is row equivalent to A.

It turns out that if we apply row operations to the augmented matrix of a linear system, we never change the underlying solution set. We will prove this later in §3.5; let's take it for granted for now (but at least do the following reading question so you believe it).

RQ

Reading Question 3D. Choose any system of linear equations and find any solution. Then perform a scale, a swap, and a replace on your system. After each operation, check that your solution is still valid.

As explained above, our goal in performing row operations is to arrive at a point where we can read off information about the solution set as readily as possible; one particular way to go about this is to put the matrix into the special form described in Definition 3.2 below. An entry in a matrix is called a **leading entry** if it is the first nonzero entry in its row (reading left to right). A row of all zeros does not have a leading entry.

What's a leading entry?

Row echelon forms

Definition 3.2. A matrix A is in **row echelon form** (REF) if the following conditions are satisfied.

What characteristics does a matrix in REF have? RREF?

① If a row contains only zeros, then so does every row below it.

② Each leading entry is to the right of any leading entry that appears in a row above it.

In REF, leading entries must stagger down and to the right.

A matrix A is said to be in **reduced row echelon form** (RREF) if it also satisfies the following additional conditions.

③ Each leading entry is a 1.

④ Each leading entry is the only nonzero entry in its column.

If a matrix is in REF, then the locations of the leading entries are called **pivot positions** or just **pivots**. For example, the pivot positions in the matrix

What's a pivot?

$$\begin{bmatrix} 0 & \boxed{1} & 2 & 0 & 4 \\ 0 & 0 & \boxed{5} & 6 & 7 \\ 0 & 0 & 0 & 0 & \boxed{8} \end{bmatrix}$$

(verify that this matrix is indeed in REF) are the locations $(1,2)$, $(2,3)$, and $(3,5)$. We'll soon discuss an algorithm for putting a matrix into a REF and into RREF; for now, we just want to show you the RREF of the above matrix:

Take care: "8" is not a pivot position. A pivot is a location, it's not the value of an entry.

$$\begin{bmatrix} 0 & 1 & 0 & -\frac{12}{5} & 0 \\ 0 & 0 & 1 & \frac{6}{5} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Make sure you believe that this matrix is in RREF.

Example 3.3 (2×2 REFs)**.** Let's write down all the possibilities for what a 2×2 matrix in REF can look like. We'll use ■ for an entry that can be any nonzero real number and $*$ for an entry that can be any real number (possibly even zero).

Suppose that A is a 2×2 matrix in REF. A can have exactly 0, 1, or 2 pivot positions; we'll consider each case.

If A has zero pivot positions, then it can't have any leading entries at all; A must be the zero matrix:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

If A has exactly 1 pivot position, then it has exactly 1 leading entry. So it has to have a row of all zeros at the bottom. The pivot position could be in either position on the top row, and you can have any real number after a pivot. There are two row echelon forms in this case:

$$\begin{bmatrix} \blacksquare & * \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & \blacksquare \\ 0 & 0 \end{bmatrix}.$$

Finally, A might have two pivot positions. Since it cannot have more than one pivot in any given row or column, and since pivot positions must stagger down and to the right, both pivot positions must be on the main diagonal. Thus,

The entries $A_{i,i}$ are the main diagonal entries of a matrix A.

$$\begin{bmatrix} \blacksquare & * \\ 0 & \blacksquare \end{bmatrix}.$$

is only REF in this case.

We have found that there are four basic forms that a 2×2 matrix in REF can have.

Exercise 3A. In the manner of the above example, find all possibilities for how a 3×4 matrix in row echelon form can look. Be sure to explain how you know your list is complete.

It turns out that any matrix A is row equivalent to REF matrices, and to an RREF matrix. In other words, A can always be converted into an REF or an RREF matrix with row operations (this process is called **row reduction**). We will not prove the following theorem, but after practicing with some row reductions you will definitely believe it.

RREF EXISTS AND IS UNIQUE

Theorem 3.4.

1. *Every matrix is row equivalent to a matrix in row echelon form.*
2. *If $A \sim B$ and $A \sim C$ and both B and C are in row echelon form, then B and C have exactly the same pivots.*
3. *Every matrix A is row equivalent to exactly one matrix in reduced row echelon form.*

If A is a matrix, then we call any REF matrix B that's row equivalent to A an REF of A. Since Theorem 3.4 tells us that A is row equivalent to a unique RREF matrix, we will call this RREF matrix *the* RREF of A.

For any matrix A, define the pivot positions of A to be the pivot positions in any REF for A. Theorem 3.4 assures us that we can use *any* REF matrix B that's row equivalent to A to find the pivot positions in A. Note that you usually can't tell where the pivots of A are just by looking at A, so you really do usually need to find an REF for A to find its pivot positions. Note also that, *by the rules of REF, any given row or column can have at most one pivot position. Thus, an $n \times k$ matrix has at most* $\min\{n, k\}$ *pivot positions.*

An $n \times k$ matrix has at most $\min\{n, k\}$ *pivot positions.*

In the following table and examples, we describe a step-by-step process for finding an REF for A and the RREF of A.

Row reduction algorithm

How do I put a matrix in REF or RREF?

To find an REF of a given matrix:

1. Swap any rows of all zeros to the bottom of the matrix.
2. Pick any row whose leading entry is furthest to the left and swap it to the top of the matrix.
3. Zero out every entry below the leading entry from the previous step using row replacements.
4. Now ignore the top row of the matrix and return to step 1.

You only need swaps and replacements to put a matrix in REF, but you can do scaling operations if you want to.

To find the RREF of a matrix already in REF:

5. For every leading entry c, scale the corresponding row by $1/c$.
6. For each leading entry, zero out every entry above the leading entry using row replacements.

This is a crucial example so study it slowly and carefully!

Example 3.5. Let's use row operations to solve a system of 4 linear equations in 4 variables. We will start by following the steps described above to obtain a REF of our augmented matrix.

System of Equations	Row Op	Augmented Matrix
$\begin{aligned} -4x_3 + 8x_4 &= 4 \\ 3x_1 + 9x_2 - 6x_3 + 12x_4 &= 9 \\ 0 &= 0 \\ 6x_1 + 18x_2 - 8x_3 + 16x_4 &= 14 \end{aligned}$		$\begin{bmatrix} 0 & 0 & -4 & 8 & 4 \\ 3 & 9 & -6 & 12 & 9 \\ 0 & 0 & 0 & 0 & 0 \\ 6 & 18 & -8 & 16 & 14 \end{bmatrix}$
	$R_3 \leftrightarrow R_4$	
$\begin{aligned} -4x_3 + 8x_4 &= 4 \\ 3x_1 + 9x_2 - 6x_3 + 12x_4 &= 9 \\ 6x_1 + 18x_2 - 8x_3 + 16x_4 &= 14 \\ 0 &= 0 \end{aligned}$		$\begin{bmatrix} 0 & 0 & -4 & 8 & 4 \\ 3 & 9 & -6 & 12 & 9 \\ 6 & 18 & -8 & 16 & 14 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

This completes Step 1 in the first pass of the row reduction algorithm.

$$R_1 \leftrightarrow R_2$$

$$\begin{aligned} 3\,x_1 + 9\,x_2 - 6\,x_3 + 12\,x_4 &= 9 \\ -4\,x_3 + 8\,x_4 &= 4 \\ 6\,x_1 + 18\,x_2 - 8\,x_3 + 16\,x_4 &= 14 \\ 0 &= 0 \end{aligned} \qquad \begin{bmatrix} 3 & 9 & -6 & 12 & 9 \\ 0 & 0 & -4 & 8 & 4 \\ 6 & 18 & -8 & 16 & 14 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This completes Step 2 in the first pass of the row reduction algorithm.

$$R_3 \mapsto R_3 - 2R_1$$

$$\begin{aligned} 3\,x_1 + 9\,x_2 - 6\,x_3 + 12\,x_4 &= 9 \\ -4\,x_3 + 8\,x_4 &= 4 \\ 4\,x_3 - 8\,x_4 &= -4 \\ 0 &= 0 \end{aligned} \qquad \begin{bmatrix} 3 & 9 & -6 & 12 & 9 \\ 0 & 0 & -4 & 8 & 4 \\ 0 & 0 & 4 & -8 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This completes Step 3 in the first pass of the row reduction algorithm.

$$R_3 \mapsto R_3 + R_2$$

$$\begin{aligned} 3\,x_1 + 9\,x_2 - 6\,x_3 + 12\,x_4 &= 9 \\ -4\,x_3 + 8\,x_4 &= 4 \\ 0 &= 0 \\ 0 &= 0 \end{aligned} \qquad \begin{bmatrix} 3 & 9 & -6 & 12 & 9 \\ 0 & 0 & -4 & 8 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Here, we restarted the process ignoring the top row. Steps 1 and 2 were unnecessary, so we just did Step 3.

At this point, the augmented matrix is in REF. It has exactly two pivot positions: $(1, 1)$ and $(2, 3)$. We know that the system, as currently written, has *exactly the same solution set* as our original system — but now the solution set is pretty easy to find! You can do this as follows. First, note that you can solve for x_3 in terms of x_4 in the second equation and then, using substitution into the first equation, you can solve for x_1 in terms of x_2 and x_4 (you should do this right now). So, x_2 and x_4 can take on any real values; they are called **free variables**. Once you pick values for x_2 and x_4, the values of x_1 and x_3 are determined; they are called **pivot variables**.

What're free variables and pivot variables?

It was easy to solve for x_3 in terms of free variables, but it was a little annoying to have to do a substitution before solving for x_1 in terms of free variables. If we go further and get our matrix into *reduced* row echelon form, this annoyance disappears.

$$R_1 \mapsto (1/3)R_1$$

$$\begin{aligned} x_1 + 3\,x_2 - 2\,x_3 + 4\,x_4 &= 3 \\ -4\,x_3 + 8\,x_4 &= 4 \\ 0 &= 0 \\ 0 &= 0 \end{aligned} \qquad \begin{bmatrix} 1 & 3 & -2 & 4 & 3 \\ 0 & 0 & -4 & 8 & 4 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$R_2 \mapsto (-1/4)R_2$$

$$\begin{aligned} x_1 + 3x_2 - 2x_3 + 4x_4 &= 3 \\ x_3 - 2x_4 &= -1 \\ 0 &= 0 \\ 0 &= 0 \end{aligned} \qquad \begin{bmatrix} 1 & 3 & -2 & 4 & 3 \\ 0 & 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

We just did Step 5 twice.

$$R_1 \mapsto R_1 + 2R_2$$

$$\begin{aligned} x_1 + 3x_2 &= 1 \\ x_3 - 2x_4 &= -1 \\ 0 &= 0 \\ 0 &= 0 \end{aligned} \qquad \begin{bmatrix} 1 & 3 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

One replacement in Step 6 brings us to RREF.

Now it's easy to read off the solution because we can solve for our pivot variables (x_1 and x_3) in terms of our free variables (x_2 and x_4) immediately:

$$\begin{cases} x_1 = 1 - 3x_2 \\ x_2 \text{ is free} \\ x_3 = -1 + 2x_4 \\ x_4 \text{ is free} \end{cases}.$$

Or, using set-bulider notation, the solution set is

$$\{(1 - 3x_2, x_2, -1 + 2x_4, x_4) \mid x_2, x_4, \in \mathbb{R}\}.$$

More significantly, we can write the solution set for this linear system using vectors. If $\mathbf{x}$ is a solution, then

Leave the free variables alone; rewrite pivot variables in terms of only free variables.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 - 3x_2 \\ x_2 \\ -1 + 2x_4 \\ x_4 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}}_{\mathbf{p}} + x_2 \underbrace{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{\mathbf{v}_1} + x_4 \underbrace{\begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}}_{\mathbf{v}_2}.$$

How do we find the PVF of the solution set to a consistent linear system, once we put the augmented matrix in RREF?

The equation above is the **parametric vector form** of the solution set. We can also write our solution set like this:

$$\mathbf{p} + \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\} = \{\mathbf{p} + \mathbf{v} \mid \mathbf{v} \in \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\}\}.$$

This connection between spans and solution sets will be developed later, in Theorem 3.9.

Assuming the linear system is consistent, non-pivot columns in the coefficient matrix will always correspond to free variables (columns 2 and 4 above), and pivot columns in the coefficient matrix will always correspond to pivot variables (columns 1 and 3 above).

Go back and read the example again.

Reading Question 3E. In Example 3.5, pick some specific values of x_2 and x_4 to write down a few specific solutions to the original linear system.

RQ

Exercise 3B. Suppose a linear system's *augmented* matrix has RREF

$$\begin{bmatrix} 0 & 1 & 0 & 7 & 0 & -1 & 2 \\ 0 & 0 & 1 & 2 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 & 9 \end{bmatrix}.$$

Find the solution set for the linear system in PVF. [Hint: since this is an augmented matrix, the coefficient matrix is 3×6. So, there are 6 variables $x_1, \ldots, x_6$. Which are free? Which are pivot variables? Convert the above augmented matrix into a system of linear equations, solve for the pivot variables in terms of the free variables, and then follow the process at the end of Example 3.5 to find the solution set in PVF.]

Example 3.6. Here's an example where there are no solutions. We used exactly the same row operations here as we did in Example 3.5, but we only needed the 2nd through 5th operations.

$$\begin{aligned} -4x_3 &= 8 \\ 3x_1 + 9x_2 - 6x_3 &= 12 \\ 6x_1 + 18x_2 - 8x_3 &= 0 \end{aligned} \qquad \begin{bmatrix} 0 & 0 & -4 & 8 \\ 3 & 9 & -6 & 12 \\ 6 & 18 & -8 & 0 \end{bmatrix}$$

$$\ldots \text{ row ops } \ldots$$

$$\begin{aligned} 3x_1 + 9x_2 - 14x_3 &= 28 \\ -4x_3 &= 8 \\ 0 &= -16 \end{aligned} \qquad \begin{bmatrix} 3 & 9 & -14 & 28 \\ 0 & 0 & -4 & 8 \\ 0 & 0 & 0 & -16 \end{bmatrix}$$

You can immediately see that the system of equations has no solutions because $0 = -16$ simply isn't true, no matter what the variable values are. It turns out that the augmented matrix of an inconsistent linear system, when reduced to REF, will always generate such an equation. *A linear system is inconsistent if and only if the last column of the augmented matrix is a pivot column.*

A system is inconsistent iff there's a pivot in the augmentation column.

Example 3.7. Let's do one where there's a unique solution.

$$\begin{aligned} 2x_2 + 4x_3 &= -1 \\ 2x_1 - 8x_3 &= 2 \\ -8x_1 + x_2 - x_3 &= -1 \end{aligned} \qquad \begin{bmatrix} 0 & 2 & 4 & -1 \\ 2 & 0 & -8 & 2 \\ -8 & 1 & -1 & -1 \end{bmatrix}$$

$$R_1 \leftrightarrow R_2$$

$$\begin{aligned} 2x_1 - 8x_3 &= 2 \\ 2x_2 + 4x_3 &= -1 \\ -8x_1 + x_2 - x_3 &= -1 \end{aligned} \qquad \begin{bmatrix} 2 & 0 & -8 & 2 \\ 0 & 2 & 4 & -1 \\ -8 & 1 & -1 & -1 \end{bmatrix}$$

$$R_3 \mapsto R_3 + 4R_1$$

$$\begin{aligned} 2x_1 - 8x_3 &= 2 \\ 2x_2 + 4x_3 &= -1 \\ x_2 - 33x_3 &= 7 \end{aligned} \qquad \begin{bmatrix} 2 & 0 & -8 & 2 \\ 0 & 2 & 4 & -1 \\ 0 & 1 & -33 & 7 \end{bmatrix}$$

$$R_3 \mapsto R_3 - (1/2)R_2$$

$$\begin{aligned} 2x_1 - 8x_3 &= 2 \\ 2x_2 + 4x_3 &= -1 \\ -35x_3 &= \tfrac{15}{2} \end{aligned} \qquad \begin{bmatrix} 2 & 0 & -8 & 2 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -35 & \tfrac{15}{2} \end{bmatrix}$$

$$R_1 \mapsto \tfrac{1}{2}R_1,\ R_2 \mapsto (1/2)R_2,\ R_3 \mapsto (-1/35)R_3$$

$$\begin{aligned} x_1 - 4x_3 &= 1 \\ x_2 + 2x_3 &= -\tfrac{1}{2} \\ x_3 &= -\tfrac{3}{14} \end{aligned} \qquad \begin{bmatrix} 1 & 0 & -4 & 1 \\ 0 & 1 & 2 & -\tfrac{1}{2} \\ 0 & 0 & 1 & -\tfrac{3}{14} \end{bmatrix}$$

$$R_2 \mapsto R_2 - 2R_3$$

$$\begin{aligned} x_1 - 4x_3 &= 1 \\ x_2 &= -\tfrac{1}{14} \\ x_3 &= -\tfrac{3}{14} \end{aligned} \qquad \begin{bmatrix} 1 & 0 & -4 & 1 \\ 0 & 1 & 0 & -\tfrac{1}{14} \\ 0 & 0 & 1 & -\tfrac{3}{14} \end{bmatrix}$$

$$R_1 \mapsto R_1 + 4R_3$$

$$\begin{aligned} x_1 &= \tfrac{1}{7} \\ x_2 &= -\tfrac{1}{14} \\ x_3 &= -\tfrac{3}{14} \end{aligned} \qquad \begin{bmatrix} 1 & 0 & 0 & \tfrac{1}{7} \\ 0 & 1 & 0 & -\tfrac{1}{14} \\ 0 & 0 & 1 & -\tfrac{3}{14} \end{bmatrix}$$

In a consistent system, there's a unique solution iff there's a pivot in every column of the coefficient matrix.

The solution set is $\{(1/7, -1/14, -3/14)\}$. Notice that the RREF of the augmented matrix in this example has a pivot position in every column of the coefficient matrix. *For a consistent linear system, there is a unique solution if and only if every column in the coefficient matrix is a pivot column.*

As we said in Example 3.5, in a consistent system free variables correspond to non-pivot columns in the coefficient matrix and pivot variables correspond to pivot columns in the coefficient matrix. A consistent linear system has a unique solution if and only if there are no free variables, which happens if and only if every column in the coefficient matrix is a pivot column. We summarize these observations in the next theorem.

Solution sets for linear systems

Theorem 3.8. *Consider the matrix-vector equation $A\mathbf{x} = \mathbf{b}$ with augmented matrix $[A\ \mathbf{b}]$.*

1. *The equation is inconsistent if and only if the augmentation column is a pivot column.*
2. *If A has a pivot position in every column and the equation is consistent, then it has a unique solution.*
3. *If A has a non-pivot column and the equation is consistent, then it has infinitely many solutions.*

You can check (1) using any REF of $[A\ \mathbf{b}]$.

You can check (2) and (3) using any REF of A.

Reading Question 3F. Put the matrix

$$A = \begin{bmatrix} 2 & 1 & 8 \\ -1 & 1 & -1 \\ -2 & 5 & 5 \end{bmatrix}$$

into RREF by hand.

RQ

Reading Question 3G. Each matrix below is the *augmented* matrix for a linear system. For each system, determine whether there is no solution, a unique solution, or infinitely many solutions.

RQ

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

Exercise 3C. Consider the linear system

$$\begin{aligned} 5x - 5ky &= h \\ -5x + 3y &= h - k. \end{aligned}$$

① Find all values of k and h so that the system has no solutions.
② Find all values of k and h so that the system has a unique solution.
③ Find all values of k and h so that the system has infinitely many solutions.

The preferred way to write down the solution set to a linear system is parametric vector form, described at the end of Example 3.5. In that example, if you replace the augmentation vector with the zero vector, then you will obtain that the solution set is just $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\}$. In general, there is a particularly nice relationship between the solution set for $A\mathbf{x} = \mathbf{b}$ and the solution set for the homogeneous linear system $A\mathbf{x} = \mathbf{0}$, which we record in the next theorem.

Remember, a solution need not exist, in which case this theorem says nothing about $A\mathbf{x} = \mathbf{b}$.

Homogeneous vs. inhomogeneous solution sets

Theorem 3.9. *Let $\mathbf{p}$ be any fixed solution to $A\mathbf{x} = \mathbf{b}$. Then, the entire solution set is*

$$\{\mathbf{p} + \mathbf{v} \mid \mathbf{v} \text{ is a solution to } A\mathbf{x} = \mathbf{0}\}.$$

Solutions to $A\mathbf{x} = \mathbf{b}$ = $\mathbf{p}$ + *Solutions to* $A\mathbf{x} = \mathbf{0}$

Proof. First we need to prove that any solution to $A\mathbf{x} = \mathbf{b}$ has the form $\mathbf{p} + \mathbf{v}$ for some solution $\mathbf{v}$ to $A\mathbf{x} = \mathbf{0}$. Then, we need to prove that if $\mathbf{v}$ is a solution to $A\mathbf{x} = \mathbf{0}$, then $\mathbf{p} + \mathbf{v}$ is a solution to $A\mathbf{x} = \mathbf{b}$.

Let $\mathbf{p}$ be any fixed solution to $A\mathbf{x} = \mathbf{b}$. Suppose $\mathbf{q}$ is some other solution to $A\mathbf{x} = \mathbf{b}$. Then,

$$A(\mathbf{q} - \mathbf{p}) = A\mathbf{q} - A\mathbf{p} = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

So, $\mathbf{v} = \mathbf{q} - \mathbf{p}$ is a solution to $A\mathbf{x} = \mathbf{0}$ and $\mathbf{q} = \mathbf{p} + \mathbf{v}$.

Conversely, if we take a solution $\mathbf{v}$ to $A\mathbf{x} = \mathbf{0}$, then

$$A(\mathbf{p} + \mathbf{v}) = A\mathbf{p} + A\mathbf{v} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

This completes the proof. ■

You can visualize the difference between the homogeneous and inhomogeneous solution sets in low dimensions. In Figure 3.4 below, the solution set to the homogeneous equation is the plane through the origin and the inhomogeneous equation's solution set is obtain by translating that plane to the parallel plane through a fixed solution vector $\mathbf{p}$.

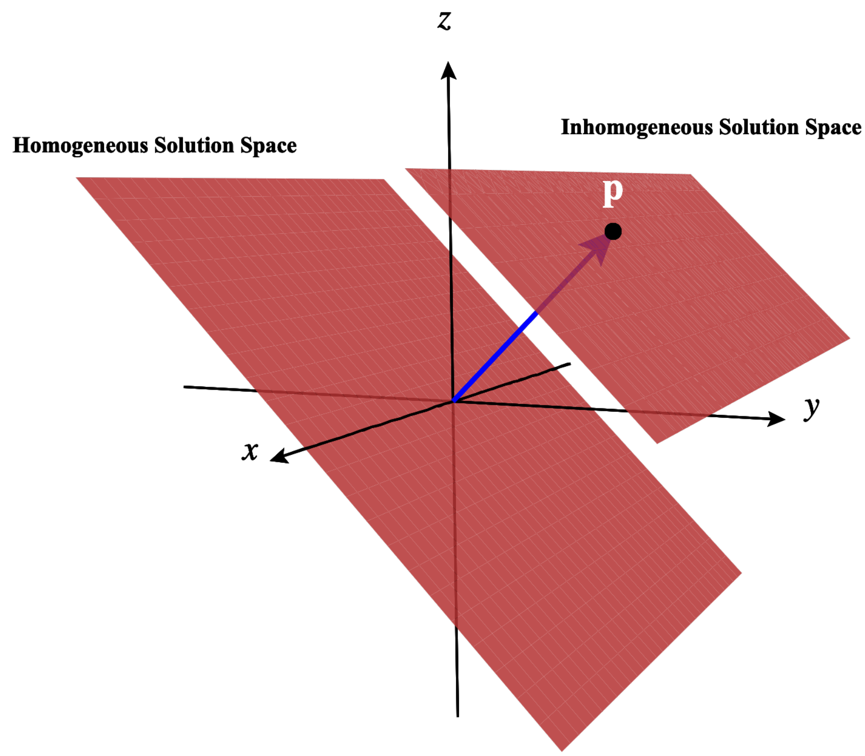


Figure 3.4: Homogeneous vs. inhomogeneous solution sets

Reading Question 3H. Suppose $\mathbf{v} = (1, 1, -2)$ is a solution to $A\mathbf{x} = \mathbf{0}$ and $\mathbf{w} = (4, 5, 6)$ is a solution to $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ is some nonzero vector. Find two new solutions $\mathbf{p}_1$ and $\mathbf{p}_2$ to $A\mathbf{x} = \mathbf{b}$ that are distinct from eachother and different than $\mathbf{w}$. RQ

Theorem 3.9 suggests that solution sets of homoegeneous equations are especially important. Here is some vocabulary.

THE KERNEL OF A MATRIX OR LINEAR TRANSFORMATION

What's the kernel of a matrix or linear transformation?

Definition 3.10. Let A be an $n \times k$ matrix. The set

$$\ker A = \{\mathbf{x} \in \mathbf{R}^k \mid A\mathbf{x} = \mathbf{0}\}$$

is called the **kernel** or **nullspace** of A. If $T\colon \mathbf{R}^k \to \mathbf{R}^n$ is the linear transformation defined by $T(\mathbf{x}) = A\mathbf{x}$, then we will also call this set the **kernel of** T and denote it by $\ker T$.

You should immediately observe that, by definition of kernel, the columns of A form a linearly dependent set if and only if $\ker A$ contains a nonzero vector.

Equivalently: the columns of A form a linearly independent set if and only if $\ker A = \{\mathbf{0}\}$.

RQ

Reading Question 3I. Find $\ker A$ in PVF, where

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

Then, suppose $(3, 1)$ is a solution to $A\mathbf{x} = \mathbf{b}$. Find $\mathbf{b}$. Write down the solution set to $A\mathbf{x} = \mathbf{b}$ in PVF (this last bit should be immediate—look back at Theorem 3.9).

The following theorem says that the parametric vector form of a solution set is nicely "efficient."

THE PARAMETRIC VECTOR FORM FOR $\ker A$

Theorem 3.11. *Suppose you solved $A\mathbf{x} = \mathbf{0}$ by computing the RREF of $[A\ \mathbf{0}]$ and writing down the PVF of the solution set:*

$$\mathbf{x} = x_{i_1}\mathbf{v}_1 + \cdots + x_{i_m}\mathbf{v}_m,$$

where $x_{i_1}, \ldots, x_{i_m}$ are the free variables. The set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ spans $\ker A$ and is linearly independent.

We will later see that S is a "basis" for $\ker A$.

Proof. By assumption, the vectors

$$\mathbf{x} = x_{i_1}\mathbf{v}_1 + \cdots + x_{i_m}\mathbf{v}_m,$$

where $x_{i_1}, \ldots, x_{i_m}$ vary over all real numbers, give the solution set to $A\mathbf{x} = \mathbf{0}$. This implies that S spans the kernel. To prove that S is linearly independent, we must show that the only solution to

$$x_{i_1}\mathbf{v}_1 + \cdots + x_{i_m}\mathbf{v}_m = \mathbf{0}$$

is $x_{i_1} = \cdots = x_{i_m} = \mathbf{0}$. Indeed, the left hand side of this equation is the solution vector $\mathbf{x}$ to $A\mathbf{x} = \mathbf{0}$, where the free variables appear in some of the entries. (Look again at Example 3.5.) Thus, if this vector $\mathbf{x}$ is zero, each free variable must be zero. ■

Exercise 3D. Each matrix below is also given in reduced row echelon form. Express the solution set to the associated homogeneous linear system $M_i\mathbf{x} = \mathbf{0}$ in parametric vector form. Describe the solution set geometrically.

(1) $M_1 = \begin{bmatrix} 2 & -1 & -5 \\ -3 & 1 & 6 \\ 1 & 0 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$

(2) $M_2 = \begin{bmatrix} 2 & 4 & -2 \\ -1 & -2 & 1 \\ -3 & -6 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(3) $M_3 = \begin{bmatrix} 3 & 8 & 1 \\ 2 & -5 & 6 \\ 0 & 1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(4) $M_4 = \begin{bmatrix} 0 & 2 & -8 \\ 0 & -1 & 4 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & -4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Exercise 3E. This is a follow-up to Exercise 3D. For $i = 1, 2$, solve the linear systems $M_i\mathbf{x} = \mathbf{w_i}$, where w_i is given below (express the solution set in parametric vector form). If I tell you that s_i (given below) is a solution, why does that make your job easier? Make sure you understand the difference, geometrically, between the solutions sets in Exercise 3D and the solution sets in this exercise.

(1) $w_1 = (-4,\ 4,\ 0)\,, s_1 = (1,\ 1,\ 1)$

(2) $w_2 = (2,\ -1,\ -3)\,, s_2 = (0,\ 1,\ 1)$

Exercise 3F. Consider the linear system with augmented matrix

$$\begin{bmatrix} 1 & -1 & 3 & a \\ 2 & 0 & 4 & b \end{bmatrix}.$$

Write down the two equations for the associated linear system. What can you say about the solution set for this linear system (your answer may or may not depend on the values of a and b)? In order to think qualitatively about the solution set for this linear system, you should put the augmented matrix into REF (RREF isn't necessary unless you want to get an explicit description of the solution set).

Exercise 3G. A linear system with more variables than equations is called **underdetermined**. Given an example of an inconsistent underdetermined system. Give an example of a consistent underdetermined system. It may help to pick some numbers here; how about three variables, and then either one or two equations?

Exercise 3H. If S is a linear system with n variables and m equations, where $n > m$, then what can you say about the solution set for the linear system, and why? Such a system may or may not be consistent. If it is consistent, then can you say anything more about the nature of the solution set?

Exercise 3I. What do you think **overdetermined** should mean? Give a consistent example.

§3.3 When the consistency of $A\mathbf{x} = \mathbf{b}$ does not depend on $\mathbf{b}$

Not all consistent systems are created exactly equal. Some systems are consistent only by virtue of the augmentation column. As a silly example, consider the systems

$$\begin{aligned} x_1 + x_2 &= 1 \\ 0x_1 + 0x_2 &= 0 \end{aligned} \quad \text{and} \quad \begin{aligned} x_1 + x_2 &= 1 \\ 0x_1 + 0x_2 &= 1 \end{aligned}$$

(they have the same coefficient matrix, but the first is consistent while the second is not). By contrast, the system

$$\begin{aligned} x_1 + x_2 &= b_1 \\ x_1 - x_2 &= b_2. \end{aligned}$$

is consistent no matter what b_1 and b_2 are, as you can check.

It turns out that when A has a pivot in every row, $A\mathbf{x} = \mathbf{b}$ will be consistent no matter what $\mathbf{b}$ is!

When the coefficient matrix has a pivot in every row

Theorem 3.12. *Let A be an $n \times k$ matrix. Then, A has a pivot position in every row if and only if $A\mathbf{x} = \mathbf{b}$ is consistent for all $\mathbf{b} \in \mathbf{R}^n$.*

Proof. Suppose A has a pivot position in every row and take any $\mathbf{b} \in \mathbf{R}^n$. The augmented matrix $[A\ \mathbf{b}]$ cannot have a pivot in the augmentation column because you can't have more than one pivot in any given row. Thus, $A\mathbf{x} = \mathbf{b}$ is consistent.

Conversely, suppose $A\mathbf{x} = \mathbf{b}$ is consistent for all $\mathbf{b} \in \mathbf{R}^n$. Let U denote the RREF of A. Then, by reversing the row operations from A to U and applying them to the matrix $[U\ \mathbf{e}_n]$, we obtain a row equivalent matrix $[A\ \mathbf{b}]$. We do not know what the vector $\mathbf{b}$ is, but since $A\mathbf{x} = \mathbf{b}$ is consistent for *any* $\mathbf{b}$, we know that $[U\ \mathbf{e}_n]$ corresponds to a consistent linear system (remember that row operations leave solution sets unchanged!). So, the bottom entry of the augmentation column of $[U \ \mathbf{e}_n]$ can't be a pivot position, meaning that U must have a pivot on the bottom row. This forces U (and hence A) to have a pivot in *every* row, because every row above the bottom row of U must also have a leading entry, and U is in RREF. ■

Reading Question 3J. Suppose A is an $n \times k$ matrix with $n > k$. Explain why there must be a $\mathbf{b} \in \mathbf{R}^n$ such that $A\mathbf{x} = \mathbf{b}$ is inconsistent. [Hint: think about pivot positions. In this case, what's the maximum number of pivot positions A can have?]

RQ

Exercise 3J. Consider the linear system

$$\begin{aligned} -3x_1 + 6x_2 &= a \\ 4x_1 - 8x_2 &= b. \end{aligned}$$

Describe the set of points (a, b) in the plane $\mathbf{R}^2$ such that the above linear system is consistent.

- We are interested in when the system is consistent, so you should find a REF for the augmented matrix.
- Stare at the REF you found. What must be true in order for the system to be consistent? This should lead you to an equation in a and b.
- Describe the equation you found geometrically.

Exercise 3K. Consider the linear system

$$\begin{aligned} x_1 + wx_2 &= b_1 \\ 2x_1 - 3x_2 &= b_2. \end{aligned}$$

Describe the numbers w for which this system is consistent no matter what b_1 and b_2 are.

Tying everything together §3.4

The concept that ties together everything we've learned so far is the linear transformation. Recall that a linear transformation is a function

$$T : \mathbf{R}^k \to \mathbf{R}^n$$

that satisfies Definition 2.2. This map has a matrix representation

$$A = [T(\mathbf{e}_1) \ T(\mathbf{e}_2) \ \cdots \ T(\mathbf{e}_k)]$$

that satisfies

$$T(\mathbf{x}) = A\mathbf{x}.$$

Conversely, given any $n \times k$ matrix A, the function $T(\mathbf{x}) = A\mathbf{x}$ is a linear transformation. Note that

$$A\mathbf{x} = \mathbf{b} \iff T(\mathbf{x}) = \mathbf{b}.$$

The following statements are equivalent.

- The equation $A\mathbf{x} = \mathbf{b}$ is consistent.
- The matrix $[A\ \mathbf{b}]$ has no pivot in the last column.
- The vector $\mathbf{b}$ is in the image of T.
- The vector $\mathbf{b}$ is in the span of the columns of A.

RQ

Reading Question 3K. Let T be a linear transformation with matrix representation

$$A = \begin{bmatrix} -6 & 5 & -27 \\ -6 & -1 & -9 \\ 4 & -9 & 35 \end{bmatrix}.$$

Use a computer to determine whether each of the following vectors is in the image of T: (i) $(-3, 2, 1)$; (ii) $(-28, -16, 30)$.

The summary below follows from Theorems 2.9, 2.17, 3.8, and 3.12.

The Concept Connector

Let $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_k]$ be an $n \times k$ matrix and let $T\colon \mathbf{R}^k \to \mathbf{R}^n$ be the linear transformation defined by $T(\mathbf{x}) = A\mathbf{x}$.

Theorem 3.13 (The Onto Theorem)**.** *The following are equivalent:*

1. *T is onto.*
2. *A has a pivot position in every row.*
3. *$A\mathbf{x} = \mathbf{b}$ is consistent for all $\mathbf{b} \in \mathbf{R}^n$.*
4. *The columns of A span $\mathbf{R}^n$.*

Theorem 3.14 (The One-to-One Theorem)**.** *The following are equivalent:*

1. *T is one-to-one.*
2. *A has a pivot position in every column.*
3. *$A\mathbf{x} = \mathbf{0}$ has a unique solution (the zero vector, $\mathbf{x} = \mathbf{0}$).*[a]
4. *The columns of A are linearly independent.*

[a] Or, we could say it this way: for any $\mathbf{b}$ in $\mathbf{R}^n$, $A\mathbf{x} = \mathbf{b}$ is either inconsistent or has a unique solution.

Theorems 3.13 and 3.14 place some strong restrictions on the domains and codomains of onto and one-to-one linear transformations, as you are asked to consider in the following very important exercise.

Exercise 3L. Let $T\colon \mathbf{R}^k \to \mathbf{R}^n$ be a linear mapping. If T is onto, what can you say about the relationship between k and n? What if T is one-to-one? What if it's both?

Exercise 3M. Suppose S is a set of k vectors in $\mathbf{R}^n$. Argue that if $k > n$ then S must be linearly dependent. (This is logically equivalent to the statement that if S is linearly independent, then $k \leq n$.)

Reading Question 3L. Suppose the columns of a 59×56 matrix form a linearly independent set. How many pivot positions does the matrix have? RQ

Reading Question 3M. Remember: each theorem in The Concept Connector above says that, for any T, either every statement is true or every statement is false. So it's important to RQ be able to logically negate each statement. As carefully as you can, write down the negation of Theorem 3.13 item (3).

Exercise 3N. Give an example of a 3×3 matrix, not in row echelon form, whose columns do not span $\mathbf{R}^3$.

Exercise 3O. Give an example of a 3×3 matrix, not in row echelon form, whose columns span $\mathbf{R}^3$.

Exercise 3P. Can a set of 3 vectors in $\mathbf{R}^4$ span $\mathbf{R}^4$?

Exercise 3Q. Must a set of 4 or more vectors in $\mathbf{R}^4$ necessarily span $\mathbf{R}^4$?

Exercise 3R. Suppose A is a square matrix and $\mathbf{b} \in \mathbf{R}^n$ is a vector such that $A\mathbf{x} = \mathbf{b}$ has a unique solution. What is the span of the columns of A? Given all this, if you take another vector $\mathbf{b}'$, what can you say about the equation $A\mathbf{x} = \mathbf{b}'$?

Exercise 3S. Find the value(s) of h such that the following vectors form a linearly independent set:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -2 \\ -9 \\ 6 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 3 \\ h \\ -9 \end{bmatrix}$$

Perhaps you can start by rephrasing this question as an equivalent problem that you already know how to solve.

Exercise 3T. Suppose T is a linear transformation from $\mathbf{R}^4$ to $\mathbf{R}^4$ that is onto. Argue that T must also be one-to-one.

Row operations do not change solution sets §3.5

In this section we fulfill our promise to show that row operations do not change solution sets. Our strategy is to first prove that if A and B are row equivalent, then the homogeneous equations $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have the same solution set. We will then use this to prove that if $[A\ \mathbf{b}]$ is row equivalent to $[B\ \mathbf{c}]$, then the equations $A\mathbf{x} = \mathbf{b}$ and $B\mathbf{x} = \mathbf{c}$ have the same solution set (no matter what the vectors $\mathbf{b}$ and $\mathbf{c}$ are).

The theorem below says that row equivalent matrices have the same kernel; hence, if A and B are row equivalent, then the columns of A form a linearly independent set if and only if the columns of B form a linearly independent set.

Row equivalent matrices have the same kernel

Theorem 3.15. *If A and B are row equivalent, then* $\ker A = \ker B$.

Proof. If two matrices are row equivalent, then there is a sequence of row operations from one to the other. Thus, to prove the theorem, it suffices to prove that each row operation leaves the kernel unchanged.

Let

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,k} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,k} \end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{bmatrix}$$

be an $n \times k$ matrix where $R_1, \ldots, R_n$ denote the rows of A. Each row is a $1 \times k$ matrix. Note that

$$A\mathbf{x} = \begin{bmatrix} x_1 a_{1,1} + x_2 a_{1,2} + \cdots + x_k a_{1,k} \\ x_1 a_{2,1} + x_2 a_{2,2} + \cdots + x_k a_{2,k} \\ \vdots \\ x_1 a_{n,1} + x_2 a_{n,2} + \cdots + x_k a_{n,k} \end{bmatrix} = \begin{bmatrix} R_1\mathbf{x} \\ R_2\mathbf{x} \\ \vdots \\ R_n\mathbf{x} \end{bmatrix}.$$

The formula given here is sometimes called the row-column rule for matrix-vector multiplication.

Hence, $\mathbf{x}$ is in the kernel of A if and only if $\mathbf{x}$ is in the kernel of R_i for each i. From this, it is clear that swapping the order of two rows will not change the kernel. For $c \neq 0$,

$$cR_i\mathbf{x} = \mathbf{0} \iff R_i\mathbf{x} = \mathbf{0}$$

(scale the first equation by $1/c$), so scaling a row will also not change the kernel. Finally, suppose we execute a replacement operation, replacing row R_i with $R_i + cR_j$ (leaving all other rows the same) to obtain a matrix B. If $\mathbf{x} \in \ker A$, then

To prove that two sets C and D are equal, prove that $x \in C$ implies $x \in D$ and $x \in D$ implies $x \in C$.

$R_i\mathbf{x} = R_j\mathbf{x} = \mathbf{0}$, and so $(R_i + cR_j)\mathbf{x} = \mathbf{0}$ which implies $\mathbf{x} \in \ker B$. On the other hand, if $\mathbf{x} \in \ker B$, then $(R_i + cR_j)\mathbf{x} = \mathbf{0}$ and $R_j\mathbf{x} = \mathbf{0}$ (because the jth row of B is the same as the jth row of A), from which it follows that $R_i\mathbf{x} = \mathbf{0}$ and hence $\mathbf{x} \in \ker A$. ■

Row operations preserve solution sets to SLEs

Theorem 3.16. *Suppose the linear systems $A\mathbf{x} = \mathbf{b}$ and $B\mathbf{x} = \mathbf{c}$ have row equivalent augmented matrices. Then, they have the same solution set.*

Proof. Given a vector $\mathbf{x} \in \mathbf{R}^k$, we will use the following notation:

$$\begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ -1 \end{bmatrix}.$$

Now you can check, using the definition of the matrix-vector product, that

$$[A\ \mathbf{b}] \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix} = A\mathbf{x} - \mathbf{b}.$$

Since $[A\ \mathbf{b}]$ is row equivalent to $[B\ \mathbf{c}]$, they have the same kernel, and so

$$\begin{aligned} A\mathbf{x} = \mathbf{b} &\iff A\mathbf{x} - \mathbf{b} = \mathbf{0} \\ &\iff [A\ \mathbf{b}] \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix} = \mathbf{0} \\ &\iff \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix} \in \ker[A\ \mathbf{b}] \\ &\iff \begin{bmatrix} \mathbf{x} \\ -1 \end{bmatrix} \in \ker[B\ \mathbf{c}] \\ &\iff B\mathbf{x} = \mathbf{c}. \end{aligned}$$

This completes the proof. ■

Exercise 3U (concept review)**.** Let $S = \{\mathbf{a}_1, \dots, \mathbf{a}_k\}$ be a set of k vectors in $\mathbf{R}^n$, let

$$A = [\mathbf{a}_1\ \cdots\ \mathbf{a}_k]$$

(note that A is $n \times k$), and let $T\colon \mathbf{R}^k \to \mathbf{R}^n$ be the linear transformation $T(\mathbf{x}) = A\mathbf{x}$. Explain how to use row reduction to answer each question (1) – (7) (for example, you might say that you need to form a particular matrix, put it in REF, and then do some sort of pivot-related analysis of the result). Then answer the remaining questions, and give reasons.

1. Is $A\mathbf{x} = \mathbf{b}$ consistent? If yes, does it have a unique solution? How do you find the set of all solutions $\mathbf{x}$?
2. Is $\mathbf{b} \in \operatorname{im} T$? If yes, how do you find $\mathbf{x}$ with $T(\mathbf{x}) = \mathbf{b}$?
3. Is $\mathbf{b} \in \operatorname{span} S$? If yes, how do you find the weights needed to build $\mathbf{b}$ as a linear combination of the vectors in S?
4. Does S span $\mathbf{R}^n$?
5. Is S linearly independent? If not, how do you find a specific dependence relation?
6. Is T one-to-one?
7. Is T onto?
8. How do we check whether a given vector $\mathbf{x}$ lies in $\ker A$? How do we find *all* vectors in $\ker A$?
9. Can 7 vectors in $\mathbf{R}^8$ span $\mathbf{R}^8$? Can such a set be linearly independent?
10. Can 7 vectors in $\mathbf{R}^6$ span $\mathbf{R}^6$? Can such a set be linearly independent?
11. If T is onto and $k = n$, must T be one-to-one?
12. If T is one-to-one and $k = n$, must T be onto?
13. If $k = 5$ and $n = 3$, can T be one-to-one? Can T be onto?
14. If $k = 3$ and $n = 5$, can T be one-to-one? Can T be onto?
15. Suppose A is 4×4. If $A\mathbf{x} = \mathbf{b}$ has a unique solution for some $\mathbf{b}$, must $A\mathbf{x} = \mathbf{b}$ have a solution for all $\mathbf{b}$?
16. Suppose A is 4×4. If $A\mathbf{x} = \mathbf{b}$ has a solution for some $\mathbf{b}$, must $A\mathbf{x} = \mathbf{b}$ have a solution for all $\mathbf{b}$?
17. Suppose A is 4×4. If $A\mathbf{x} = \mathbf{b}$ has a solution for all $\mathbf{b}$, must $A\mathbf{x} = \mathbf{b}$ have a unique solution for all $\mathbf{b}$?

Chapter 4 Matrix algebra

Key concepts

- Matrix operations: scalar multiplication, addition, multiplication, and transpose
- Properties that the key operations satisfy
- The relationship between the matrix product and function composition
- $M_{n,k}(\mathbf{R})$ as a set with a lot of algebraic structure, especially when $n = k$
- Examples of discrete dynamical systems
- Invertible matrices and their algebraic properties
- How to compute the matrix inverse
- The Isomorphism Theorem

Summary. Many applications of linear algebra require us to look not only at a single linear transformation, but at a composite of two or more linear transformations. This could be because you are combining geometric operations (to rotate, then scale, then ...). Or, you might be using a discrete dynamical system to model growing subpopulations of lionfish, where you apply the same transformation over and over again.

If T is a linear transformation with matrix representation A and S is a linear transformation with matirx representation B, then the matrix representation for $T \circ S$ is the matrix product AB. Though this product shares many features with the ordinary product of real numbers, it is also quite different. For example, AB and BA are usually not the same.

The Isomorphism Theorem is the theoretical highlight of this chapter; it unifies nearly every concept that we've discussed up to this point. It characterizes what it means for an $n \times n$ matrix A to be invertible, relating this concept to properties of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$.

Chapter 4

We already learned that rotating the plane counter-clockwise by $\pi/4$ radians about the origin is a linear transformation from $\mathbf{R}^2$ to itself with matrix representation

$$A = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

Here is another example of a geometric transformation that you met in Exercise 2B: reflection over the line $y = x$. This transformation maps $\mathbf{e}_1$ to $\mathbf{e}_2$ and vice-versa (draw pictures to check this). Thus, its matrix representation is

Go back and look at Theorem 2.5.

$$B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

What if we were to combine the two operations? For example, what is the matrix for the linear transformation that first reflects over the line $y = x$ and then rotates by $\pi/4$, and how is it related to the matrices A and B?

This question is about function composition. After all, if $T(\mathbf{x}) = A\mathbf{x}$ and $S(\mathbf{x}) = B\mathbf{x}$, then we are asking: what is the matrix representation of $T \circ S$ (this function first does S, then does T)? It turns out that it's the matrix product AB, which we will define and explore in the next section.

Function compostion acts right-to-left; you evaluate the inner function first!

Matrix operations §4.1

The motivation given in the introduction to this chapter will tell us how we *must* define the product of two matrices if we want it to be consistent with function composition. Let

$$S\colon \mathbf{R}^k \to \mathbf{R}^n$$

and

$$T\colon \mathbf{R}^n \to \mathbf{R}^p$$

be linear transformations. Note that we made the codomain of S the same as the domain of T. We did this so that the composite $T \circ S$ makes sense; it gives a linear transformation

$$T \circ S \colon \mathbf{R}^k \to \mathbf{R}^p.$$

RQ

Reading Question 4A. Verify that $T \circ S$ is a linear transformation. Use the fact that T and S are each linear, and start your computation like this:

$$(T \circ S)(s\mathbf{v} + t\mathbf{w}) = T(S(s\mathbf{v} + t\mathbf{w})) = \cdots$$

(where s and t are scalars and $\mathbf{v}$ and $\mathbf{w}$ are vectors).

If T has matrix representation A (a $p \times n$ matrix) and S has matrix representation

$$B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_k]$$

(an $n \times k$ matrix) then we want to define AB so that it is exactly the matrix representation of $T \circ S$. But we already know how to find the matrix representation of $T \circ S$. The columns are

$$(T \circ S)(\mathbf{e}_i) = T(S(\mathbf{e}_i)) = T(B\mathbf{e}_i) = T(\mathbf{b}_i) = A\mathbf{b}_i$$

for $i = 1, \ldots, k$. So, AB should be the $p \times k$ matrix whose ith column is $A\mathbf{b}_i$.

The matrix product

How's the matrix product defined?

Definition 4.1. Let A be a $p \times n$ matrix and let $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_k]$ be an $n \times k$ matrix. The matrix product AB is defined by

$$AB = [A\mathbf{b}_1 \ \cdots \ A\mathbf{b}_k].$$

The definition of the matrix product uses the matrix-vector product; since you know how to compute the matrix-vector product, you know how to compute the product of two matrices. However, there is a "row-column rule" for matrix multiplication that can be more convenient for by-hand computations. If we let $R_1, \ldots, R_p$ denote the rows of A and write

$$A = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_p \end{bmatrix},$$

then

$$AB = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_p \end{bmatrix} [\mathbf{b}_1 \ \cdots \ \mathbf{b}_k] = \begin{bmatrix} R_1\mathbf{b}_1 & \cdots & R_1\mathbf{b}_k \\ \vdots & \vdots & \vdots \\ R_p\mathbf{b}_1 & \cdots & R_p\mathbf{b}_k \end{bmatrix}.$$

You might want to look back at the proof of Theorem 3.15 here.

To multiply a row vector by a column vector, you multiply corresponding entries and add the result:

$$\begin{bmatrix} r_1 & r_2 & \cdots & r_n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = r_1c_1 + r_2c_2 + \cdots + r_nc_n.$$

Example 4.2. Let's multiply a 4×3 with a 3×2:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & -5 \\ -7 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 4 & -1 \\ -1 & 3 \\ 8 & -5 \end{bmatrix} = \begin{bmatrix} 1(4)+2(-1)+\ \ 3(8) & 1(-1)+2(3)+\ \ 3(-5) \\ 4(4)+0(-1)+-5(8) & 4(-1)+0(3)+-5(-5) \\ -7(4)+1(-1)+\ \ 2(8) & -7(-1)+1(3)+\ \ 2(-5) \\ 1(4)+1(-1)+\ \ 1(8) & 1(-1)+1(3)+\ \ 1(-5) \end{bmatrix}$$

$$= \begin{bmatrix} 26 & -10 \\ -24 & 21 \\ -13 & 0 \\ 11 & -3 \end{bmatrix}$$

Note that the result is 4×2, as we expect.

Reading Question 4B. Compute

$$\begin{bmatrix} 3 & -3 \\ 2 & 1 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 4 & -7 \\ 3 & 0 & -1 \end{bmatrix}$$

and then compute the product of these two matrices in the other order.

RQ

Example 4.3. Returning to the discussion at the start of the chapter, we can compute the matrix representation for the linear transformation that first reflects over the line $y = x$ and then rotates $\pi/4$ counter-clockwise about the origin using matrix multiplication:

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

You should also track what the transformation does to $\mathbf{e}_1$ and $\mathbf{e}_2$ to verify this another way.

What's the transpose of a matrix?

For any pair of positive integers n and k, we will denote the set of all $n \times k$ matrices with real entries by $\mathrm{M}_{n,k}(\mathbf{R})$. If $n = k$, we can just write $\mathrm{M}_n(\mathbf{R})$ for the set of $n \times n$ matrices. You can add matrices of the same dimensions by adding their corresponding entries, and you can scale a matrix by a real number by scaling its entries. The **transpose** of a matrix is an operation that swaps its columns and rows. The transpose of a vector is defined by

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}.$$

We can now transpose a matrix by

$$\begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_k \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{a}_k^{\mathrm{T}} \end{bmatrix}.$$

Example 4.4. A quick example of matrix transposition:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & -5 \\ -7 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} 1 & 4 & -7 & 1 \\ 2 & 0 & 1 & 1 \\ 3 & -5 & 2 & 1 \end{bmatrix}$$

Again, notice that what we're doing taking the columns of the original matrix and making them the rows of the new matrix.

Let's summarize the key matrix operations.

Key Matrix Operations

① Scalar multiplication	αA	A: any dimensions α: any real number scale each coordinate
② Addition	$A + B$	A, B: same dimensions add coordinate-wise
③ Multiplication	AB	$(p \times n) \cdot (n \times k)$ AB: $p \times k$ $AB = [A\mathbf{b}_1 \; \cdots \; A\mathbf{b}_k]$
④ Transpose	A^{T}	A: $n \times k$ A^{T}: $k \times n$ columns ↔ rows

What're the key matrix operations?

The theorem below summarizes all the properties of $\mathrm{M}_{n,k}(\mathbf{R})$ that don't involve matrix multiplication or the transpose. We'll write 0 to denote the zero matrix; its size will be clear from context. As you can see from the theorem, addition and scalar multiplication in $\mathrm{M}_{n,k}(\mathbf{R})$ and $\mathbf{R}^n$ are basically identical. This reflects the fact that both sets are examples of vector spaces (an object that we'll define later).

ALGEBRAIC PROPERTIES OF $M_{n,k}(\mathbf{R})$

Later, we'll see that these properties make $M_{n,k}(\mathbf{R})$ a vector space.

Theorem 4.5. *The following statements hold for all $A, B, C \in M_{n,k}(\mathbf{R})$ and all $s, t \in \mathbf{R}$.*

① *Matrix addition is commutative:*

$$A + B = B + A.$$

② *Matrix addition is associative:*

$$(A + B) + C = (A + B) + C.$$

③ *The zero matrix is the additive identity:*

$$A + 0 = 0 + A = A.$$

④ *Every matrix A has additive inverse $-A = -1 \cdot A$:*

$$A + (-A) = -A + A = 0.$$

⑤ *Scalar multiplication distributes over matrix addition:*

$$s(A + B) = sA + sB.$$

⑥ *Scalar multiplication distributes over real number addition:*

$$(s + t)A = sA + tA.$$

⑦ *The real number 1 acts as it should:*

$$1A = A.$$

⑧ *Scalar multiplication is associative:*

$$s(tA) = (st)A.$$

What's an identity matrix?

The next theorem gives the basic properties of matrix multiplication. Define the $m \times m$ **identity matrix** to be the $m \times m$ matrix with 1s down the main diagonal and zeros everywhere else:

$$I_m = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix} = [\mathbf{e}_1 \ \cdots \ \mathbf{e}_m].$$

Algebraic properties of the matrix product

Theorem 4.6. *Let A be an $n \times k$ matrix, let t be a scalar, and let B and C be any matrices such that the quantities below are defined.*

How does matrix multiplication interact with other operations?

(1) *Matrix multiplication is associative:*

$$(AB)C = A(BC).$$

(2) *Matrix multiplication is left distributive:*

$$A(B + C) = AB + AC.$$

(3) *Matrix multiplication is right distributive:*

$$(B + C)A = BA + CA$$

(4) *Matrix multiplication associates with scalar multiplication:*

$$t(AB) = (tA)B = A(tB).$$

(5) *The identity matrices satisfy*

$$I_n A = A = A I_k.$$

Reading Question 4C. Find two square matrices A, B such that $AB \neq BA$. [This says that matrix multiplication is not commutative.] Even so, can you find a specific 2×2 matrix A such that $AB = BA$ **does** hold for *any* 2×2 matrix B? How many such matrices A can you find? RQ

Reading Question 4D. Find two nonzero (meaning, not every entry is zero) square matrices A and B such that $AB = 0$. [Such matrices are called zero-divisors.] RQ

Reading Question 4E. Let A, B be square matrices of the same dimension. Simplify $(A + B)(A - B)$. How does what you found compare to what's true for real numbers? RQ

How does the transpose interact with other operations?

Algebraic properties of the matrix transpose

Theorem 4.7. *Let t be a scalar and let A and B be any matrices such that the quantities below are defined.*

① *The transpose distributes over addition:*

$$(A+B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}.$$

② *The transpose commutes with scalar multiplication:*

$$(tA)^{\mathrm{T}} = tA^{\mathrm{T}}.$$

③ *The transpose of the transpose is the original:*

$$(A^{\mathrm{T}})^{\mathrm{T}} = A.$$

④ *The transpose is product-reversing:*

$$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}.$$

Exercise 4A. Prove Theorem 4.7 ④ in the following steps (you can use the first three properties).

① Prove that if A is $n \times k$ and $\mathbf{x}$ is a k-vector, then $(A\mathbf{x})^{\mathrm{T}} = \mathbf{x}^{\mathrm{T}}A^{\mathrm{T}}$.

② Prove that if C is a $p \times k$ matrix and M is $k \times n$, where

$$C = \begin{bmatrix} R_1 \\ \vdots \\ R_p \end{bmatrix},$$

then

$$CM = \begin{bmatrix} R_1M \\ \vdots \\ R_pM \end{bmatrix}.$$

③ Now let A be $n \times k$ and let $B = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_p]$ be $k \times p$. Prove that $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.

Exercise 4B. Let A be an $n \times m$ matrix and let $\mathbf{x} = [x_1 \ldots x_n]$ (so, $\mathbf{x}$ is a row vector). Use the transpose to explain why finding solutions to $\mathbf{x}A = [0 \ldots 0]$ is the same as finding solutions to a more familiar homogeneous linear system of the form $B\mathbf{y} = \mathbf{0}$.

Discrete dynamical systems §4.2

Discrete dynamical system

Definition 4.8. A **discrete dynamical system** is any function $f : X \to X$ that maps a set X to itself. Given any x_0 in X, the recursively defined sequence

$$\begin{aligned} &x_0 \\ &x_1 = f(x_0) \\ &x_2 = f(x_1) = f^2(x_0) \\ &\quad\vdots \\ &x_k = f(x_{k-1}) = f^k(x_0) \\ &\quad\vdots \end{aligned}$$

is called the **solution of the dynamical system with initial condition x_0.** (Here, note carefully that f^k means f *composed* with itself k times, not multiplied by itself k times.)

Example 4.9. Consider the discrete dynamical system $f : \mathbf{R} \to \mathbf{R}$, where $f(x) = x^2$. The solution with initial condition 2 is

$$(2, 4, 16, 256, \ldots);$$

the solution with initial condition −1 is

$$(-1, 1, 1, 1, \ldots);$$

and the solution with initial condition 0.5 is

$$(0.5, 0.25, 0.0625, 0.00390625, \ldots).$$

Several remarks about this definition are in order.

- It may seem odd at first that a “solution” of a dynamical system is a *sequence*. In mathematics, a “solution” is any object that makes a specified relationship hold: so in basic high-school algebra solutions are numbers; in linear algebra solutions are vectors; in ordinary differential equations solutions are functions; and in discrete dynamical systems solutions are sequences.

- It may also seem odd that we are giving a new, fancy name (*discrete dynamical system!*) to a very familar object: a function from a set to itself. When we call a function $f : X \to X$ a "discrete dynamical system," we are announcing that our primary interest is in how the solutions $(x_0, x_1, x_2, \ldots)$ behave.
- We use dynamical systems to model change over time: so if x_0 is the current value of some quantity, we assume that $x_k = f^k(x_0)$ is the value after k time periods. The adjective "discrete" comes from the fact that we envision time as moving in discrete steps. (In *continuous dynamical systems* we image time moving continuously; the best-known continuous dynamical systems are differential equations.) When no confusion can arise we will often drop the adjective "discrete" and just talk about "dynamical systems."

Many dynamical systems take the form of linear transformations $T : \mathbf{R}^n \to \mathbf{R}^n$; so for such dynamical systems, we are interested in applying powers of T to an initial vector $\mathbf{v}$:

$$(\mathbf{v}, T(\mathbf{v}), T^2(\mathbf{v}), \ldots).$$

One of the great triumphs of linear algebra is that these dynamical systems are very well understood: we can say *a lot* about the powers $T^k(\mathbf{v})$. We will see this quite vividly over the next few chapters.

Below we briefly present three linear dynamical systems.

RQ

Reading Question 4F. Consider the dynamical system $T : \mathbf{R}^2 \to \mathbf{R}^2$ given by the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Find the solution with initial condition $(2, 1)$. (Find enough points of the solution that you fully understand how it behaves.)

---

§4.2.1 Lionfish again

In Example 1.2, we introduced a model for lionfish populations. In Example 2.7, we recognized that this model could be viewed as saying that the vector (L', J', A') of larvae, juveniles, and adults one month in the future is a linear transformation of the current population vector (L, J, A):

$$\begin{bmatrix} L' \\ J' \\ A' \end{bmatrix} = M \begin{bmatrix} L \\ J \\ A \end{bmatrix}, \quad \text{where } M = \begin{bmatrix} 0 & 0 & 35315 \\ 0.00003 & 0.777 & 0 \\ 0 & 0.071 & 0.949 \end{bmatrix}.$$

Suppose the current population is (L_0, J_0, A_0) and we want to predict the population (L_k, J_k, A_k) k months from now. We do this by beginning with our vector (L_0, J_0, A_0) and applying our linear transformation k times:

$$\begin{bmatrix} L_k \\ J_k \\ A_k \end{bmatrix} = \underbrace{T \circ \cdots \circ T}_{k \text{ times}} \begin{bmatrix} L_0 \\ J_0 \\ A_0 \end{bmatrix}.$$

Since we have defined the matrix product to agree with composition of linear transformations, we also have the formula

$$\begin{bmatrix} L_k \\ J_k \\ A_k \end{bmatrix} = M^k \begin{bmatrix} L_0 \\ J_0 \\ A_0 \end{bmatrix}.$$

We now recongize our lionfish model as a linear dynamical system.

Given (L_0, J_0, A_0), we can easily have a computer use the above formula to compute and plot the sequences (L_k), (J_k), and (A_k) for us. We have done this in Figure 4.1 below, assuming initial values of 1 unit for each of L_0, J_0, and A_0. (Since the numbers of larvae are so much larger than the numbers of juveniles and adults, we have actually plotted the natural logs of the population components so they can be viewed on the same scale.) Note carefully that, in the plot, the actual population values are given by the dots; the dots are joined by lines to make the movement over time easier to see, but the lines themselves do not represent predicted population values.

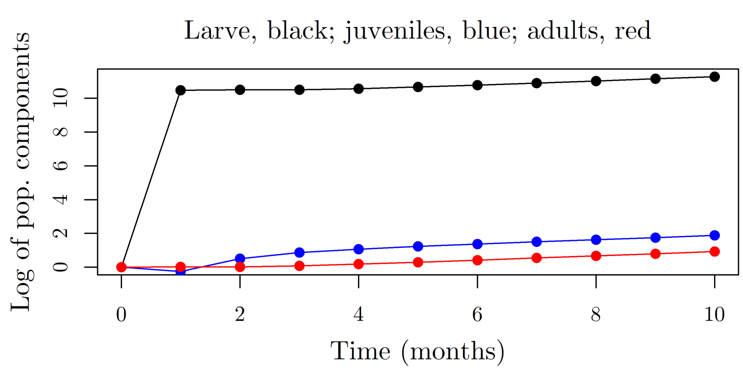


Figure 4.1: Log of subpopulations vs. time

It turns out that the situation is much better than our merely being able to recognize that an interesting quantity is given by a matrix product. We can actually write down an (approximate) explicit formula for the matrix M^k that gives us full insight into why the populations behave the way they do in the long term. We will learn how to do this later.

RQ **Reading Question 4G.** Using a computer, calculate M^5, where M is the lionfish model matrix given above. If the current population vector is $(3, 2, 1)$, what is the predicted population vector five months from now?

§4.2.2 *The Fibonacci sequence*

The Fibonacci sequence

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots$$

is defined by the following recursive formula:

$$F_0 = 0,\ F_1 = 1,\ \text{and } F_n = F_{n-1} + F_{n-2} \text{ for } n \geq 2.$$

The Fibonacci sequence has been studied all over the world for more than two millennia. It comes up often and unexpectedly in pure mathematics, but it also arises in other fields like biology and computer science. There is an entire journal devoted to it — *The Fibonacci Quarterly* — where new discoveries continue to emerge.

Now, consider the linear dynamical system $T : \mathbf{R}^2 \to \mathbf{R}^2$ given by the following matrix:

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

As you can (and should) check, if we start with the initial condition $\mathbf{v}_0 = (1, 0)$ we get the following solution:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \ldots.$$

More generally, we have the formula

$$\mathbf{v}_k = M^k \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}.$$

In other words, the second entry of the kth vector in our solution is just the kth Fibonacci number!

Now, as with the lionfish, the usefulness of looking at the Fibonacci sequence this way goes way beyond an opportunity to appreciate an unexpected appearance of matrix algebra. If you are thinking about the Fibonacci sequence in the usual way and want to compute (say) $F_{98211345}$, you have to compute the 98211345 terms before it (or program a computer to do so). We repeat, though, that humankind has an excellent understanding of linear algebra — so excellent that, in later chapters, we will be able to write down an explicit formula for M^k (and hence for F_k).

Reading Question 4H. Since $\mathbf{v}_0$ above is just $\mathbf{e}_1$, $M^n\mathbf{v}_0$ is just the first column of M^n. Compute M^n for $n = 1, 2, 3, 4$ to verify this. RQ

Reading Question 4I. Compute the first few terms of some "Fibonacci-like" sequences by looking at solutions of the dynamical system $\mathbf{x} \mapsto M\mathbf{x}$ with different initial conditions. Initial conditions involing entries other than non-negative integers are highly encouraged. RQ

Directed graphs §4.2.3

A **directed graph** G is a collection of **vertices** $\{1, \ldots, n\}$ and a collection of **edges** (i.e. arrows) connecting the vertices. The term *directed graph* is often abbreviated *digraph*.

Here is a digraph with 3 vertices and 5 edges.

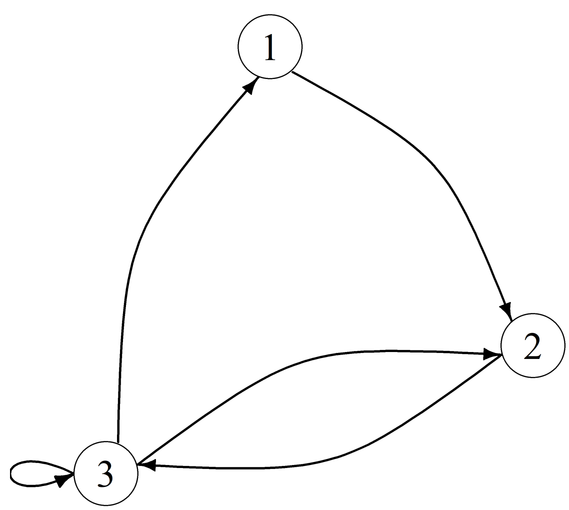


Figure 4.2: Digraph with 3 vertices and 5 edges

We use digraphs to model networks of various kinds: for example, cities with roads between them, or people with communication channels between them. We are interested in the ways that it is possible to move around the network.

RQ **Reading Question 4J.** In the digraph above, how many different ways (maybe none) are there to move, following the arrows,

- from vertex 2 to vertex 1 in exactly one step?
- from vertex 2 to vertex 2 in exactly one step?
- from vertex 2 to vertex 3 in exactly one step?

Repeat the above for "exactly two steps" and "exactly three steps".

Suppose that G is a digraph with n vertices. The **adjacency matrix of** G is the $n \times n$ matrix A defined as follows:

Read this **very** *carefully.*

- A has a 1 in row i and column j if there is an arrow from j to i in the digraph;
- all the other entries of A are zero.

Here is the adjacency matrix of the digraph illustrated in Figure 4.2.

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

In most applications of linear algebra, matrices play one of two roles: they function either as matrices of linear transformations, or as data sets. Unlike the matrices that we studied in the last two examples, this matrix A definitely feels more like a data set than a linear transformation: it contains the same information about our network that the digraph does.

Nevertheless, it turns out that A *also* gives us information about the network if we view it as a linear transformation. For example, if we multiply the initial vector $\mathbf{e}_2 = (0, 1, 0)$ by powers of A, as you can (and should!) check we get

$$A \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad A^2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad A^3 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

Comparing these vectors to what we found in the last Reading Question, we see that the ith entry of $A^k\mathbf{e}_2$ seems to count the number of ways to go from vertex 2 to vertex i in exactly k steps! This turns out to be true. Moreover since the columns of A^k are just the vectors $A^k\mathbf{e}_j$, we have the following proposition.

Powers of adjacency matrices

Theorem 4.10. *Suppose that the $n \times n$ matrix A is the adjacency matrix for a digraph G. Then the $[i, j]$-entry of A^k is the number of distinct ways to move in G from vertex j to vertex i in exactly k steps.*

Exercise 4C. Prove Theorem 4.10.

Theorem 4.10 isn't much use if we are interested in analyzing short paths in small networks — it's quicker and easier to example the corresponding digraph by eye. Even for a modestly sized network like this one, though,

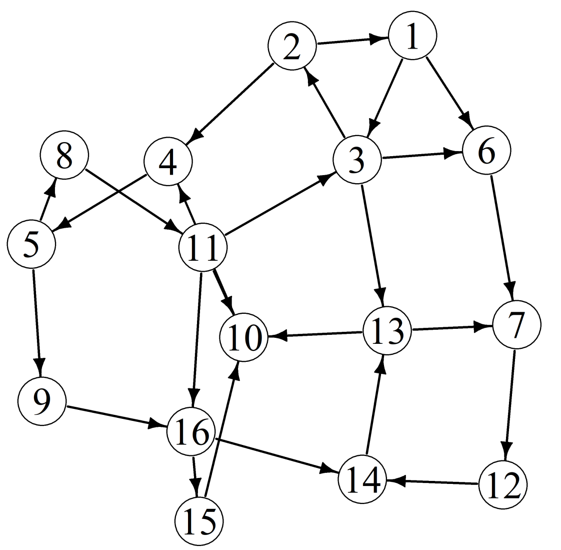


Figure 4.3: A modestly sized network

counting paths is much harder than asking a computer to find powers of the corresponding adjacency matrix.

Exercise 4D. Suppose that A is the adjacency matrix of a graph G.

(a) Describe the graph G^{T} that has adjacency matrix A^{T}.

(b) Describe what G looks like if A is **symmetric** — that is, if $A^{\mathrm{T}} = A$.

What's a symmetric matrix?

---

§4.3 Invertible matrices

At this point, you have probably noticed that matrix algebra has a lot in common with ordinary arithmetic in the real numbers: you can add, you can multiply, multiplication distributes over addition, etc. Just look back at Theorem 4.5; all the items in this theorem are true when A, B, and C are ordinarly real numbers.

But there are also striking differences! For example, you can only multiply matrices of appropriate dimensions, matrix multiplication is not commutative, and there are nonzero matrices that multiply to zero. Since matrix multiplication seems to be the operation that behaves the most differently, let's consider another point of comparison. Every nonzero real number has a multiplicative inverse: for any nonzero real number x, there is a real number y such that $xy = 1$. Is this true for matrices? It can't be: in Reading Question 4D, you found nonzero matrices A and B with $AB = 0$. If "A^{-1}" exists, and if it behaves at all reasonably, then we would expect something like:

$$\begin{aligned} AB &= 0 \\ A^{-1}(AB) &= A^{-1}(0) \\ (A^{-1}A)B &= 0 \\ B &= 0. \end{aligned}$$

But the B you found isn't zero!

To thicken the plot further, remember that a matrix A determines a linear transformation $T\colon \mathbf{R}^k \to \mathbf{R}^n$ defined by $T(\mathbf{x}) = A\mathbf{x}$. Now, some functions have inverses. If $f\colon X \to Y$ is a function, then the **inverse** of f, if it exists, is the (unique) function $g\colon Y \to X$ such that

What's the inverse of a function?

$$g(f(x)) = x \text{ for all } x \in X$$

and

$$f(g(y)) = y \text{ for all } y \in Y.$$

In this situation, we write f^{-1} for the function g. It turns out (we'll prove this later on) that a function has an inverse if and only if it is both one-to-one and onto. If the linear transformation T above has an inverse, then Exercise 3L implies that $k = n$. So A is square. For this reason, when we define the inverse of a matrix, we will insist that the matrix be square.

RQ **Reading Question 4K.** Find the inverse of the function in Reading Question 2G.

Keeping in mind that matrix multiplication isn't commutative and the identity matrix plays the role of the real number 1, we have the following definition.

INVERTIBLE MATRIX

Definition 4.11. An $n \times n$ matrix A is **invertible** if there is an $n \times n$ matrix B such that

$$AB = BA = I_n.$$

When A is invertible, we will write A^{-1} for the matrix B above.

What's an invertible matrix?

In the above definition, you might worry that the notation A^{-1} is ambiguous: what if there is more than one way to choose the matrix B? Well, there isn't, and you can prove it in the next exercise.

Exercise 4E. Suppose B and C are both inverses of the matrix A. Write down the equations that B and C must satisfy, and then try to argue that $B = C$.

When is a matrix invertible? If a matrix *is* invertible, how can we find its inverse? The next theorem answers these questions.

FINDING THE INVERSE OF A MATRIX

Theorem 4.12. *Let A be an $n \times n$ matrix. The matrix A is invertible if and only if A is row equivalent to the identity matrix I_n. If A is invertible, then the RREF of $[A\ I_n]$ is $[I_n\ A^{-1}]$.*

How do I find the inverse of a matrix?

Proof. Suppose A^{-1} exists. Then, for any $\mathbf{b} \in \mathbf{R}^n$, the vector $\mathbf{x} = A^{-1}\mathbf{b}$ is a solution to the equation $A\mathbf{x} = \mathbf{b}$:

$$\begin{aligned} A(A^{-1}\mathbf{b}) &= (AA^{-1})\mathbf{b} \\ &= I_n\mathbf{b} \\ &= \mathbf{b}. \end{aligned}$$

This implies that A has a pivot in every row, so A has n pivots. Since A is $n \times n$, it must also have a pivot in every column. By the rules of REF, this forces the pivot positions in A to be the main diagonal entries. In RREF, these entries must be 1s and all other entries (every non-diagonal entry is either above or below a pivot) must be 0s. Thus, the RREF of A is the identity matrix I_n.

Conversely, suppose A is row equivalent to I_n. Then, A has a pivot in every row, so $A\mathbf{x} = \mathbf{e}_i$ has a solution $\mathbf{x} = \mathbf{b}_i$ for $i = 1, \dots, n$. Let $B = [\mathbf{b}_1\ \cdots \mathbf{b}_n]$. We now have

$$AB = [A\mathbf{b}_1 \ \cdots \ A\mathbf{b}_n] = [\mathbf{e}_1 \ \cdots \ \mathbf{e}_n] = I_n.$$

So it looks like B must be the inverse of A. Unfortunately, we can't conclude this until we prove that $BA = I_n$.

First, let's prove that the columns of B are linearly independent. To do so, we can show that the zero vector is the only solution to $B\mathbf{x} = \mathbf{0}$. Well, if $\mathbf{v}$ is a solution to this equation, then

$$\begin{aligned} B\mathbf{v} = \mathbf{0} &\implies A(B\mathbf{v}) = A(\mathbf{0}) \\ &\implies (AB)\mathbf{v} = \mathbf{0} \\ &\implies I_n\mathbf{v} = \mathbf{0} \\ &\implies \mathbf{v} = \mathbf{0}. \end{aligned}$$

So indeed the only solution to $B\mathbf{x} = \mathbf{0}$ is the zero vector.

Since the columns of B are linearly independent, B has a pivot in every column, so since B is square it is row equivalent to I_n. Using *exactly the same argument above that we applied to* A, there must be an $n \times n$ matrix C such that $BC = I_n$. Now the fun begins:

$$\begin{aligned} BA &= BAI_n \\ &= BABC \\ &= BI_nC \\ &= BC \\ &= I_n. \end{aligned}$$

This completes the proof that A is invertible.

Recall that above, $A\mathbf{b}_i = \mathbf{e}_i$. So, the RREF of $[A \ \mathbf{e}_i]$ is $[I_n \ \mathbf{b}_i]$. From this it follows that if you augment A with all the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ to obtain $[A \ I_n]$, then the RREF of this matrix will be

$$[I_n \ \mathbf{b}_1 \ \cdots \mathbf{b}_n] = [I_n \ B].$$

Since $B = A^{-1}$, the proof is complete. ■

RQ

Reading Question 4L. At the start of the proof of Theorem 4.12, we used the fact that A^{-1} exists to prove that, for any $\mathbf{b}$, $A\mathbf{x} = \mathbf{b}$ has a solution (and thus A has a pivot in every row). Use a similarly direct argument to show that if A^{-1} exists, then the only solution to $A\mathbf{x} = \mathbf{0}$ is the zero vector (and thus A has a pivot in every column).

Exercise 4F. Using a computer, but only the command for putting a matrix into RREF, compute the inverse of each of the 3×3 matrices

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix}$$

(if the inverse exists).

Example 4.13 (invertible 2×2 matrices)**.** Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

be an invertible 2×2 matrix. Since A is row equivalent to the identity, its first column can't be zero, so either $a \neq 0$ or $c \neq 0$. Let's assume $a \neq 0$; the argument is similar in the other case.

Let's compute the RREF of

$$[A\ I_2] = \begin{bmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{bmatrix}.$$

First, do the replacement $R_2 \mapsto R_2 - (c/a)R_1$ to obtain

$$\begin{bmatrix} a & b & 1 & 0 \\ 0 & d - bc/a & -c/a & 1 \end{bmatrix}.$$

To make this look a bit nicer, do $R_2 \mapsto aR_2$:

$$\begin{bmatrix} a & b & 1 & 0 \\ 0 & ad - bc & -c & a \end{bmatrix}.$$

The matrix is now in REF, and since A is invertible we must have $ad - bc \neq 0$. This is a simple condition you can now use to check whether a 2×2 matrix is invertible. This quantity is important enough to have a name; it's called the **determinant** of A, written

$$\det A = ad - bc.$$

Continuing with our row reduction:

$$R_1 \mapsto R_1 - (b/\det A)R_2 \qquad \begin{bmatrix} a & 0 & 1 + bc/\det(A) & -ba/\det(A) \\ 0 & \det(A) & -c & a \end{bmatrix}$$

$$\begin{matrix} R_1 \mapsto (1/a)R_1 \\ R_2 \mapsto (1/\det(A))R_2 \end{matrix} \qquad \begin{bmatrix} 1 & 0 & d/\det(A) & -b/\det(A) \\ 0 & 1 & -c/\det(A) & a/\det(A) \end{bmatrix}.$$

We now have a formula for A^{-1}:

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

For square matrices larger than 2×2, the best way to compute A^{-1} is to row reduce $[A\ I_n]$ using a computer.

Let's record what we learned about the determinant for $2{\times}2$ matrices below, because it will be useful later on.

What's the formula for the inverse of a 2×2 matrix?

Invertibility of 2 by 2 matrices

Theorem 4.14. *A 2×2 matrix*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is invertible if and only if $\det A = ad - bc \neq 0$. *If A is invertible, then*

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

There is a determinant for $n \times n$ matrices of any size, but the algorithm for computing it is difficult to describe and prove the correctness of, so we will not discuss it here.

RQ

Reading Question 4M. Write the system of linear equations

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

As a matrix vector equation $A\mathbf{x} = \mathbf{b}$. Compute A^{-1} and use it to find the solution to this equation.

RQ

Reading Question 4N. Write down the matrix for rotation counter-clockwise $\pi/4$ radians about the origin. Now, use the formula from Example 4.13 to compute its inverse. Confirm using some other method that this is rotation by $\pi/4$ *clockwise* about the origin.

What's a singular matrix?

The next theorem summarizes the key properties of the matrix inverse. A square matrix that is not invertible is called **singular**.

Properties of the matrix inverse

Theorem 4.15. *Let A and B be invertible $n \times n$ matrices and let C and D be arbitrary $n \times n$ matices.*

① *Invertible matrices can be cancelled:*

$$AC = AD \implies C = D.$$

② *AB is invertible and the matrix inverse is product reversing:*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

③ *The inverse of the inverse is the original:*

$$(A^{-1})^{-1} = A.$$

④ *A^{T} is invertible and inversion commutes with transposition:*

$$(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}.$$

Warning: don't cancel a singular matrix!

These follow fairly readily from the definition of inverse. For example, consider property ②. Let A and B be invertible matrices. The property says that $B^{-1}A^{-1}$ is the inverse of AB. To prove this, we'll verify that the definition holds:

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1}\\ &= AI_nA^{-1}\\ &= AA^{-1}\\ &= I_n,\end{aligned}$$

and a nearly identical computation shows that $(B^{-1}A^{-1})(AB) = I_n$.

Reading Question 4O. Prove Theorem 4.15 item ④. Given that A is invertible, the claim here is that the inverse of A^{T} is $(A^{-1})^{\mathrm{T}}$. So, you should compute the product of A^{T} and $(A^{-1})^{\mathrm{T}}$ both ways and show that you get I_n. A property of the transpose may be useful here. RQ

Exercise 4G. Is every nonzero square matrix invertible? If A and B are invertible matrices, you now have a formula for $(AB)^{-1}$. What about the inverse of a product of three invertible matrices, ABC?

Exercise 4H. Must the sum of two invertible matrices be invertible? [Hint: find an example that shows this isn't even true in **R**.]

Exercise 4I. Suppose P, Q, R are invertible and $Q^{-1}X^{\mathrm{T}}R^{-1} = P^{\mathrm{T}}R$. Solve for X.

§4.4 The Isomorphism Theorem

What's an isomorphism?

An invertible linear transformation from $\mathbf{R}^n$ to itself is called an **isomorphism**. We have already defined what it means for a function to be invertible, and it turns out that a function is invertible if and only if it is both one-to-one and onto.

A function is invertible iff it's one-to-one and onto

Theorem 4.16. *A function $f: X \to Y$ is invertible if and only if it is both one-to-one and onto.*

To understand this proof, review how to show functions are one-to-one and onto in §2.2 and §2.4.

Proof. First, suppose f is invertible. Then, for any $y \in Y$, we have

$$f(f^{-1}(y)) = y,$$

so every element of the codomain is in the image of f. This means f is onto. If $f(x) = f(y)$, then $f^{-1}(f(x)) = f^{-1}(f(y))$, which implies $x = y$. Thus, f is one-to-one as well.

Conversely, let f be a one-to-one and onto function. Define a function $g: Y \to X$ as follows. Given $y \in Y$, we know there is an element $x \in X$ such that $f(x) = y$, and since f is one-to-one there is only one way to pick x. So, if we define $g(y)$ to be x, we have $f(g(y)) = f(x) = y$ for all $y \in Y$. Now take $x \in X$ and consider $g(f(x))$. By definition of g, we have $g(f(x)) = x$. This proves that g is the inverse of f. ■

Invertible matrices, invertible linear transformations

Theorem 4.17. *Let $T: \mathbf{R}^n \to \mathbf{R}^n$ be a linear transformation with matrix representation A. The following statements are equivalent.*

① *The matrix A is invertible.*
② *The function T is an isomorphism.*
③ *The function T is one-to-one and onto.*

You will see in the proof below that when $T(\mathbf{x}) = A\mathbf{x}$ is an isomorphism, $T^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}$.

Proof. We proved that the second two items are equivalent in the previous theorem because isomorphism just means “invertible linear transformation”, so let's prove that the first two items are equivalent.

First, suppose that A is invertible and define $S\colon \mathbf{R}^n \to \mathbf{R}^n$ by $S(\mathbf{x}) = A^{-1}\mathbf{x}$. Then, for all $\mathbf{x} \in \mathbf{R}^n$ we have

$$T(S(\mathbf{x})) = A(A^{-1}\mathbf{x}) = (AA^{-1})\mathbf{x} = I_n\mathbf{x} = \mathbf{x}.$$

Similarly, $S(T(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in \mathbf{R}^n$. Thus, S is the inverse of T and hence T is an isomorphism.

Conversely, suppose T has an inverse T^{-1}. Then, T is one-to-one and onto by Theorem 4.16. The matrix A therefore has a pivot in every row and a pivot in every column. Since A is square it must be row equivalent to the identity and thus invertible. ■

The following theorem ties invertiblity to most of the ideas we've seen in linear algebra so far. For any $n \times n$ matrix A, the theorem says that either every statement is true or every statement is false.

Warning: it is absolutely critical that the domain and codomain of T are the same in this theorem!

The Isomorphism Theorem

Theorem 4.18. *Let $T : \mathbf{R}^n \to \mathbf{R}^n$ be a linear transformation with $n \times n$ matrix representation A (note that we are assuming* $\mathrm{domain}(T) = \mathrm{codomain}(T)$*). The following 17 statements are equivalent to one another:*

1. *T is an isomorphism.*

These are equivalent by Theorem 3.13.

2. *T is one-to-one.*
3. *A has a pivot position in every column.*
4. *$A\mathbf{x} = \mathbf{0}$ has a unique solution (the zero vector).*
5. *The columns of A are linearly independent.*

These are equivalent by Theorem 3.14.

6. *T is onto.*
7. *A has a pivot position in every row.*
8. *$A\mathbf{x} = \mathbf{b}$ is consistent for all $\mathbf{b} \in \mathbf{R}^n$.*
9. *The columns of A span $\mathbf{R}^n$.*

10. *A is invertible as a matrix.*
11. *There is an $n \times n$ matrix C such that $AC = I_n$.*
12. *There is an $n \times n$ matrix D such that $DA = I_n$.*

13. *T is invertible as a function.*
14. *There is a linear map $S : \mathbf{R}^n \to \mathbf{R}^n$ such that $T(S(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbf{R}^n$.*
15. *There is a linear map $S : \mathbf{R}^n \to \mathbf{R}^n$ such that $S(T(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x}$ in $\mathbf{R}^n$.*

16. *A is row equivalent to I_n.*
17. *A^{T} is invertible.*

Again: Theorem 4.18 applies only when A is square! *The concepts "one-to-one" and "onto" are not equivalent concepts when the domain and codomain of T are not the same.*

Proof. To prove this theorem, we must prove enough implications so that it is possible, starting with any statement, to follow a sequence of established implications to any other statement. As you read this proof, get out a piece of paper and keep track of the numbers 1–17. Put an arrow from one number to the other every time we prove an implication; put a double-headed arrow if we prove that two statements are equivalent.

By Theorem 4.17, (1), (10), and (13) are equivalent. Statement (10) is equivalent to (16) by Theorem 4.12.

Items (2)–(5) are equivalent by Theorem 3.14 and (6)-(9) are equivalent by Theorem 3.13. All of these statements are implied by (1) since an isomorphism is both one-to-one and onto.

Statement (10) implies statements (11) and (12) by definition of invertible matrix.

Statement (11) is equivalent to (14): given (11), define $S(\mathbf{x}) = C\mathbf{x}$; given (14), define C to be the matrix representation of S. Similarly, (12) is equivalent to (15).

So far, we have established the implications in the diagram below:

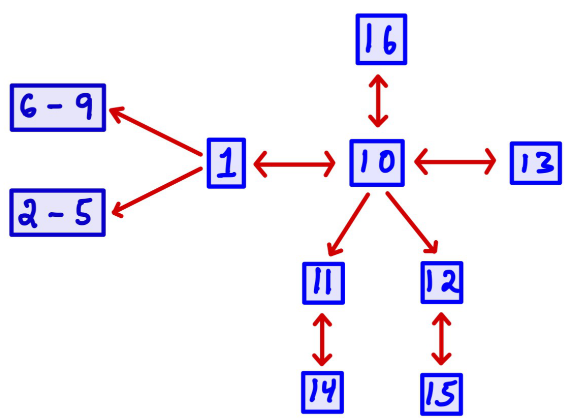


Figure 4.4: Isomorphism Theorem implications so far

There is more to do to finish the proof; this is outlined in Exercise 4J below. ■

You have probably noticed that we more often make statements about the columns of a matrix than the rows. Statement (17) enables us to get rows involved. Here's why: suppose A is an invertible matrix. Then, by (17), so is A^{T}.

Hence, the columns of A^{T} both span $\mathbf{R}^n$ and form a linearly independent set. But, the columns of A^{T} are precisely (the transposes of) the rows of A. So a square matrix A is invertible if and only if the rows of A span $\mathbf{R}^n$, if and only if the rows of A are linearly independent.

Exercise 4J. Look back at Figure 4.4. To finish the proof of Theorem 4.18, you can do the following:

- Show that any one of items (2)–(5) implies a statement equivalent to (1).
- Show that any one of items (6)–(9) implies a statement equivalent to (1).
- Show that (11) implies a statement equivalent to (1).
- Show that (12) implies a statement equivalent to (1).
- Show that (17) is equivalent to a statement equivalent to (1).

Accomplish the above tasks.

RQ

Reading Question 4P. Let Q be an $n \times n$ matrix with linearly independent columns. Explain why the columns of Q^2 span $\mathbf{R}^n$.

Exercise 4K. Let Q be an $n \times n$ matrix and suppose the linear transformation

$$T(\mathbf{x}) = (Q^{\mathrm{T}})^3\mathbf{x}$$

is not onto. Argue that $Q\mathbf{x} = \mathbf{0}$ has a nontrivial solution.

Exercise 4L. Let A be an $n \times n$ matrix and suppose there is some vector $\mathbf{b}$ in $\mathbf{R}^n$ such that $A\mathbf{x} = \mathbf{b}$ has two distinct solutions. Show that A is not row equivalent to the identity matrix.

Exercise 4M. Suppose A and B are $n \times n$ matrices and suppose the product AB is invertible. Must the matrix A also be invertible? Why or why not?

Exercise 4N. Let A and B be $n \times n$ matrices. Though AB is not always the same as BA, if AB is invertible then so is BA. Prove this.

Chapter 5 Geometry in the plane

Key concepts

- The area interpretation of the determinant
- A linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ scales areas by $|\det A|$
- The basic geometric transformations (rotating, scaling, shearing, projecting, and reflecting)
- Similar matrices: A is similar to B if they represent the same transformation in "different coordinates"
- Affine linear transformations $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$
- Homogeneous coordinates
- Fractals via iterated function systems
- Classification of distance-preserving transformations (linear isometries)

Summary. This chapter is about how linear transformations capture certain aspects of geometry in the plane. The area of the parallelogram determined by the columns of a matrix A is the absolute value of its determinant, and $\mathbf{x} \mapsto A\mathbf{x}$ scales areas of regions by $|\det A|$.

We discuss several key types of geometric transformations—rotations, reflections, projections, shears, and scalings—finding their matrix representations and illustrating their geometry. The concept of similar matrices is introduced to describe how different matrices can represent the same transformation in different coordinate systems (but we'll talk a lot more about this in much greater detail later).

We can extend linear transformations to "affine linear transformations", which include translations. Affine maps can be represented using homogeneous coordinates and 3×3 matrices. This framework allows us to build self-similar geometric figures called fractals, like the Sierpinski triangle.

We conclude by classifying the linear isometries of the plane, identifying precisely which transformations are distance preserving: the rotations and reflections. The matrix of a linear isometry is an orthogonal matrix; we'll study these in more detail later.

Chapter 5

We have already met several linear transformations (and their matrix representations) with geometric content: rotating the plane by θ radians counterclockwise about the origin (Example 2.6), uniformly scaling the plane (Example 2.3), and reflecting over the line $y = x$ (see the introduction to Chapter 4). Before we revisit these transformations (and introduce some new ones), we will first show you how the area of a region is related to the area of its image under a linear transformation.

§5.1 Area and the determinant

It turns out that the determinant controls the way a region's area changes when a linear transformation is applied to it. When we say "region in the plane", we just mean any subset of the plane where the concept of area makes sense.

How do linear transformations affect the areas of regions they transform?

Determinants and areas

Theorem 5.1. *Let $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ be a linear transformation with matrix representation $A = [\mathbf{v}_1\ \mathbf{v}_2]$.*

1. *The area of the parallelogram determined by $\mathbf{v}_1$ and $\mathbf{v}_2$ is $|\det A|$.*
2. *For any region S in the plane with finite area,*

$$\mathrm{Area}(T(S)) = |\det A| \cdot \mathrm{Area}(S).$$

Proof. Let's write

$$A = [\mathbf{v}_1\ \mathbf{v}_2] = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

and let P denote the parallelogram determined by $\mathbf{v}_1$ and $\mathbf{v}_2$. If $\{\mathbf{v}_1, \mathbf{v}_2\}$ is linearly dependent, then both the determinant and the area of P are zero. (In this situation, P is just a line segment.) So let's assume this set is linearly independent.

Call a square matrix M **diagonal** if $M_{i,j} = 0$ whenever $i \neq j$. If A is diagonal, then $b = c = 0$ and

What's a diagonal matrix?

$$A = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}.$$

The parallelogram P is a square with side lengths $|a|$ and $|d|$ (the lengths of $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively), so in this case we indeed have that the area of P is $|ad| = |\det A|$. Now the question is: how can we connect the general situation to the situation where A is diagonal?

We will use *column* operations applied to A. These are exaclty the row operations, but applied to columns instead of rows. Doing a column operation on A is the same as doing a row operation on A^{T}. In an exercise below, you will verify that for a 2×2 matrix A, we have $\det B = \det A$ if B is obtained from A using a column replacement operation, and $\det B = -\det A$ if B is obtained from A using a column swap. Thus, neither operation changes $|\det A|$.

Further, it is clear that a column swap does not change the parallelogram P determined by the columns. What about column replacement? Let P' be the parallelogram determined by the columns of $[\mathbf{v}_1 \ \mathbf{v}_2 + t\mathbf{v}_1]$.

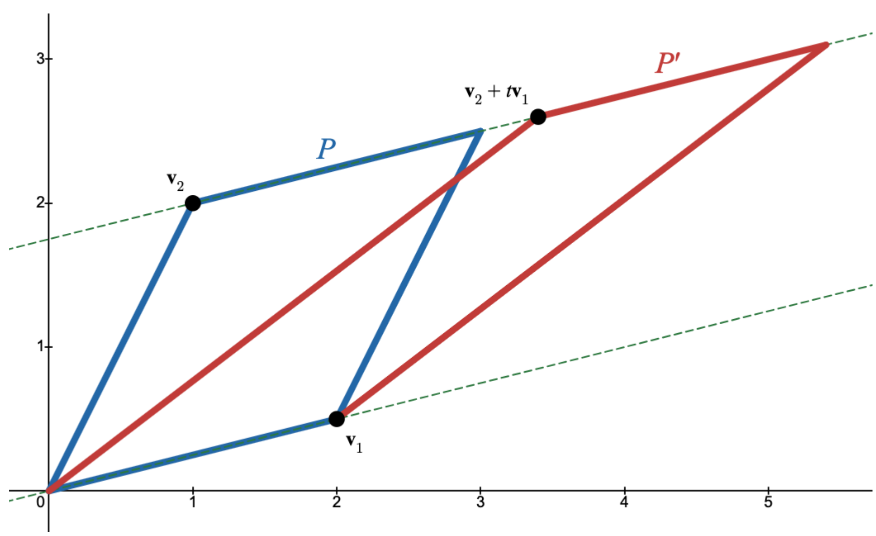


Figure 5.1: Parallelograms with the same base and height

Note that the parallelograms P and P' in Figure 5.1 have exactly the same base and height; hence, they have the same area. So column swaps and replacements change neither $|\det A|$ nor the area of the parallelogram determined by the columns of the matrix.

Do this with $\begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$.

To finish the proof, we just need to be sure we can use column operations to change A into a diagonal matrix. To do so, first use a column swap (if necessary) to make the upper left entry nonzero. Then, do a replacement in column 2 (if necessary) to make the upper right entry zero. Now, the lower right entry is nonzero, for otherwise the new matrix would not be invertible. However, it must be invertible, because it's equivalent to the invertible matrix A via column operations that do not change the absolute value of $\det A$, which is nonzero. Finally, do a replacement in column 1 (if necessary) to zero out the bottom left entry.

We leave the proof of the second item, in the case where S is a parallelogram, to the exercises. For general regions S, we will omit the proof entirely. The rough idea is to approximate the region using squares, use the fact that the theorem holds for squares, and then finish the argument using limit techniques from calculus. ■

RQ

Reading Question 5A. Let $A = [\mathbf{v}_1\ \mathbf{v}_2]$ be a 2×2 matrix and suppose B is obtained from A by swapping the columns. Check that $\det A = -\det B$. Suppose C is obtained from A by doing the column replacement $\mathbf{v}_2 \mapsto \mathbf{v}_2 + t\mathbf{v}_1$. Check that $\det A = \det C$.

Exercise 5A. Let S be a parallelogram in the plane. It's fairly easy to see that translating S doesn't change its area, so to prove Theorem 5.1 item (2) for parallelograms it suffices to consider parallelograms with a vertex at the origin. Hence, we can assume that S is determined by two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$, in which case $T(S)$ is the parallelogram determined by $A\mathbf{v}_1$ and $A\mathbf{v}_2$. To finish the argument, prove that $\det[A\mathbf{v}_1\ A\mathbf{v}_2] = |\det A| \cdot |\det[\mathbf{v}_1\ \mathbf{v}_2]|$.

Example 5.2 (the area of an ellipse)**.** The area of the unit circle is π. A typical ellipse centered at the origin has equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

where $a, b > 0$. Can we find a linear transformation that maps the unit circle onto this ellipse? Yes:

$$T(x, y) = (ax, by).$$

Because:

$$x_0^2 + y_0^2 = 1 \iff \frac{(ax_0)^2}{a^2} + \frac{(by_0)^2}{b^2} = 1$$
$$(x_0, y_0) \text{ lies on the circle} \iff T(x_0, y_0) \text{ lies on the ellipse}$$

So applying Theorem 5.1 to this situation (with the unit circle in the role of S) yields the following formula:

$$\text{area of ellipse } = \left|\det \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}\right| \cdot \pi = ab\pi.$$

Which do you like better, finding the area this way, or finding the area via integration?

Basic geometric transformations §5.2

Recall that the matrix representation of a linear transformation $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ is given by

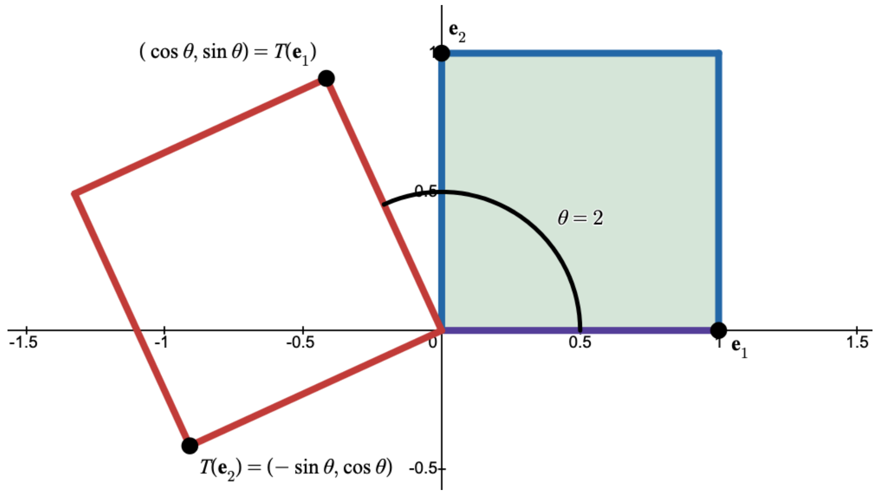
$$A = [T(\mathbf{e}_1)\ T(\mathbf{e}_2)].$$

So if you want to find the matrix of a geometric linear transformation, you just need to figure out where T maps $\mathbf{e}_1$ and $\mathbf{e}_2$.

Rotations §5.2.1

We've already covered rotations. The matrix for the linear transformation that rotates the plane counter-clockwise θ radians about the origin is given by

$$\mathrm{R}_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$



www.desmos.com/calculator/ouh3bg4slx

Figure 5.2: CCW rotation (Desmos)

Since every rotation matrix has determinant 1, rotating a region does not change its area.

Exercise 5B. It is clear geometrically that $\mathrm{R}_\alpha \mathrm{R}_\theta = \mathrm{R}_{\alpha+\theta}$. Multiply the two matrices on the left hand side and set the result equal to the right hand side. You have just discovered the angle sum trig identities.

§5.2.2 *Shear transformations*

Shearing in the x-direction with shear parameter t takes the unit box pictured below and "shifts" it to the right (if $t > 0$) or to the left (if $t < 0$) in such a way that the bottom of the box remains fixed. The matrix representation is

$$\mathrm{H}_t^x = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

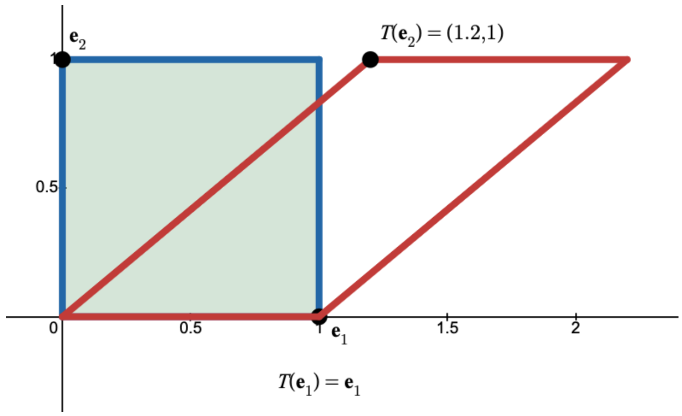


Figure 5.3: Shearing in the x-direction (Desmos)

www.desmos.com/calculator/9qe4wnn8g2

Similarly, shearing in the y-direction with shear parameter t shifts the box up (if $t > 0$) or down (if $t < 0$) in such a way that the left side remains fixed. The matrix representation is

$$\mathrm{H}_t^y = \begin{bmatrix} 1 & 0 \\ t & 1 \end{bmatrix}.$$

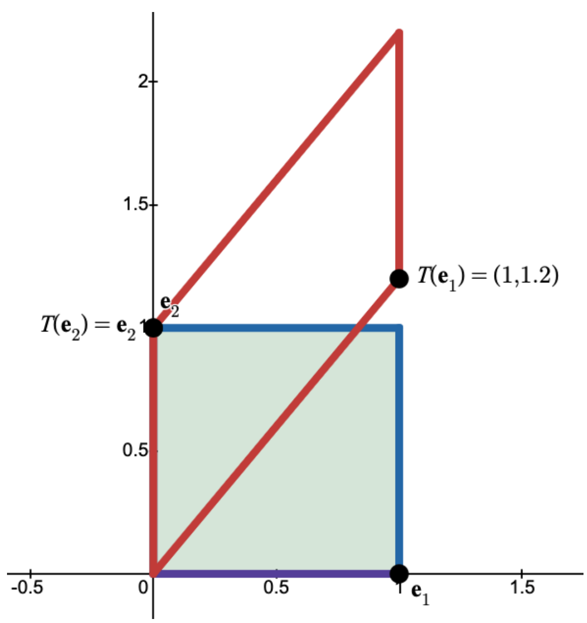


Figure 5.4: Shearing in the y-direction (Desmos)

www.desmos.com/calculator/v2mttabtly

Since shearing matrices have determinant 1, shearing a region does not change its area. This is also clear because the area of a parallelogram is base times height.

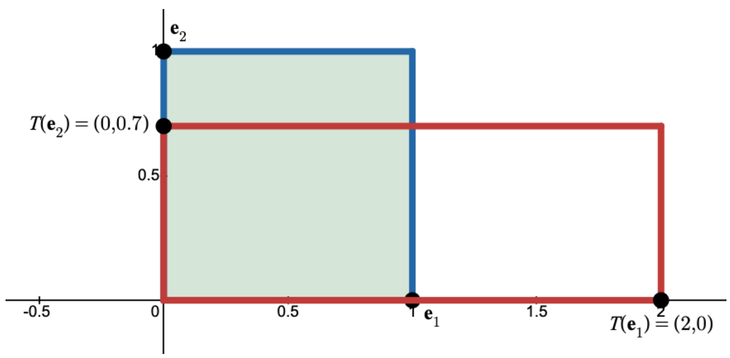
RQ

Reading Question 5B. Suppose $T: \mathbf{R}^2 \to \mathbf{R}^2$ is a linear transformation with $T(\mathbf{e}_2) = \mathbf{e}_2$ and $T(\mathbf{e}_1) = \mathbf{e}_1 - 5\mathbf{e}_2$. Identify T as a shear transformation using the notation from this section.

§5.2.3 *Scaling transformations*

The transformation that scales in the x-direction by s and the y-direction by t has matrix representation

$$\mathrm{S}_{s,t} = \begin{bmatrix} s & 0 \\ 0 & t \end{bmatrix}.$$



www.desmos.com/calculator/vwbt8rmazc

Figure 5.5: Scaling (Desmos)

When $s = t$, we will just write S_t for this matrix. Since scaling matrices have determinant st, scaling a region multiplies its area by $|st|$. Note that if s or t is negative, then this transformation actually involves reflection over the x-axis, the y-axis, or both.

RQ

Reading Question 5C. Write down the matrices $\mathrm{S}_{1,-1}, \mathrm{S}_{-1,1}$, and $\mathrm{S}_{-1,-1}$. What does each matrix do geometrically?

Projections §5.2.4

Projection onto the x-axis maps a point (x, y) to $(x, 0)$, and projection onto the y-axis maps a point (x, y) to $(0, y)$. Their matrix representations are

$$\pi_x = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } \pi_y = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Since projection matrices map all regions into lines, the projection of any region has area zero (and the determinants of the above matrices are zero).

Reflections §5.2.5

Let M_L denote reflection over the line L. What does reflection do? Given any point P in the plane, reflection over L maps P to the point P' that (1) lies on the line through P that is perpendicular to L; (2) lies on the other side of the line from P; and (3) is the same distance to the line as P. If P lies on L, then $P' = P$.

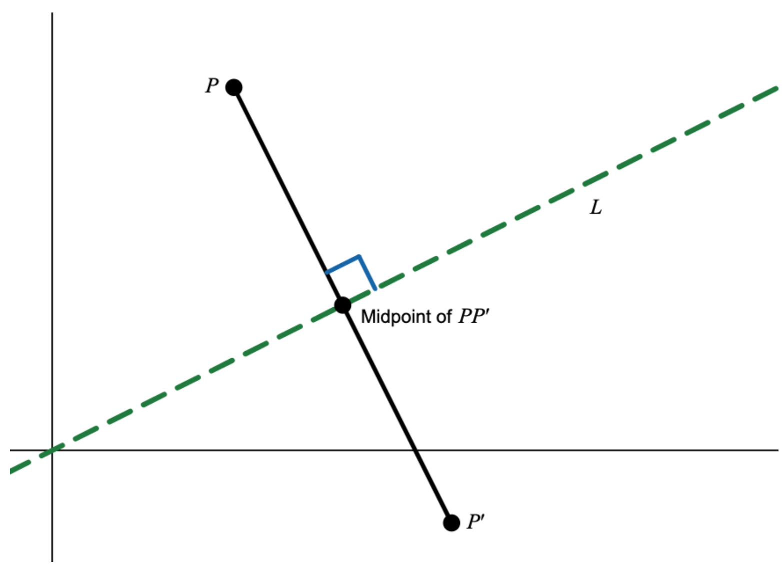


Figure 5.6: Reflection over the line L

For example, $\mathrm{M}_{y=0}$ is reflection over the x-axis. It maps $\mathbf{e}_1$ to itself and $\mathbf{e}_2$ to $-\mathbf{e}_2$. Its matrix representation is

$$\mathrm{M}_{y=0} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Now lets suppose that L is an arbitrary line through the origin whose angle with the positive x-axis, measured counter-clockwise, is θ. To do the reflection

M_L we can (1) rotate by $-\theta$ to move L to the x-axis; (2) reflect over the x-axis; and (3) rotate by θ to move the x-axis back to L. So,

Remember: matrix multiplication mirrors function composition so the order in which the transformations happen is right to left!

$$M_L = R_\theta \, M_{y=0} \, R_{-\theta}.$$

We leave it to you to find a general formula by computing this product in the next exercise. In specific situations, it's sometimes easier to just track what the reflection does to $\mathbf{e}_1$ and $\mathbf{e}_2$. For example, reflection over the line $y = x$ swaps $\mathbf{e}_1$ and $\mathbf{e}_2$, so

$$M_{y=x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Exercise 5C. Compute the matrix $M_L = R_\theta \, M_{y=0} \, R_{-\theta}$ and use trig identities to make it as simple as possible.

There is an interesting point here worth lingering on. Since rotation by $-\theta$ is the inverse of rotation by θ, we have

$$M_L = R_\theta \, M_{y=0} \, R_\theta^{-1}.$$

This relationship between M_L and $M_{y=0}$ has a name.

Similar matrices

What are similar matrices?

Definition 5.3. Suppose that A and B are $n \times n$ matrices. We say that A is **similar** to B if there is invertible $n \times n$ matrix P for which $A = PBP^{-1}$.

RQ

Reading Question 5D. In this problem, A, B, and C are $n \times n$ matrices.

(a) Show that A is similar to A. [Hint: saying that A is similar to A means that $A = PAP^{-1}$ for **some** invertible matrix P — not necessarily for **all** invertible matrices P. Look for an invertible P that makes $A = PAP^{-1}$ obvious!]

(b) Show that if A is similar to B, then B is similar to A. [Hint: $A = PBP^{-1}$. Solve for B.]

(c) Show that A is similar to B and B is similar to C, then A is similar to C. [Hint: $A = PBP^{-1}$ and $B = QCQ^{-1}$. Can you combine the equations to find an invertible matrix R such that $A = RCR^{-1}$?]

In light of point (b) above, if A is similar to B (and so B is similar to A), we usually just say that A **and** B **are similar**.

When we investigate the concept of similarity in detail later, we will come to understand that when A and B are similar, they "represent the same linear transformation with respect to different sets of coordinates." To get some sense of what this means, look back at the relationship $\mathrm{M}_L = \mathrm{R}_\theta\, \mathrm{M}_{y=0}\, \mathrm{R}_\theta^{-1}$ that we discovered above. This relationship says that "reflection over any line L is similar to reflection over the x-axis." Now, suppose we were allowed to rotate our coordinate system so that the x-axis lined up with L; then reflection over L would literally be the same as reflection over the x-axis!

Reading Question 5E. To show that two matrices A and B are similar, you must find an invertible matrix P such that $A = PBP^{-1}$. This equation is equivalent to $AP = PB$ which is often easier to verify. Let H_s^x and H_t^x be shear transformations with $s, t \neq 0$. Use the invertible matrix

$$P = \begin{bmatrix} 1 & 0 \\ 0 & t/s \end{bmatrix}$$

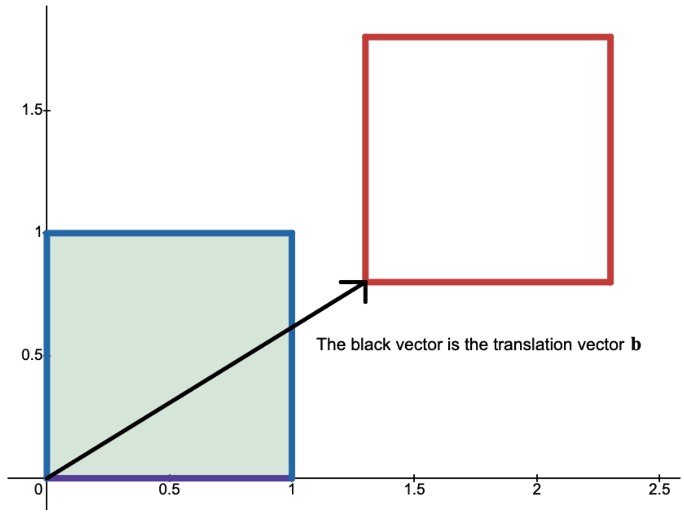
to show that H_s^x and H_t^x are similar matrices by verifying the equation $\mathrm{H}_s^x\, P = P\, \mathrm{H}_t^x$. RQ

Reading Question 5F. Show that if A is similar to B, then A^2 is similar to B^2. Can you generalize? RQ

Exercise 5D. Suppose that A is a 2×2 matrix that is not similar to any matrix other than itself. What must A look like?

Affine linear transformations §5.3

Take $\mathbf{b} \in \mathbf{R}^2$. **Translation** by $\mathbf{b}$ is the function $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ defined by $T(\mathbf{x}) = \mathbf{x} + \mathbf{b}$. Please note that translation is *not* linear unless $\mathbf{b} = 0$ because $T(\mathbf{0}) = \mathbf{b}$, and linear transformations must send zero to zero. *What's a translation?*



www.desmos.com/calculator/giajhnyrbe

Figure 5.7: Translation by a vector (Desmos)

In general, if A is a 2×2 matrix and $\mathbf{b} \in \mathbf{R}^2$, we call a function $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ of the form

$$T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$$

What's an affine linear transformation?

an **affine linear transformation**. Affine linear transformations allow us to increase the variety of geometric transformations we can create because we can mix translations with the others via function composition.

RQ

Reading Question 5G. In high school we learn to call functions of the form $y = mx + b$ *linear functions*. Comment on this.

Example 5.4. Suppose I want to first rotate the plane by $\pi/3$ counter-clockwise, then translate by the vector $(1, 2)$, and finally shear in the x-direction by a factor of 2. What we seek is the composite of these transformations in the correct order:

$$\mathbf{x} \mapsto \mathrm{R}_{\pi/3}\,\mathbf{x} \qquad \text{[First]}$$

$$\mathbf{x} \mapsto \mathbf{x} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \text{[Second]}$$

$$\mathbf{x} \mapsto \mathrm{H}_2^x\,\mathbf{x} \qquad \text{[Third]}$$

The composite transformation is

$$\begin{aligned}\mathbf{x} \mapsto \mathrm{H}_2^x\left((\mathrm{R}_{\pi/3}\,\mathbf{x}) + \begin{bmatrix}1\\2\end{bmatrix}\right) \\ = (\mathrm{H}_2^x\,\mathrm{R}_{\pi/3})\mathbf{x} + \mathrm{H}_2^x\begin{bmatrix}1\\2\end{bmatrix}.\end{aligned}$$

This is an affine linear transformation where $A = \mathrm{H}_2^x\,\mathrm{R}_{\pi/3}$ and $\mathbf{b} = \mathrm{H}_2^x\begin{bmatrix}1\\2\end{bmatrix}$.

Since affine linear transformations are usually not linear, they do not have matrix representations, and this is irritating. However, there is a trick for representing affine linear transformations with 3×3 matrices in such a way that in order to combine affine linear transformations using function composition, you can just multiply together their associated 3×3 matrices.

Let T be the affine linear transformation $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$, where

$$A = \begin{bmatrix}a & b\\c & d\end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix}p\\q\end{bmatrix}.$$

Define the **affine matrix representation** of T to be the 3×3 matrix

What's an affine matrix representation?

$$E = \begin{bmatrix}a & b & p\\c & d & q\\0 & 0 & 1\end{bmatrix}.$$

To simplify the notation for E, we will write it as

$$E = \begin{bmatrix}A & \mathbf{b}\\0 & 1\end{bmatrix}.$$

In the matrix above, the "0" is actually the 1×2 zero matrix.

We can't multiply E by our input vector $\mathbf{x} = (x_1, x_2)$, so we will convert this vector into **homogeneous coordinates**. The homogeneous coordinate vector of $\mathbf{x}$ is the 3-vector

$$\begin{bmatrix}\mathbf{x}\\1\end{bmatrix} = \begin{bmatrix}x_1\\x_2\\1\end{bmatrix}.$$

Informally, we're thinking of $\mathbf{R}^2$ as the plane with equation $x_3 = 1$ in $\mathbf{R}^3$. Why did we do this? Watch what happens when we multiply our homogeneous coordinate vector by E:

$$\begin{bmatrix} A & \mathbf{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & p \\ c & d & q \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} ax_1 + bx_2 + p \\ cx_1 + dx_2 + q \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} A\mathbf{x} + \mathbf{b} \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} T(\mathbf{x}) \\ 1 \end{bmatrix}.$$

The "1" in our homogeneous coordinate vector is just a placeholder—we really can use E to compute the values of T.

RQ

Reading Question 5H. Define $A, \mathbf{b}, T$, and E as in the above discussion.

1. What is E when T is translation by the vector $(-3, 4)$?
2. What is E when T is the linear transformation that scales in the x-direction by a factor of s and in the y-direction by a factor of t?
3. What is E when T first rotates by $\pi/2$, then translates by $(1, 3)$, and finally shears in the y-direction with shear factor -5?

Exercise 5E. Suppose $T(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ has affine matrix representation E and $S(\mathbf{x}) = C\mathbf{x} + \mathbf{d}$ has affine matrix F. Simplify both $T(S(\mathbf{x}))$ and EF. You should see that the matrix for $T \circ S$ is exactly EF.

The transformations in an IFS need to be sufficiently nice; roughly, they need to decrease distances between points.

We can use affine linear transformations to construct geometric figures called **fractals**. To create a fractal, we need a list of affine linear transformations $\{T_1, \ldots, T_n\}$ called an **iterated function system**. To illustrate the process used to create a fractal from these transformations, we need an initial figure in the plane to which we can apply the transformations. We'll represent a figure as a list of points $P_1, P_2, \ldots, P_k$ in homogeneous coordinates. For convenience, we will collect these points into a single matrix

$$F = [P_1 \ P_2 \ \cdots \ P_k].$$

When drawing the figure, we will connect consecutive points with a line segment and then close the figure off by connecting P_k to P_1.

For example, the initial figure usually used to create the **Sierpinski triangle** is the equilateral triangle with vertices $(0, 0)$, $(1, 0)$, and $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$.

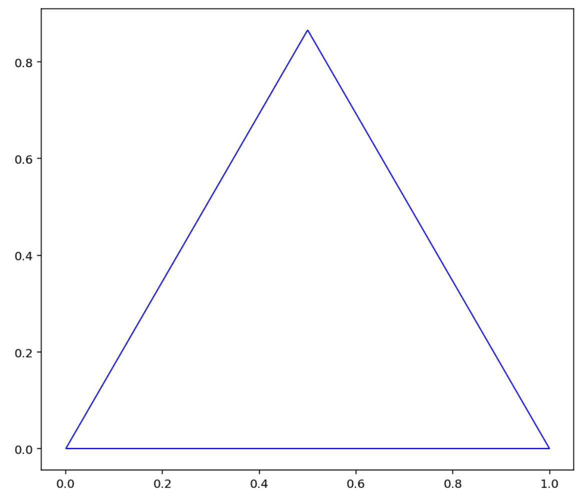


Figure 5.8: Sierpinski triangle initial figure

As a matrix, this figure is

$$F = \begin{bmatrix} 0 & 1 & 1/2 \\ 0 & 0 & \sqrt{3}/2 \\ 1 & 1 & 1 \end{bmatrix}$$

We can alter this figure by applying an affine linear transformation T to it. To do so, we apply the transformation to each point to obtain a new figure. If E is the affine matrix representation of T, then

$$E[P_1 \ P_2 \ \cdots \ P_k] = [EP_1 \ EP_2 \ \cdots \ EP_k],$$

so we can transform an entire figure using matrix multiplication!

To build the Sierpinski triangle, we will use three affine linear transformations $\{T_1, T_2, T_3\}$. Transformation T_1 will map the original equilateral triangle F (the big outer triangle in Figure 5.9 below) to the smaller triangle labeled **1**. Transformations T_2 and T_3 will map F to the triangles labeled **2** and **3**, respectively. Each of the three new triangles has half the width of the original.

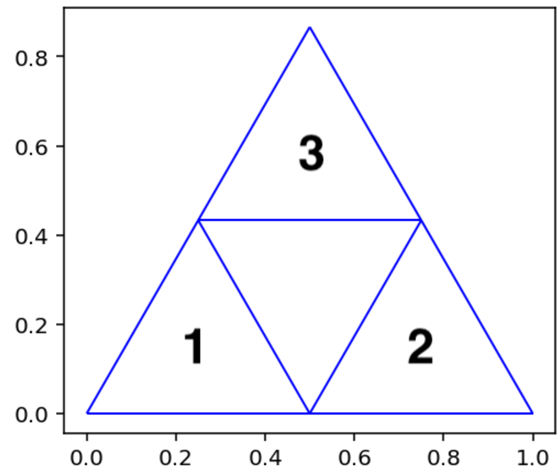


Figure 5.9: Building Sierpinski: three transformations

Triangle **1** in the lower left corner is obtained by scaling the original triangle by a factor of $\frac{1}{2}$. So our first transformation T_1 is

$$\mathbf{x} \mapsto S_{1/2}\,\mathbf{x}.$$

Triangle **2** in the lower right corner is obtained by first scaling the original triangle by a factor of $\frac{1}{2}$ and then translating it to the right $\frac{1}{2}$ a unit (so, the translation vector is $(1/2, 0)$). Our second transformation T_2 is therefore

$$\mathbf{x} \mapsto S_{1/2}\,\mathbf{x} + (1/2, 0).$$

Triangle **3** is obtained in three steps: first translate the top vertex to the origin (with translation vector $(-1/2, -\sqrt{3}/2)$), then scale by $\frac{1}{2}$, and finally translate the top vertex back to its original location (with translation vector $(1/2, \sqrt{3}/2)$). Our third transformation T_3 is

$$\mathbf{x} \mapsto S_{1/2}(\mathbf{x} + (-1/2, -\sqrt{3}/2)) + (1/2, \sqrt{3}/2).$$

Exercise 5F. Draw pictures to illustrate the three steps that are used to define T_3 above. Find the affine matrix representations for T_1, T_2, and T_3.

Our iterated function system (IFS) will create a fractal as the limit of the following process. Choose a figure F in the plane. Starting with $L_0 = \{F\}$, recursively define the list of figures L_s by

$$L_s = \{T_i(Y) \mid Y \in L_{s-1}, i = 1, \ldots, n\}.$$

To compute L_s, you take every figure in L_{s-1} and apply every transformation to it. At each step in the process, we can make a plot that contains all the figures in L_s. It turns out that these plots converge to a limiting figure $\hat{F}$. Moreover, this limiting figure $\hat{F}$ has the property that if you do each of our affine transformations T_i to $\hat{F}$ and plot all of the figures $T_i(\hat{F})$ together, the resulting plot is exactly $\hat{F}$ again! This property of $\hat{F}$ is sometimes referred to as **self-similarity**.

The following plots illustrate L_1, L_2, L_3, and L_8 for the Sierpinski triangle. If you take this further than $s = 8$ at the fixed scale given in the images, you will not notice any changes (though you certainly would if you zoomed in on any part of the figure).

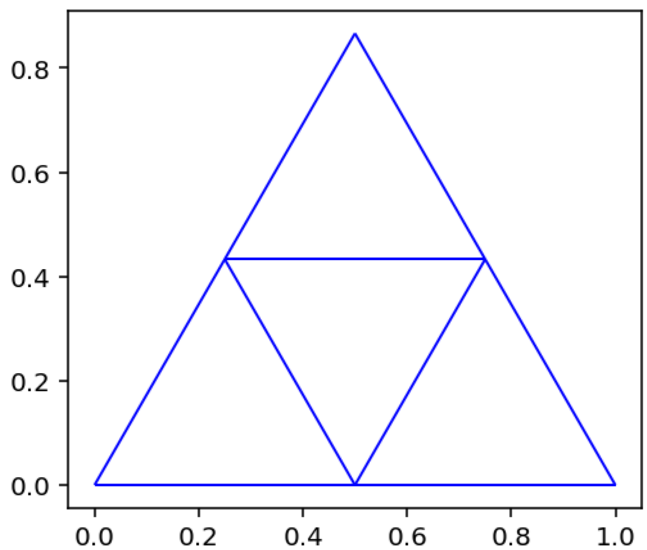


Figure 5.10: Sierpinski triangle L_1

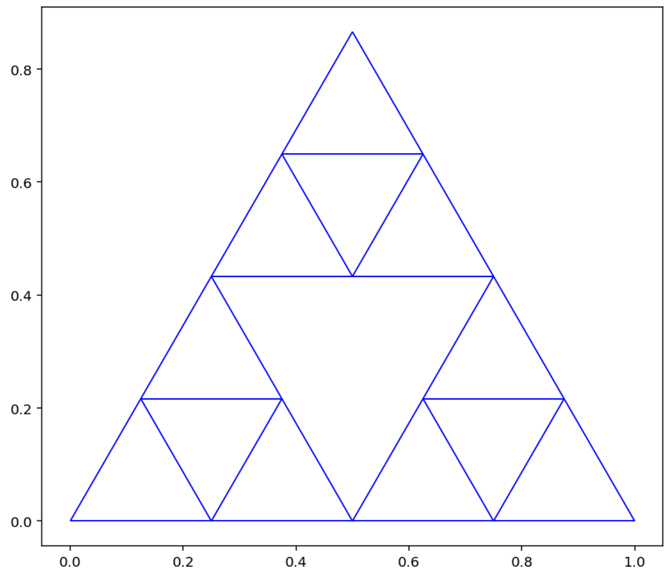


Figure 5.11: Sierpinski triangle L_2

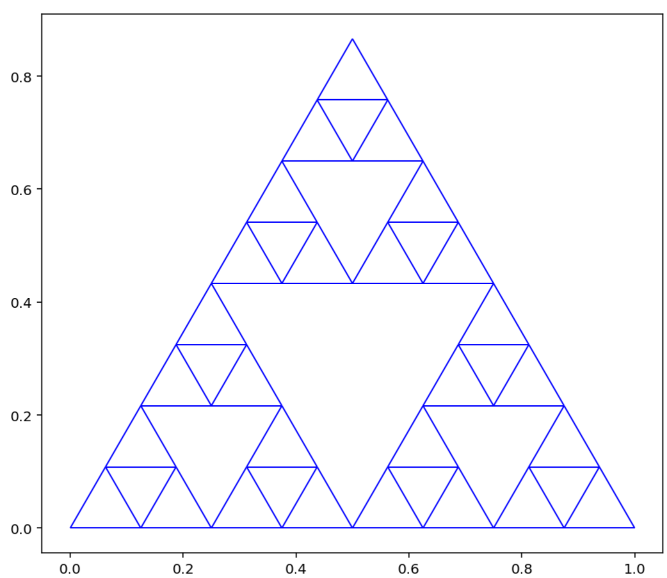


Figure 5.12: Sierpinski triangle L_3

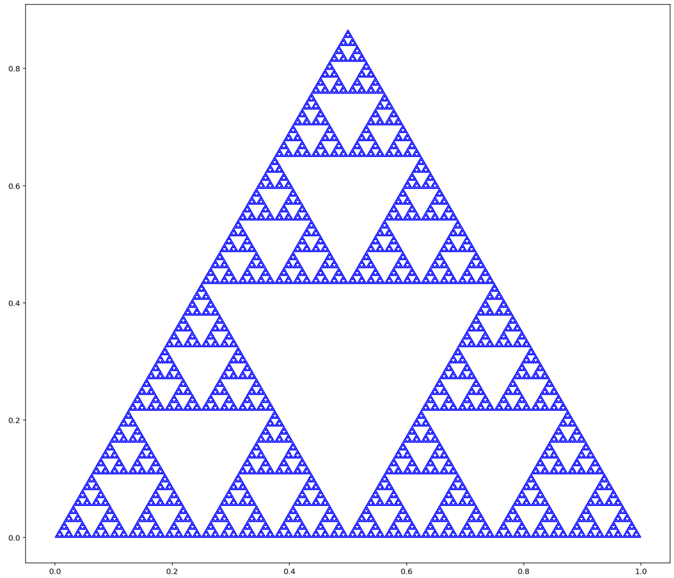
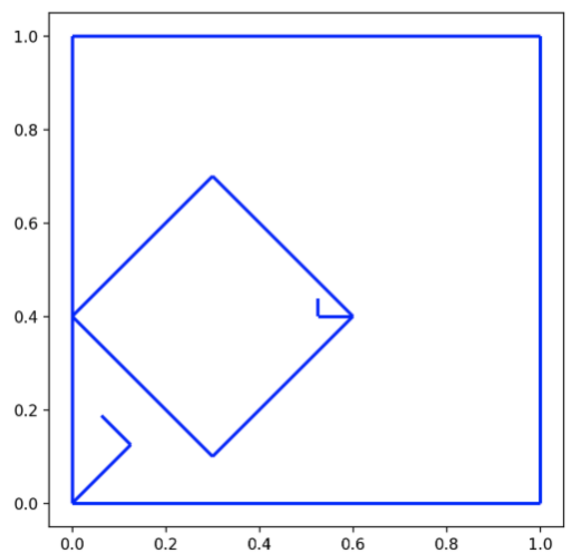


Figure 5.13: Sierpinski triangle L_8

We want to emphasize that when you plot the figures in L_s for any given s, you're looking at an *approximation* to the actual fractal which is (roughly speaking) the limit of this sequence of images. Surprisingly, the fractal (the limit of the process, not the approximations) is the same for essentially any choice of initial figure!

Exercise 5G. In the diagram below, let's call the larger outer box with an "L" in the corner A and let's call the smaller box with a (backwards) "L" in the corner B.



Find an affine linear transformation T such that $T(A) = B$. Express your answer as a composite of basic geometric transformations and/or translations and as a single transformation (give the affine matrix).

§5.4 Distance preserving linear transformations

In Euclidean geometry, two triangles are congruent if there is a way to pick one up and place it on top of the other so that they coincide. You can translate the triangle, you can rotate the triangle, and you can flip it over, but you can't stretch it or scale it. Another way to put this is that there needs to be a function from the plane to itself that takes one triangle to the other and does not change distances between points. Such a transformation is called an isometry, and in this section we will classify the isometries of the plane that are also linear transformations.

In the plane, the distance between two points (a, b) and (c, d) is given by

$$\sqrt{(a-c)^2 + (b-d)^2}.$$

What's the length of a vector in the plane?

Define the **length** or **magnitude** of a vector $\mathbf{v} = (a, b)$, written $|\mathbf{v}|$, to be the distance from $\mathbf{v}$ to the origin:

$$|\mathbf{v}| = \sqrt{a^2 + b^2}.$$

The distance between two vectors $\mathbf{v}, \mathbf{w} \in \mathbf{R}^2$ is $|\mathbf{v} - \mathbf{w}|$.

RQ

Reading Question 5I. For any scalar t and vector $\mathbf{v}$, show that $|t\mathbf{v}| = |t||\mathbf{v}|$. Does this equation make sense to you geometrically?

RQ

Reading Question 5J. Suppose that x and y are real numbers (that is, vectors in $\mathbf{R}^1$) and let $|x - y|$ denote the usual absolute value of $x - y$. Make a few different choices of x and y (with varying signs) and verify that, in each case, $|x - y|$ is the distance between x and y and $|x|$ is the distance from x to 0. This is how you should think about absolute value for the rest of your life.

What's an isometry?

Call a function $f\colon \mathbf{R}^2 \to \mathbf{R}^2$ an **isometry** if it preserves distances: for all $\mathbf{v}, \mathbf{w} \in \mathbf{R}^2$,

$$|f(\mathbf{v}) - f(\mathbf{w})| = |\mathbf{v} - \mathbf{w}|.$$

Such functions must preserve angles as well. Consider a triangle with vertices A, B, and C in the plane and let α denote the angle at vertex A. The transformed triangle with vertices $f(A)$, $f(B)$, and $f(C)$ has the same side lengths since f is an isometry, so the angle at $f(A)$ must also be α because the two triangles are congruent by the side-side-side congruence theorem from geometry.

The next theorem classifies the linear isometries of the plane.

LINEAR ISOMETRIES OF THE PLANE

Theorem 5.5. *Let $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ be a linear transformation. T is an isometry if and only if its matrix representation is R_θ or $\mathrm{R}_\theta\,\mathrm{M}_{y=0}$ for some $\theta \in \mathbf{R}$.*

Recall that $\mathrm{M}_{y=0}$ is reflection over the x-axis.

Proof. We will take it as intuitively clear that rotations and reflections do not change distances between points. So our job is to show that the listed matrices are the only ones that give isometries.

Let T be a linear transformation from the plane to itself, and let $A = [\mathbf{v}\ \mathbf{w}]$ be its matrix representation. Since $\mathbf{v} = T(\mathbf{e}_1)$ and $\mathbf{w} = T(\mathbf{e}_2)$, the columns of A must be unit length vectors because $\mathbf{e}_1$ and $\mathbf{e}_2$ have unit length (and T preserves distances). Further, since $\mathbf{e}_1$ and $\mathbf{e}_2$ are perpendicular, so are $\mathbf{v}$ and $\mathbf{w}$, because isometries are angle preserving. So both of these vectors are points on the unit circle that are $\pi/2$ radians apart. If we write

$$\mathbf{v} = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix},$$

then either

$$\mathbf{w} = \begin{bmatrix} \cos(\theta + \pi/2) \\ \sin(\theta + \pi/2) \end{bmatrix} = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix},$$

in which case A is R_θ, or

$$\mathbf{w} = \begin{bmatrix} \cos(\theta - \pi/2) \\ \sin(\theta - \pi/2) \end{bmatrix} = \begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix}.$$

In the second case,

$$\begin{aligned} A &= \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ &= \mathrm{R}_\theta\,\mathrm{M}_{y=0}. \end{aligned}$$

■

It is a fact that the matrix $\mathrm{R}_\theta\,\mathrm{M}_{y=0}$ is reflection over the line whose counterclockwise angle with the positive x-axis is $\theta/2$, but we leave this as an exercise. Thus, the linear isometries of the plane are exactly the rotations (about the origin) and the reflections (over lines through the origin).

RQ

Reading Question 5K. If A is the matrix representation of a linear isometry, how can we use the determinant to determine whether it's a reflection or a rotation? [Hint: compute the determinant of the matrices in Theorem 5.5.]

Exercise 5H. Let θ be a real number and let L denote the line through the origin whose counter-clockwise angle with the x-axis is $\theta/2$. Show that $\mathrm{M}_L = \mathrm{R}_\theta \, \mathrm{M}_{y=0}$. [Hint: in §5.2.5 we observed that $\mathrm{M}_L = \mathrm{R}_{\theta/2} \, \mathrm{M}_{y=0} \, \mathrm{R}_{\theta/2}^{-1}$.]

In Chapter 11, we will study a special class of matrices called orthogonal matrices, where the columns are pair-wise perpendicular and have unit length. The above theorem is a classification of the 2×2 orthogonal matrices.

Exercise 5I. Let A be a 2 by 2 matrix and define $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ by $T(\mathbf{x}) = A\mathbf{x}$. If $\det A = \pm 1$, does that necessarily mean that T is an isometry? If $\det A = 3$, can T be an isometry?

Exercise 5J. Factor the matrix

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

as a product of basic transformation matrices (from §5.2).

Chapter 6	Subspaces of R^n

Key concepts

- Subspaces are nonempty subsets of $\mathbf{R}^n$ that are closed under + and $\cdot$
- Spans, kernels, and images are all subspaces
- Bases are linearly independent spanning sets for subspaces
- Every subspace has a basis, and all bases for a subspace have the same size
- The dimension of a subspace is the number of vectors in any basis
- How to find bases for $\operatorname{im} A$, $\ker A$ and $\operatorname{row} A$
- A linear transformation is completely determined by its action on a basis
- The rank of a matrix is the dimension of its image (or row space)
- The Rank Theorem: $\operatorname{rank} A + \dim \ker A =$ number of columns of A

Summary. This chapter introduces subspaces—subsets of $\mathbf{R}^n$ that are closed under vector addition and scalar multiplication. Spans are subspaces, and familiar sets like kernels and images of linear transformations are subspaces. Lines and planes through the origin are low-dimensional examples of subspaces.

A basis is a set of vectors that spans a subspace without any redundancy. Every subspace has a basis, and all bases for a given subspace have the same size, which is called the dimension of the subspace. The pivot columns of A give a basis for $\operatorname{im} A$, and the vectors that arise when you compute $\ker A$ in PVF form a basis for $\ker A$. A basis is also important because if you know what a linear transformation does to a basis, then you know the linear transformation! In future chapters, we will see that choosing the right basis is a critical step in many applications of linear algebra.

The rank of a matrix is the dimension of its image. The Rank Theorem says that, for an $n \times k$ matrix A,

$$\operatorname{rank} A + \dim \ker A = k.$$

Chapter 6

Every line through the origin has the form

$$L = \{t\mathbf{v} \mid t \in \mathbf{R}\}$$

Pick a nonzero vector $\mathbf{v}$ *on* L *and replace it with* $\mathbf{v}/|\mathbf{v}|$ *to make it unit length.*

for some unit vector $\mathbf{v}$. The real parameter t really drives home the point that the line L feels like a copy of $\mathbf{R}$ situated in space in a particular way. The origin on L corresponds to $t = 0$, and the sign of t tells you which "side" of L you're on relative to the origin. Since $|t\mathbf{v}| = |t|$, the number $|t|$ tells you how far away you are from the origin.

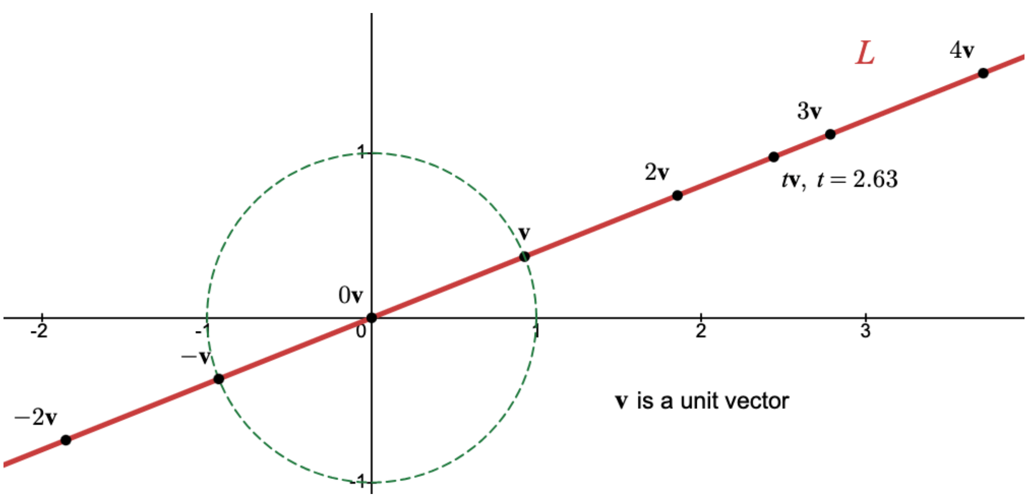


Figure 6.1: A line as a subspace of $\mathbf{R}^2$

We will later formalize this observation by expanding the concept "isomorphism", and we will refer to $\mathbf{R}$ and L as isomorphic. Isomorphic means "the same" in almost every way that matters in linear algebra. Lines such as L, and spans more generally, are called **subspaces** of $\mathbf{R}^n$. What makes L a subspace is that it is "closed" under taking linear combinations; if you take $c, d \in \mathbf{R}$ and

$t_1\mathbf{v}, t_2\mathbf{v} \in L$, then

$$c(t_1\mathbf{v}) + d(t_2\mathbf{v}) = (ct_1 + dt_2)\mathbf{v}$$

is *also* an element of L.

Definition and examples §6.1

Just as the linearity property captures what's important about linear transformations, the following definition captures what's important about subspaces.

Subspaces

Definition 6.1. A subset V of $\mathbf{R}^n$ is called a **subspace** if it satisfies the following two properties:

① V contains the zero vector.

② For all $s, t \in \mathbf{R}$ and $\mathbf{x}, \mathbf{y} \in V$,

$$s\mathbf{x} + t\mathbf{y} \in V.$$

What's a subspace?

Compare this to Definition 2.2.

A few observations:

- For the second item, we could have said:
$$\mathbf{x}, \mathbf{y} \in V \implies \operatorname{span}\{\mathbf{x}, \mathbf{y}\} \subseteq V.$$
- If a subspace contains a nonzero vector, then it contains *the whole line* spanned by that vector.
- If a subspace contains a pair of linearly independent vectors, then it contains contains *the whole plane* spanned those vectors.

Reading Question 6A. Suppose a subspace V of $\mathbf{R}^3$ contains $(0, 1, 2)$ and $(1, 3, 0)$. Find several other vectors that *must* be in V, and find one more vector that may or may not be in V (where there's not enough information to tell).

RQ

Reading Question 6B. Check that $\{\mathbf{0}\}$ and $\mathbf{R}^n$ are subspaces of $\mathbf{R}^n$ (for any n).

RQ

A subspace V is **closed under addition** since (take $s = t = 1$ above)

$$\mathbf{x}, \mathbf{y} \in V \implies \mathbf{x} + \mathbf{y} \in V.$$

What's closure under addition and scalar multiplication?

A subspace V is **closed under scalar multiplication** since (take $t = 0$ above)

$$s \in \mathbf{R}, \mathbf{x} \in V \implies s\mathbf{x} \in V.$$

To check ② in Definition 6.1, you can check closure under addition and scalar multiplication separately if you'd like.

Exercise 6A. Show that if Theorem 6.1 item ② holds and V contains a nonzero vector $\mathbf{x}$, then V contains $\mathbf{0}$. [Hint: build the zero vector using $\mathbf{x}$, scalar multiplication, and addition.] This means when you show V is a subspace, you don't have to prove it contains $\mathbf{0}$, you just have to prove V is nonempty (and verify ②).

The next theorem is crucial. It's an example factory!

Spans are subspaces

You proved this in Exercise 1K! But we should review it.

Theorem 6.2. *The span of a finite set of vectors in* $\mathbf{R}^n$ *is a subspace of* $\mathbf{R}^n$.

Proof. Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a finite set of vectors in $\mathbf{R}^n$ and let $V = \operatorname{span} S$. Since the zero vector is a linear combination of any set of vectors, V certainly contains $\mathbf{0}$.

Next take two vectors in V:

$$\mathbf{x} = s_1\mathbf{v}_1 + \cdots + s_k\mathbf{v}_k$$
$$\mathbf{y} = t_1\mathbf{v}_1 + \cdots + t_k\mathbf{v}_k.$$

Since

$$\mathbf{x} + \mathbf{y} = (s_1 + t_1)\mathbf{v}_1 + \cdots + (s_k + t_k)\mathbf{v}_k,$$

$\mathbf{x} + \mathbf{y} \in V$.

Finally, take $t \in \mathbf{R}$ and $\mathbf{x}$ above. Then,

$$t\mathbf{x} = (ts_1)\mathbf{v}_1 + \cdots + (ts_k)\mathbf{v}_k,$$

so $t\mathbf{x} \in V$. ∎

Example 6.3 (lines and planes). The span of a linearly independent set of one vector is a line, and the span of a linearly independent set of two vectors is a plane. So lines and planes *through the origin* are subspaces of $\mathbf{R}^n$. Lines and planes that *do not* go through the origin are not subspaces because they don't contain $\mathbf{0}$, though each is just a translation away from a parallel line or plane that *does* go through the origin.

Exercise 6B. Let L be the line $3x - 2y = 6$ in $\mathbf{R}^2$. Find a subspace W of $\mathbf{R}^2$ and a vector $\mathbf{b}$ in $\mathbf{R}^2$ such that L is the translation of W by $\mathbf{b}$. (Draw pictures.) Is there more than one way to choose the vector $\mathbf{b}$?

Example 6.4 (the xy-plane)**.** The xy-plane in $\mathbf{R}^3$ is a subspace since it's a plane through the origin; it's the span of $\{\mathbf{e}_1, \mathbf{e}_2\}$:

$$\left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \in \mathbf{R}^3 \;\middle|\; x, y \in \mathbf{R} \right\}.$$

Please understand that this subspace of $\mathbf{R}^3$ is *not* equal to $\mathbf{R}^2$. It is a particular set of 3-vectors, but $\mathbf{R}^2$ is the set of all 2-vectors, so they just can't be equal. We will later see that they are "isomorphic", which means they are structurally identical in almost every meaningful way.

Example 6.5 (kernels)**.** The kernel of a matrix (or of its associated linear transformation) is a subspace since kernels are spans by Theorem 3.11.

Kernels are subspaces.

Exercise 6C. Directly use the definition of subspace and the definition of kernel to prove that $\ker A$ is a subspace without using Theorem 3.11. To do so, take $s, t \in \mathbf{R}$ and $\mathbf{x}, \mathbf{y} \in \ker A$. Write down the equations $\mathbf{x}$ and $\mathbf{y}$ satisfy by definition of kernel. You want to prove that $s\mathbf{x} + t\mathbf{y}$ also lies in the kernel. What equation must you prove is true, again by definition of kernel? Do it!

Example 6.6 (images)**.** If T is a linear transformation with matrix representation $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_k]$, then

Images are subspaces.

$$\begin{aligned} \operatorname{im} T &= \{A\mathbf{x} \mid \mathbf{x} \in \mathbf{R}^k\} \\ &= \{x_1\mathbf{a}_1 + \cdots + x_k\mathbf{a}_k \mid \mathbf{x} = (x_1, \ldots, x_k) \in \mathbf{R}^k\} \\ &= \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_k\}. \end{aligned}$$

So images are spans and hence subspaces. We will also refer to this set as the **image** of A, written $\operatorname{im} A$. Others call this the **column space** of A, written $\operatorname{Col} A$.

The image or column space of a matrix is the span of its columns.

RQ **Reading Question 6C.** Show that

$$\left\{ \begin{bmatrix} 2s - t \\ -s + 5t \\ 7t \end{bmatrix} \middle| s, t \in \mathbf{R} \right\}$$

is a subspace of $\mathbf{R}^3$ by showing that it's the span of a set of vectors. [Hint: write the 3-vector in the set above as s times a vector plus t times a vector.]

Exercise 6D. If A is row equivalent to B, then do they necessarily have the same kernel? How about the same image?

RQ **Reading Question 6D.** Show that

$$\left\{ \begin{bmatrix} 2s - t \\ -s + 5t \\ 7t + 1 \end{bmatrix} \middle| s, t \in \mathbf{R} \right\}$$

is *not* a subspace of $\mathbf{R}^3$ by arguing that it does not contain $(0, 0, 0)$.

Exercise 6E. Determine whether each set below is a subspace of $\mathbf{R}^2$. If the set is not a subspace, determine exactly which of the following fails: (1) containing zero, (2) closure under addition, (3) closure under scalar multiplication.

1. The line $y = 7x$.
2. The line $y = 7x + 1$.
3. $\{(x, y) \mid x \geq 0, y \geq 0\}$ (the first quadrant)
4. $\{(x, y) \mid x = 0 \text{ or } y = 0\}$ (the union of the x- and y-axes)

§6.2 Bases

According to Theorem 2.14, you can always replace a spanning set with a subset that's linearly independent. This is the best kind of spanning set to have, and it's called a basis.

Basis

Definition 6.7. Let V be a subspace of $\mathbf{R}^n$. A set of vectors S in V is called a **basis** if it is (1) linearly independent and (2) spans V.

What's a basis?

The set of vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is linearly independent and spans $\mathbf{R}^n$. It is called the **standard basis** for $\mathbf{R}^n$. This basis has size n, and in fact every basis for $\mathbf{R}^n$ must have size n.

What's the standard basis for $\mathbf{R}^n$?

The size of linearly independent and spanning sets in $\mathbf{R}^n$

Theorem 6.8. *Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbf{R}^n$.*

1. *If S is linearly independent, then $k \le n$.*
2. *If $\operatorname{span} S = \mathbf{R}^n$, then $k \ge n$.*
3. *A basis for $\mathbf{R}^n$ must contain exactly n vectors.*

Proof. Let $A = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k]$. The matrix A is $n \times k$. If S is linearly independent, then A must have a pivot in every column–that's k pivot positions. Since it can't have more than one pivot in any given row, $n \ge k$. If S spans $\mathbf{R}^n$, then A has a pivot in every row–that's n pivot positions. Since it can't have more than one pivot in any given column, $k \ge n$.

For a set of vectors in $\mathbf{R}^n$ to be a basis, both inequalities must hold, forcing $k = n$. ■

Given a set $\{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ of n vectors in $\mathbf{R}^n$, how can we check whether it's a basis? Let $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$. According to The Isomorphism Theorem (4.18), this set is linearly independent if and only if it spans $\mathbf{R}^n$, if and only if A is invertible. So we just need to verify that an REF of A has n pivots, by for example using a computer to compute the RREF of A to check that it's I_n.

How do you check that a set of n vectors in $\mathbf{R}^n$ is a basis?

RQ

Reading Question 6E. Use a computer to check whether

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 5 \\ 9 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 8 \\ 12 \end{bmatrix}, \begin{bmatrix} -5 \\ -4 \\ -3 \\ -2 \end{bmatrix}, \begin{bmatrix} 15 \\ 12 \\ 7 \\ 0 \end{bmatrix} \right\}$$

is a basis for $\mathbf{R}^4$.

We know $\mathbf{R}^n$ has a basis and that every basis for $\mathbf{R}^n$ has size n. What's true for subspaces of $\mathbf{R}^n$? First we will show that every subspace of $\mathbf{R}^n$ *has* a basis.

Extending linearly independent subsets in a subspace

Theorem 6.9. *Let $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be a set of linearly independent vectors in a subspace V of $\mathbf{R}^n$. Unless* $\operatorname{span} S = V$*, there is a vector $\mathbf{v}_{k+1} \in V$ such that $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ is linearly independent.*

This step in the proof shows why subspaces are important: they allow you to build up LI sets in this way.

Proof. Suppose $\operatorname{span} S \neq V$. Then, take *any* vector $\mathbf{v}_{k+1} \in V$ that is NOT in $\operatorname{span} S$. Since V is a subspace, it is closed under taking linear combinations of vectors in V, so the span of $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ is still a subset of V. Further, this set is linearly independent. If it weren't, then we'd have a dependence relation:

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} = \mathbf{0}.$$

We can't have $c_{k+1} = 0$, because if we did, then the equation would *also* give a dependence relation for S (but S is linearly independent!). So, $c_{k+1} \neq 0$ and we can solve for $\mathbf{v}_{k+1}$:

$$\mathbf{v}_{k+1} = \frac{-c_1}{c_{k+1}}\mathbf{v}_1 + \cdots + \frac{-c_k}{c_{k+1}}\mathbf{v}_k.$$

But *this* implies that $\mathbf{v}_{k+1}$ is actually in $\operatorname{span} S$ (a contradiction). So indeed: $\{\mathbf{v}_1, \dots, \mathbf{v}_k, \mathbf{v}_{k+1}\}$ is a linearly independent extension of S in V. ∎

Here's how to show that a subspace V of $\mathbf{R}^n$ must have a basis. If $V = \{\mathbf{0}\}$ then we'll declare its basis to be the empty set since we already established the convention that the empty set is linearly independent and spans $\{\mathbf{0}\}$. If V contains a nonzero vector, then pick one and call it $\mathbf{v}_1$. If $V = \operatorname{span}\{\mathbf{v}_1\}$, then we've found a basis for V (the linearly independent set $\{\mathbf{v}_1\}$). Otherwise, use Theorem 6.9 to obtain a linearly independent set $\{\mathbf{v}_1, \mathbf{v}_2\}$ in V. If $V = \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\}$, then we've found a basis (the set $\{\mathbf{v}_1, \mathbf{v}_2\}$). Otherwise, we can continue to expand our linearly independent set until we eventually find one that spans V; the process must terminate because we can't have more than n linearly independent vectors in $\mathbf{R}^n$ by Theorem 6.8.

We already knew that spans are always subspaces; now we know they're the *only* subspaces.

Every subspace has a basis

Theorem 6.10. *Every subspace of $\mathbf{R}^n$ has a basis.*

Example 6.11 (kernel basis)**.** If $\ker A = \{\mathbf{0}\}$, then the empty set is its basis. Otherwise, the solution set for $A\mathbf{x} = \mathbf{0}$ in PVF yields a spanning set that is always linearly independent (and hence a basis). For example, suppose a matrix A is row equivalent to

See Theorem 3.11.

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

The solution set in PVF is

It would be good practice to check that this is the correct PVF of the solution set.

$$\mathbf{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -5 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_7 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

The four column vectors above form a basis for $\ker A$.

Example 6.12 (image basis)**.** In Theorem 6.13, we will prove that the pivot columns of a matrix A form a basis for $\operatorname{im} A$. For example, the matrix

$$A = \begin{bmatrix} 1 & 2 & 2 & 12 & 7 & 5 & -5 \\ 2 & 4 & 5 & 28 & 18 & 13 & -14 \\ -2 & -4 & -4 & -24 & -13 & -9 & 9 \\ -1 & -2 & 2 & 4 & 8 & 6 & -10 \end{bmatrix}$$

Warning: do not take pivot columns from a REF of A, take the columns from the original matrix A!

has RREF

$$B = \begin{bmatrix} 1 & 2 & 0 & 4 & 0 & 0 & 2 \\ 0 & 0 & 1 & 4 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

This means the first, third, and fifth columns of A are pivot columns. Hence, the set

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ -2 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ -4 \\ 2 \end{bmatrix}, \begin{bmatrix} 7 \\ 18 \\ -13 \\ 8 \end{bmatrix} \right\}$$

is a basis for $\operatorname{im} A$.

RQ

Reading Question 6F. In Example 6.12, you really do need to take the pivot columns of A and not B. Show that A and B do not have the same image by finding a vector that's in the image of one but not the other.

Kernel and image bases

How do we find bases for the kernel and image of a matrix?

Theorem 6.13. *Let A be an $n \times k$ matrix.*

1. *The vectors that appear next to the free variables in the PVF of the solution set for $A\mathbf{x} = \mathbf{0}$ form a basis for* $\ker A$.
2. *The pivot columns of A form a basis for* $\operatorname{im} A$.

Proof. Item (1) is just a restatement of Theorem 3.11.

You can look back at A and B in Example 6.12 as you read this proof.

For item (2), let B be the RREF of A. Each pivot column of B is one of the standard basis vectors $\mathbf{e}_i$, and they are all distinct, so they are certainly linearly independent. Since the matrix containing just the pivot columns of A is row equivalent to the matrix containing the just the pivot columns of B, the pivot columns of A are linearly independent. So if we can show that the non-pivot columns of A are in the span of the pivot columns of A, then the proof will be complete.

Make a matrix from the pivot columns of A in Example 6.12 and use a computer to RREF it.

Suppose $\mathbf{a}_i$ is a non-pivot column in A. Then, $\mathbf{b}_i$ is a non-pivot column of B. If the jth entry in $\mathbf{b}_i$ is nonzero, then it's to the right a pivot. Put another way, if the jth entry of $\mathbf{b}_i$ is nonzero, then $\mathbf{e}_j$ is a pivot column of B. From this it follows that $\mathbf{b}_i$ is in the span of the pivot columns of B. This means that there is a dependence relation

Look again at B in Example 6.12. Check that each non-pivot column is in the span of the pivot columns.

$$c_1\mathbf{b}_1 + \cdots + c_i\mathbf{b}_i + \cdots + c_k\mathbf{b}_k = \mathbf{0}$$

where $c_i \neq 0$ and *all the weights in front of the other non-pivot columns of B are zero.* But the equations $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have the same solution set (A and B are row equivalent), so there is a dependence relation

$$c_1\mathbf{a}_1 + \cdots + c_i\mathbf{a}_i + \cdots + c_k\mathbf{a}_k = \mathbf{0}$$

where, again, $c_i \neq 0$ and all the weights in front of the other non-pivot columns in A are zero. This means the non-pivot column $\mathbf{a}_i$ of A is in the span of A's pivot columns, completing the proof. ■

Exercise 6F. Make up a 5 by 6 matrix and use the computer to compute its RREF. Use this to find a basis for the kernel and image. Again: make sure you draw the basis vectors for the image from the columns of the original matrix!

Exercise 6G. Let

$$\mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} \right\}.$$

Find a matrix whose image has basis $\mathcal{B}$. Now find a matrix whose kernel has basis $\mathcal{B}$.

Exercise 6H. Let P be the plane in $\mathbf{R}^3$ spanned by

$$\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \text{ and } \mathbf{w} = \begin{bmatrix} 0 \\ 3 \\ -4 \end{bmatrix}.$$

Find a vector $\mathbf{r}$ such that $\{\mathbf{v}, \mathbf{w}, \mathbf{r}\}$ is a basis for $\mathbf{R}^3$. Let T be the linear transformation defined by $T(\mathbf{e}_1) = \mathbf{v}$, $T(\mathbf{e}_2) = \mathbf{w}$, and $T(\mathbf{e}_3) = \mathbf{r}$. Is T invertible? Note that T maps the xy-plane onto the plane P (do you see why?). Where does T^{-1} map P? Now, find a linear transformation T' that maps the xy-plane onto the plane P' spanned by

$$\begin{bmatrix} 0 \\ 5 \\ -1 \end{bmatrix} \text{ and } \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Can you use T and T' to cook up a linear transformation that maps P onto P'?

Exercise 6I. Let

$$W = \left\{ \begin{bmatrix} 2s + 3t - 4r \\ -s - t + 2r \end{bmatrix} \middle| r, s, t \in \mathbf{R} \right\}.$$

Find a basis for W.

Bases are also valuable because a linear transformation is completely determined by its values on the elements of any given basis. To see why, suppose $T\colon \mathbf{R}^n \to \mathbf{R}^k$ is a linear transformation and $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a basis for $\mathbf{R}^n$. Given any $\mathbf{x} \in \mathbf{R}^n$, we know we can write $\mathbf{x}$ as a linear combination of the basis vectors:

A linear transformation is determined by what it does to a basis.

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n.$$

Applying T to both sides of the equation and using the fact that T is linear, we obtain

$$T(\mathbf{x}) = c_1 T(\mathbf{b}_1) + \cdots + c_n T(\mathbf{b}_n).$$

So if you know the values $T(\mathbf{b}_1), \ldots, T(\mathbf{b}_n)$, you can compute the value of T at any vector in $\mathbf{R}^n$. If A is the matrix represenation of T, then this says

$$A\mathbf{x} = c_1 A\mathbf{b}_1 + \cdots + c_n A\mathbf{b}_n.$$

This is a key exercise for later work; we had better do it!

Exercise 6J. Let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

The set $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis for $\mathbf{R}^2$ (what's the easiest way to check this?). Let

$$A = \begin{bmatrix} 4 & 8 \\ 8 & 4 \end{bmatrix}.$$

Compute $A\mathbf{v}_i$ for each i, and in each case find a scalar λ_i such that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$. Finally, using the ideas discussed before this exercise, compute

$$A^{11} \begin{bmatrix} 9 \\ -1 \end{bmatrix}$$

without computing the matrix A^{11}. [Hint: for the last part, write $(9, -1)$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$.]

§6.3 Dimension

Except for the subspace $\{\mathbf{0}\}$, subspaces have a staggering number of bases. For example, the columns of any invertible $n \times n$ matrix form a basis for $\mathbf{R}^n$ because, according to the Isomorphism Theorem (4.18), the columns of such a matrix are both linearly independent and span $\mathbf{R}^n$. But must all bases for fixed subspace have the same size? The answer is *yes*.

All bases of a given subspace have the same size

Theorem 6.14. *Let V be a subspace of $\mathbf{R}^n$. Any two bases for V must have the same number of vectors.*

This seems intuitively reasonable; for example, it doesn't seem that a plane can be spanned with only one vector, and it seems like you can't have three linearly independent vectors trapped in a plane because their span should have

3-dimensional thickness. We won't prove this theorem now. It will be easier to prove later, after we learn about the coordinate mapping in §7.1. Given Theorem 6.14, we will define the **dimension** of a subspace V, written $\dim V$, to be the size of any basis for V.

What's the dimension of a subspace?

Reading Question 6G. In terms of free variables, pivot variables, pivot columns, etc. describe $\dim \ker A$ and $\dim \operatorname{im} A$ when A is an $n \times k$ matrix.

RQ

Based on the work we've done so far, we can say that lines through the origin have dimension 1, planes through the origin have dimension 2, and $\dim \mathbf{R}^n = n$. The dimension of $\{\mathbf{0}\}$ is zero since its basis is the empty set, which has zero elements. Here is another key property of dimension.

The dimension inequality

Theorem 6.15. *Suppose V and W are subspaces of $\mathbf{R}^n$ and $W \subseteq V$. Then,*

$$\dim W \leq \dim V.$$

The dimension of a subspace is at least as big as the dimension of any subspace it contains.

Proof. Let $\mathcal{B}$ be a basis for W. By repeatedly using Theorem 6.9, we can extend this basis to a basis for V. By definition of dimension, the inequality must hold. ∎

Reading Question 6H. Use Theorem 6.15 to list the possible dimensions of a subspace of $\mathbf{R}$. For each possible dimension on your list, find all subspaces of $\mathbf{R}$ of that dimension.

RQ

Exercise 6K. Let V and W be k-dimensional subspaces of $\mathbf{R}^n$ with $W \subseteq V$. Is it possible that $V \neq W$?

Exercise 6L. Describe the possible subspaces of $\mathbf{R}^n$ for $n = 2, 3$, stratified by dimension.

Rank §6.4

An $n \times k$ matrix has n rows; each of these rows may be viewed as as a vector in $\mathbf{R}^k$ because each row vector has k entries. The span of the rows of A is called the **row space** of A, written row A. The row space of A is the same as the image of A^{T}:

$$\operatorname{row} A = \operatorname{im} A^{\mathrm{T}}.$$

Basis for the row space

Theorem 6.16. *The nonzero rows in any REF of a matrix A form a basis for* row A.

Proof. Since pivots stagger down and to the right, the nonzero rows of a matrix in REF are linearly independent. To prove the theorem, it suffices to prove that row equivalent matrices have the same row space. To do that, we just need to check that executing a row operation does not change the row space. It's clear that swapping rows won't change their span; you'll do the rest in an exercise below. ∎

Exercise 6M. To complete the proof of Theorem 6.16, do the following things. Let $R = \{\mathbf{r}_1, \ldots, \mathbf{r}_n\}$ be a set of vectors in $\mathbf{R}^k$. Show that R has the same span as each of the following sets. In both cases we apply a row operation to $\mathbf{r}_1$; a similar proof to the one you discover would work for any row.

- $\{c\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$ (c a nonzero real number)
- $\{\mathbf{r}_1 + c\mathbf{r}_j, \mathbf{r}_2, \ldots, \mathbf{r}_n\}$ (c a real number and $j \neq 1$)

The dimension of row A is the number of nonzero rows in a REF of A, but that's just the number of pivot positions in A. The number of pivot positions in A is also the number of pivot columns, and the number of pivot columns is the dimension of im A. Since these dimensions are always the same, we give them a common name.

The rank of a matrix

What's the rank of a matrix?

Definition 6.17. Let A be an $n \times k$ matrix. Define the **rank** of A by

$$\operatorname{rank} A = \dim \operatorname{im} A = \dim \operatorname{row} A.$$

But wait, there's more! The number of pivot columns in A is the number of pivot variables in the linear system $A\mathbf{x} = \mathbf{0}$, and the number of free variables is the number of non-pivot columns. So, the number of free variables is

$$k - \dim \operatorname{im} A = k - \operatorname{rank} A.$$

The dimension of the kernel of A is *also* the number of free variables, so we have the following theorem.

The Rank Theorem

Theorem 6.18. *For any $n \times k$ matrix A,*

$$\operatorname{rank} A + \dim \ker A = k.$$

$\dim \ker A$ *is sometimes called the nullity of A.*

Reading Question 6I. If A is an $n \times n$ matrix, then what is $\operatorname{rank} A$ if A is invertible? What is $\dim \ker A$ when A is invertible? RQ

Reading Question 6J. Suppose a 6×8 matrix A has four pivot columns. What is $\dim \ker A$? What is the rank of A? Is $\operatorname{im} A = \mathbf{R}^4$? Why or why not? RQ

Exercise 6N. Suppose $T : \mathbf{R}^3 \to \mathbf{R}^3$ is a linear transformation. Geometrically, what are the qualitative possibilities for the value of $\ker T$? In each case, what can you say (geometrically) about $\operatorname{im} T$?

Exercise 6O. Let A be $n \times k$ and let B be $k \times p$. Explain why $\ker B$ is contained in $\ker AB$. What does this tell you about the dimensions of these two subspaces? Now draw a conclusion about the relationship between the rank of B and the rank of AB using the rank theorem.

Exercise 6P. Suppose you have a homogeneous linear system with 40 equations and 42 variables. You have found two solutions that are not scalar multiples of one another, and all solutions are linear combinations of the two solutions you found. Can you be sure that any associated non-homogeneous system will have a solution?

Exercise 6Q. Suppose A is an $n \times n$ matrix such that some power of A is the zero matrix. What's the largest the rank of A could be? Find a 3×3 matrix A such that some power of A is zero and the rank of A is as large as possible.

Example 6.19 (visualizing ker, im, row)**.** Suppose

$$A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = B.$$

Then:

$$\ker A = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} = y\text{-axis}$$

$$\operatorname{im} A = \operatorname{span}\left\{ \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

$$\operatorname{row} A = \operatorname{span}\left\{ \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} = xz\text{-plane}$$

In the Figure 6.2, $\operatorname{row} A \perp \ker A$ (in the domain of $\mathbf{x} \mapsto A\mathbf{x}$):

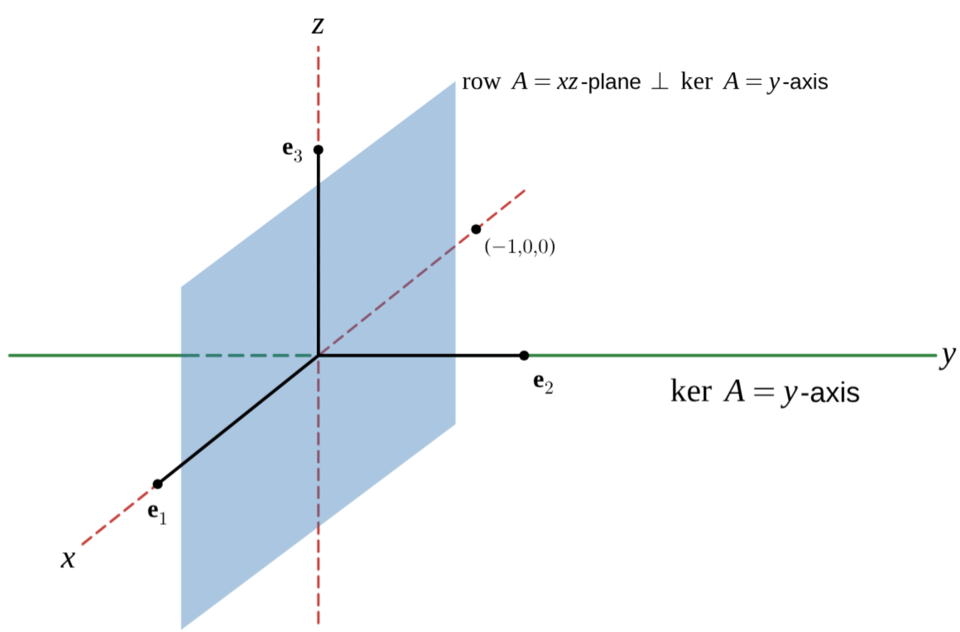


Figure 6.2: $\operatorname{row} A \perp \ker A$

In the Figure 6.3, $\operatorname{im} A = \operatorname{row} A^{\mathrm{T}} \perp \ker A^{\mathrm{T}}$ (in the codomain):

$$A^{\mathrm{T}} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \ \ker A^{\mathrm{T}} = \operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

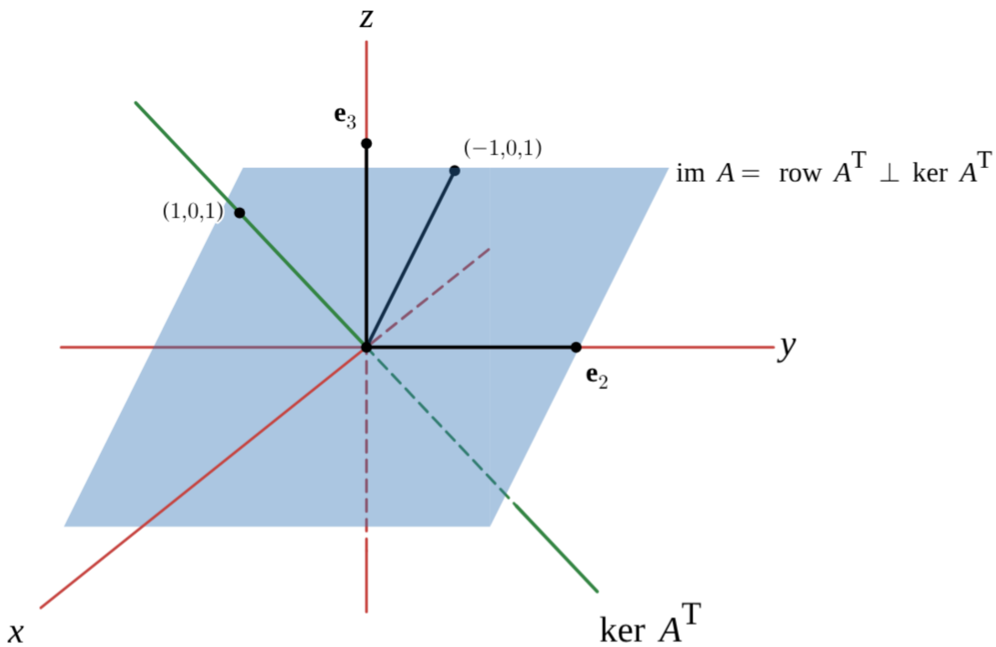


Figure 6.3: $\operatorname{im} A \perp \ker A^{\mathrm{T}}$

Chapter 7 Coordinate systems in $\mathbf{R}^n$

Key concepts

- Coordinate vectors $[\mathbf{x}]_{\mathcal{B}}$ relative to a basis $\mathcal{B}$
- The span and coordinate mappings as inverse linear transformations
- The transformation $\mathbf{x} \mapsto A\mathbf{x}$ has $\mathcal{B}$-matrix $B = P^{-1}AP$, where the columns of P are the basis vectors in $\mathcal{B}$
- The $\mathcal{B}$-matrix B tells us how A transforms coordinate vectors:

$$B[\mathbf{x}]_{\mathcal{B}} = [A\mathbf{x}]_{\mathcal{B}}$$

- Change of basis can simplify computations, especially when the $\mathcal{B}$-matrix is diagonal

Summary. This chapter is about representing vectors and linear transformations using different coordinate systems. Given a basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for $\mathbf{R}^n$, every vector $\mathbf{x}$ can be uniquely written as a linear combination $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. The tuple $(c_1, \ldots, c_n)$ is called the $\mathcal{B}$-coordinate vector of $\mathbf{x}$, denoted $[\mathbf{x}]_{\mathcal{B}}$.

There are two important related linear transformations. The *span mapping* takes a coordinate vector and returns the original vector; the span mapping is represented by the matrix $P_{\mathcal{B}}$ whose columns are the vectors in $\mathcal{B}$. The *coordinate mapping* is its inverse. It takes a vector in $\mathbf{R}^n$ and spits out its $\mathcal{B}$-coordinate vector; the coordinate mapping is represented by the matrix $P_{\mathcal{B}}^{-1}$.

Matrix representations of linear transformations also change when we change coordinates. If a transformation is represented by a matrix A in the standard basis, then in $\mathcal{B}$-coordinates it is represented by the similar matrix $B = P^{-1}AP$. This leads to the formula

$$B[\mathbf{x}]_{\mathcal{B}} = [A\mathbf{x}]_{\mathcal{B}},$$

which shows what the transformation does to $\mathcal{B}$-coordinate vectors.

This idea is especially useful when B is a diagonal matrix, since computing powers of A and related computations become much easier. Choosing a good basis is often the key to simplifying a linear algebra problem.

Chapter 7

Coordinate systems §7.1

Warming up in the plane §7.1.1

When it comes to bases, the order in which the vectors are listed matters, so you should think of $\{\mathbf{v},\mathbf{w}\}$ and $\{\mathbf{w},\mathbf{v}\}$ as different bases.

Let's examine two different bases for $\mathbf{R}^2$: the standard basis

$$\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$$

and another basis

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix} \right\}.$$

Consider the vector (say) $\mathbf{q} = (7, 4)$ in $\mathbf{R}^2$. Usually, we think of this vector in "$\mathcal{E}$-coordinates" (or "standard coordinates"); this means that, of all the ways we might express this vector as a linear combination of basis vectors, the one that is foremost in our mind is

$$\begin{bmatrix} 7 \\ 4 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 4 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 7\mathbf{e}_1 + 4\mathbf{e}_2.$$

We can view the numbers 7 and 4 as "instructions" for obtaining $\mathbf{q}$ as a linear combination of vectors in $\mathcal{E}$.

However, $\mathcal{B}$ is *also* a basis; so we know that there are unique scalars c_1 and c_2 such that

$$\begin{bmatrix} 7 \\ 4 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Solve this equation. You should find that $c_1 = 2$ and $c_2 = 5/2$:

$$\begin{bmatrix} 7 \\ 4 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + (5/2) \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

These values of c_1 and c_2 are instructions for obtaining **q** as a linear combination of vectors in $\mathcal{B}$; in particular, they are the required weights in the linear combination. These numbers c_1 and c_2 are called the **$\mathcal{B}$-coordinates** of **q**; when we put these coordinates together in a single vector, we obtain the so-called **$\mathcal{B}$-coordinate vector of q**, which we denote as $[\mathbf{q}]_{\mathcal{B}}$:

$$[\mathbf{q}]_{\mathcal{B}} = \begin{bmatrix} 7 \\ 4 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 2 \\ 5/2 \end{bmatrix}.$$

Geometrically, the $\mathcal{E}$-coordinates of **q** tell us how to reach **q** from **0** by moving in the $\mathbf{e}_1$ and $\mathbf{e}_2$ directions (that is, to the right and up). Similarly, the $\mathcal{B}$-coordinates of **q** tell us how to reach **q** from **0** by moving in the $(1, 2)$ and $(2, 0)$ directions! Let's try to visualize the situation. You may find this Desmos illustration helpful. In Figure 7.1 below, our point **q** is the big blue dot. The vectors $\mathbf{v} = (1, 2)$ and $\mathbf{w} = (2, 0)$ are also marked—find them. The thin black lines are the standard coordinate gridlines, and the thicker lines are gridlines determined by **v** and **w**.

www.desmos.com/calculator/dlmcsohuy8

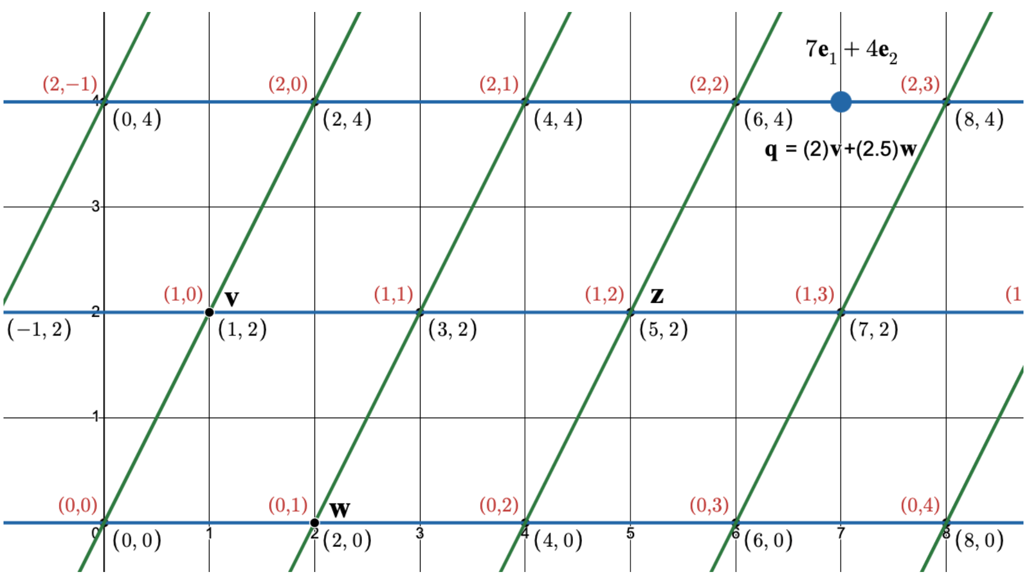


Figure 7.1: $\mathcal{E}$-coordinates (black) and $\mathcal{B}$-coordinates (red)

To see what we mean by the gridlines determined by **v** and **w**, find the point labeled both $(1, 2)$ and $(5, 2)$ (let's call it **z**). The pair $(5, 2)$ gives the location of this point using the standard coordinate system, but the pair $(1, 2)$ gives the location in $\mathcal{B}$-coordinates. This is because:

$$\begin{bmatrix} 5 \\ 2 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

To get to $\mathbf{z}$, you start at the origin and follow $\mathbf{v}$ once (up and to the right). Then, you move in the $\mathbf{w}$ direction and follow $\mathbf{w}$ to the right twice. Trace this path with your finger in Figure 7.1. We know the $\mathcal{B}$-coordinate vector of $\mathbf{q}$ is $(2, 2.5)$, so to get to $\mathbf{q}$ from $\mathbf{0}$ you move in the $\mathbf{v}$ direction twice and then in the $\mathbf{w}$ direction 2.5 times. Trace this path with your finger as well.

Let's generalize the above example. Define the function

$$\Phi : \mathbf{R}^2 \to \mathbf{R}^2$$

by the formula

$$\Phi(c_1, c_2) = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

We will call this map the **$\mathcal{B}$-span mapping** because each input is a vector of weights that Φ maps to the corresponding element in the span of $\mathcal{B}$.

What's the span mapping?

Although strictly speaking the domain and codomain of Φ are the same (namely, the set $\mathbf{R}^2$), you should think of the domain of Φ as where $\mathcal{B}$-coordinate vectors live, and the codomain of Φ as where $\mathcal{E}$-coordinate vectors live. Here's an alternative formula for Φ that emphasizes this way of thinking:

$$\Phi\left([\mathbf{x}]_{\mathcal{B}}\right) = \mathbf{x}.$$

Now, observe that

$$\Phi\left(\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}\right) = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Thus we see that Φ is actually a linear transforamtion, with matrix representation

$$P_{\mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}.$$

Since the columns of $P_{\mathcal{B}}$ are independent (they are the members of $\mathcal{B}$), we conclude that the matrix $P_{\mathcal{B}}$, and hence the linear transformation Φ, is invertible. Conceptually, it is clear what Φ^{-1} has to do: since

$$\Phi\left([\mathbf{x}]_{\mathcal{B}}\right) = \mathbf{x},$$

we must have

$$\Phi^{-1}(\mathbf{x}) = [\mathbf{x}]_{\mathcal{B}}.$$

In other words, Φ^{-1} *takes a vector to its $\mathcal{B}$-coordinate vector*; we call Φ^{-1} the **$\mathcal{B}$-coordinate mapping**. In still other words, we can find $[\mathbf{x}]_{\mathcal{B}}$ by multiplying $\mathbf{x}$ by the matrix $P_{\mathcal{B}}^{-1}$.

What's the coordinate mapping?

Let's see all this working with our example vector $\mathbf{q}$. We have

$$\Phi\left(\begin{bmatrix} 7 \\ 4 \end{bmatrix}_{\mathcal{B}}\right) = \Phi\left(\begin{bmatrix} 2 \\ 5/2 \end{bmatrix}\right) = P_{\mathcal{B}} \begin{bmatrix} 2 \\ 5/2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 5/2 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \end{bmatrix}$$

$\mathcal{B}$- to $\mathcal{E}$-coords.

and

$\mathcal{E}$- to $\mathcal{B}$-coords.

$$\begin{bmatrix}7\\4\end{bmatrix}_{\mathcal{B}} = \Phi^{-1}\left(\begin{bmatrix}7\\4\end{bmatrix}\right) = P_{\mathcal{B}}^{-1}\begin{bmatrix}7\\4\end{bmatrix} = \begin{bmatrix}0 & 1/2\\1/2 & -1/4\end{bmatrix}\begin{bmatrix}7\\4\end{bmatrix} = \begin{bmatrix}2\\5/2\end{bmatrix}.$$

To summarize with a diagram:

$$\mathbf{R}^2 = \left\{\substack{\mathcal{B}\text{−coordinate}\\ \text{vectors}}\right\} \underset{P_{\mathcal{B}}^{-1}}{\overset{P_{\mathcal{B}}}{\rightleftarrows}} \left\{\substack{\mathcal{E}\text{−coordinate}\\ \text{vectors}}\right\} = \mathbf{R}^2$$

§7.1.2 *The general situation*

Everything we did above works equally well in $\mathbf{R}^n$—there are just more coordinates. Given a basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ for $\mathbf{R}^n$, we'll call the map $\Phi\colon \mathbf{R}^n \to \mathbf{R}^n$ defined by

$$\Phi(x_1, \ldots, x_n) = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$$

the $\mathcal{B}$**-span mapping**. As above, Φ is invertible because its matrix representation

$$P_{\mathcal{B}} = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n]$$

is invertible. The inverse of the $\mathcal{B}$-span mapping is called the $\mathcal{B}$**-coordinate mapping**.

$\mathcal{B}$-COORDINATES FOR $\mathbf{R}^n$

What's a $\mathcal{B}$-coordinate vector?

Definition 7.1. Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for $\mathbf{R}^n$ and let $P_{\mathcal{B}} = [\mathbf{b}_1 \ \ldots \ \mathbf{b}_n]$. Then, for any $\mathbf{x}$ in $\mathbf{R}^n$, the $\mathcal{B}$**-coordinate vector of x**, written $[\mathbf{x}]_{\mathcal{B}}$, is defined by the equivalence

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix} \iff \mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n = P_{\mathcal{B}}\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}.$$

In particular, we have

$$[\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}^{-1}\mathbf{x}$$

and

$$P_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = \mathbf{x}.$$

This entire discussion requires you to pick a particular basis $\mathcal{B}$ for $\mathbf{R}^n$. If we want to emphasize the basis, we will say "$\mathcal{B}$-span mapping" and "$\mathcal{B}$-coordinate mapping". If the basis is understood (clear from context) then we can just say "span mapping" and "coordinate mapping".

Reading Question 7A. Find $[\mathbf{x}]_{\mathcal{B}}$, where

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\} \text{ and } \mathbf{x} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}.$$

RQ

Reading Question 7B. Find $\mathbf{x}$, where

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\} \text{ and } [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}.$$

RQ

Reading Question 7C. Find $[\mathbf{x}]_{\mathcal{B}}$, where

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\} \text{ and } \mathbf{x} = \begin{bmatrix} 6 \\ 1 \\ -3 \end{bmatrix}.$$

RQ

Reading Question 7D. Find $\mathbf{x}$, where

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\} \text{ and } [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 6 \\ 1 \\ -3 \end{bmatrix}.$$

RQ

Exercise 7A. Suppose $\mathcal{B} = \{\mathbf{v}, \mathbf{w}\}$ is a basis for $\mathbf{R}^2$. Suppose $[2\mathbf{e}_1]_{\mathcal{B}} = (-2, 4)$ and $[\mathbf{e}_1 + \mathbf{e}_2]_{\mathcal{B}} = (2, -3)$. Find the basis $\mathcal{B}$.

Exercise 7B. This exercise provides a first illustration of how changing coordinates can be useful in data analysis. While this exercise is very simple (and artificial), it illustrates a widely used and important idea.

Suppose that we are running a video streaming service, and we want to recommend movies to viewers based on their past ratings of movies they've watched. For each movie in our service we have two pieces of information: x, the number of car crashes in the movie; and y, the number of kissing scenes in the movie.

Suppose that a viewer has rated five movies on our service as follows (higher ratings means the viewer liked the movie better):

x	y	rating r
5	0	2
10	0	4
5	5	5
8	6	8
13	6	10

You should think of each movie as being described by the vector (x, y). We want to use a movie's vector (x, y) to predict the user's rating.

(a) Make plots of r vs. x and r vs. y.

(b) For each of the five movies above, find the coordinate vector

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}_{\mathcal{B}}$$

with respect to the basis

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}.$$

(c) Make plots of r vs. c_1 and r vs. c_2.

(d) Given your work so far, how do you think this user would rate a movie with an x-value of 2 and a y-value of 3?

(e) Explain what this exercise has to do with changing coordinates.

§7.2 Dynamical systems and choice of coordinates

Go back and look at Exercise 6J. In that exercise, we defined the matrix

$$A = \begin{bmatrix} 4 & 8 \\ 8 & 4 \end{bmatrix}$$

and sought to calculate

$$A^{11} \begin{bmatrix} 9 \\ -1 \end{bmatrix}.$$

Now, trying to do this calculation by performing eleven matrix-by-vector calculations would be a nightmare. In Exercise 6J, though, we saw a way to do this calculation that was quite easy (and also gave us more insight than merely having a computer do it!). The idea was to use a basis of $\mathbf{R}^2$ that was well-suited to the problem — in particular, a basis of vectors whose products by A are easy to compute.

Let's consider a similar example here, expressed in the language of dynamical systems (what happens next should remind you of what happened when we last looked at lionfish, Fibonacci numbers, and directed graphs). Let A be as above, and suppose we are interested in the discrete dynamical system $\mathbf{x} \mapsto A\mathbf{x}$ with initial condition $\mathbf{x}_0 = (8, -2)$ — so we want to examine $A^k\mathbf{x}_0$ for all positive integers k.

As in Exercise 6J, let's use the basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

We leave it to you to verify that

$$\mathbf{x}_0 = \begin{bmatrix} 8 \\ -2 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 5\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 3\mathbf{v}_1 + 5\mathbf{v}_2.$$

We claim that $\mathcal{B}$ is a "good" basis to use for this problem. To see why, observe that

$$A\mathbf{v}_1 = \begin{bmatrix} 4 & 8 \\ 8 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 12 \\ 12 \end{bmatrix} = 12\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 12\mathbf{v}_1$$

and

$$A\mathbf{v}_2 = \begin{bmatrix} 4 & 8 \\ 8 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \end{bmatrix} = -4\begin{bmatrix} 1 \\ -1 \end{bmatrix} = -4\mathbf{v}_2.$$

This is helpful because, whenever we have $A\mathbf{v} = \lambda\mathbf{v}$, we can easily compute what happens to $\mathbf{v}$ when we repeatedly apply A k times:

$$\begin{aligned} A\mathbf{v} &= \lambda\mathbf{v}; \\ A^2\mathbf{v} &= A(\lambda\mathbf{v}) \\ &= \lambda A\mathbf{v} \\ &= \lambda^2\mathbf{v}; \\ A^3\mathbf{v} &= A(\lambda^2\mathbf{v}) \\ &= \lambda^2 A\mathbf{v} \\ &= \lambda^3\mathbf{v}; \\ &\vdots \\ A^k\mathbf{v} &= \lambda^k\mathbf{v}. \end{aligned}$$

If $A\mathbf{v} = \lambda\mathbf{v}$, then $\mathbf{x}_k = \lambda^k\mathbf{v}$ is a solution to the dynamical system $\mathbf{x} \mapsto A\mathbf{x}$ with initial condition $\mathbf{v}$.

With all these observations in hand, calculating $A^k\mathbf{x}_0$ becomes a simple application of linearity:

$$\begin{aligned} A^k\mathbf{x}_0 = A^k\begin{bmatrix} 8 \\ -2 \end{bmatrix} &= A^k(3\mathbf{v}_1 + 5\mathbf{v}_2) \\ &= 3A^k\mathbf{v}_1 + 5A^k\mathbf{v}_2 \\ &= 3 \cdot (12)^k\mathbf{v}_1 + 5 \cdot (-4)^k\mathbf{v}_2 \end{aligned}$$

$$= 3 \cdot (12)^k \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 5 \cdot (-4)^k \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$
$$= \begin{bmatrix} 3 \cdot (12)^k + 5 \cdot (-4)^k \\ 3 \cdot (12)^k - 5 \cdot (-4)^k \end{bmatrix}.$$

The point is that we are able to compute the solution of the discrete dynamical system $\mathbf{x} \mapsto A\mathbf{x}$ with initial condition $(8, -2)$ *without computing the matrices* A^k. We repeat: this worked because we were able to find a special basis for $\mathbf{R}^2$ that is related to A in a very nice way.

RQ

Reading Question 7E. Find (by hand)

$$A^{10} \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad \text{where } A = \begin{bmatrix} 4 & 8 \\ 8 & 4 \end{bmatrix}.$$

[Hint: first write $(0, 2)$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ and then reread the material above.]

Here's a question: if we start with the *coordinate vector* of $\mathbf{x}$, what does the *coordinate vector* of $A\mathbf{x}$ look like? To see the answer, we can take the following steps.

Start with the coordinate vector:	$[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{B}}$.
Apply the span mapping:	$P[\mathbf{x}]_{\mathcal{B}} = \mathbf{x}$.
Perform the linear transformation:	$AP[\mathbf{x}]_{\mathcal{B}} = A\mathbf{x}$.
Get the coordinate vector of the output:	$P^{-1}AP[\mathbf{x}]_{\mathcal{B}} = [A\mathbf{x}]_{\mathcal{B}}$.

We therefore see that the matrix $P^{-1}AP$ transforms the *coordinate vector* of $\mathbf{x}$ into the *coordinate vector* of $A\mathbf{x}$.

RQ

Reading Question 7F. Memorize that last sentence and say it to yourself five times with your eyes closed. [We're serious. Do it. Right now.]

This process is illustrated in the following diagram. We have just described how to move from the top-left of the diagram ($\mathcal{B}$-coordinates of inputs) to the bottom-left of the diagram ($\mathcal{B}$-coordinates of outputs) by performing P, and then A, and then P^{-1}.

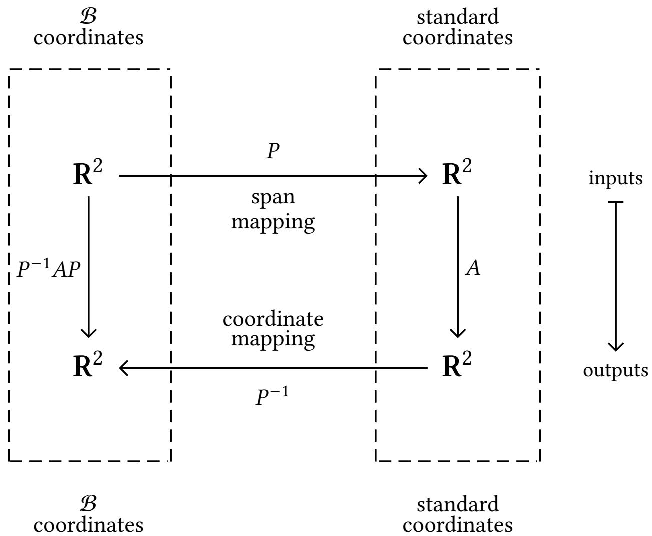


Figure 7.2: The Similarity Diagram

The point of all of this is that, as complicated as this procedure initially seems, the linear transformation $P^{-1}AP$ might really be quite simple — so simple that using it to calculate A (and the powers of A) is far preferable.

Let's figure out what $P^{-1}AP$ is in our particular example by tracking what happens to $\mathbf{e}_1$ and $\mathbf{e}_2$. Since

$$\mathbf{v}_1 = (1)\mathbf{v}_1 + (0)\mathbf{v}_2,$$

the coordinate mapping takes $\mathbf{v}_1$ to $\mathbf{e}_1 = (1, 0)$. The span mapping does the reverse. Similarly, the coordinate mapping takes $\mathbf{v}_2$ to $\mathbf{e}_2$. In terms of the representing matrices, we have

$$P^{-1}\mathbf{v}_i = [\mathbf{v}_i]_{\mathcal{B}} = \mathbf{e}_i, \quad P\mathbf{e}_i = P[\mathbf{v}_i]_{\mathcal{B}} = \mathbf{v}_i.$$

So:

$$\mathbf{e}_1 = [\mathbf{v}_1]_{\mathcal{B}} \xmapsto{P} \mathbf{v}_1 \xmapsto{A} 12\mathbf{v}_1 \xmapsto{P^{-1}} 12\mathbf{e}_1 = \begin{bmatrix} 12 \\ 0 \end{bmatrix}$$

$$\mathbf{e}_2 = [\mathbf{v}_2]_{\mathcal{B}} \xmapsto{P} \mathbf{v}_2 \xmapsto{A} -4\mathbf{v}_2 \xmapsto{P^{-1}} -4\mathbf{e}_2 = \begin{bmatrix} 0 \\ -4 \end{bmatrix},$$

and thus

$$P^{-1}AP = \begin{bmatrix} 12 & 0 \\ 0 & -4 \end{bmatrix}.$$

Let us name this matrix D (for “diagonal”):

$$P^{-1}AP = \begin{bmatrix} 12 & 0 \\ 0 & -4 \end{bmatrix} = D.$$

We defined similar matrices in Chapter 5.

Observe that A and D are *similar.*

If we multiply both sides of the equation $P^{-1}AP = D$ by P^{-1} on the right, we obtain $P^{-1}A = DP^{-1}$ and hence that

$$P^{-1}A\mathbf{x} = DP^{-1}\mathbf{x}$$

for any vector $\mathbf{x}$. Since P^{-1} is the matrix representation of the coordinate mapping, this last equation can be written

$$[A\mathbf{x}]_{\mathcal{B}} = D[\mathbf{x}]_{\mathcal{B}}.$$

This says, again, that we obtain the *coordinate vector* of $A\mathbf{x}$ by multiplying the *coordinate vector* of $\mathbf{x}$ by the matrix D. Mathematicians express this idea by saying that if we choose to “work in the coordinate system determined by $\mathcal{B}$,” then the matrix D represents our linear transformation $\mathbf{x} \mapsto A\mathbf{x}$.

Geometrically, the similarity relationship $P^{-1}AP = D$ says that if you work in the $\mathcal{B}$-coordinate plane, where $(1, 0)$ corresponds to $\mathbf{v}_1$ and $(0, 1)$ corresponds to $\mathbf{v}_2$, then the transformation scales by 12 in the “x”-direction, reflects over the “x”-axis, and then scales in the “y”-direction by 4. You just have to pretend the axes are the lines spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$, and the tick marks are at the integer multiples of these vectors.

In the Figure 7.3 below, we put gridlines parallel to $\mathbf{v}_1$ and $\mathbf{v}_2$ that intersect at the points $4m\mathbf{v}_1 + 4n\mathbf{v}_2$ (where m, n are integers). We plotted the point $Q = \mathbf{v}_1 + \mathbf{v}_2$ with $\mathcal{B}$-coordinates $(1, 1)$. The $\mathcal{B}$-coordinates of AQ are given by

$$[AQ]_{\mathcal{B}} = D[Q]_{\mathcal{B}} = \begin{bmatrix} 12 & 0 \\ 0 & -4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 12 \\ -4 \end{bmatrix}.$$

Trace the path described in the diagam. A scales by 12 in the $\mathbf{v}_1$ direction and by -4 in the $\mathbf{v}_2$ direction.

To get from $[Q]_{\mathcal{B}} = (1, 1)$ to $[AQ]_{\mathcal{B}} = (12, -4)$, you start at the origin and move 12 $\mathbf{v}_1$-units up and to the right and 4 $\mathbf{v}_2$-units up and to the left (in the direction opposite to $\mathbf{v}_2$).

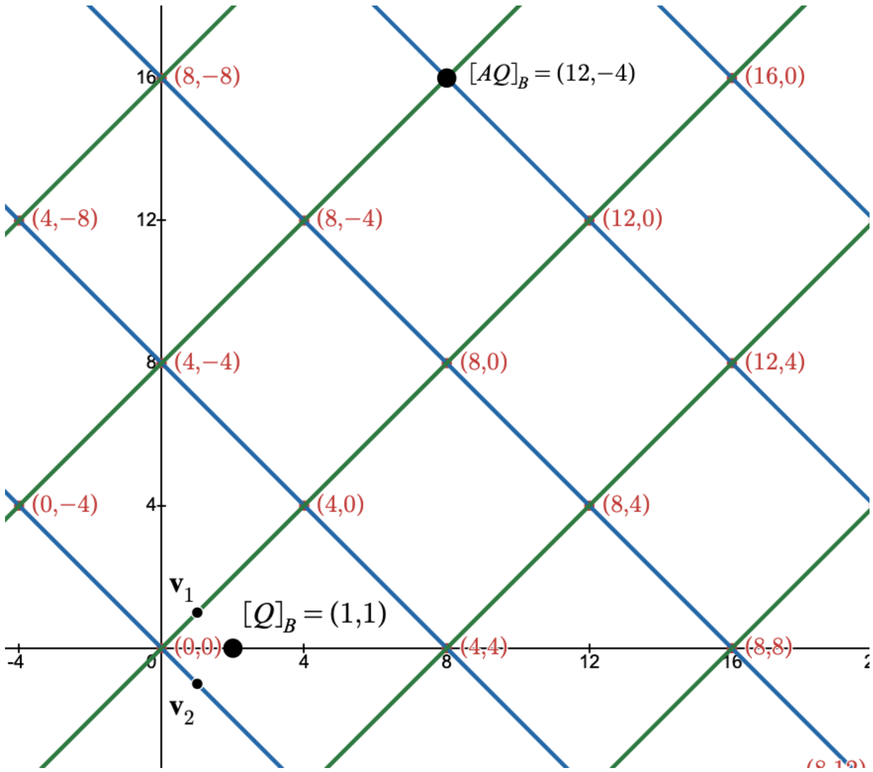


Figure 7.3: What A does to $\mathcal{B}$-coordinate vectors

Here is one reason we are better off if we work with respect to the basis $\mathcal{B}$. Observe that the powers of D are easy to compute:

$$D^k = \begin{bmatrix} 12 & 0 \\ 0 & -4 \end{bmatrix}^k = \begin{bmatrix} 12^k & 0 \\ 0 & (-4)^k \end{bmatrix}.$$

This fact lets us write down a formula for A^k. If we multiply both sides of the equation $P^{-1}AP = D$ by P on the left and P^{-1} on the right, we obtain $A = PDP^{-1}$. Now look what happens when we take powers of both sides: since $P^{-1}P = I$,

$$\begin{aligned} A^2 &= PDP^{-1}PDP^{-1} = PD^2P^{-1}; \\ A^3 &= PD^2P^{-1}PDP^{-1} = PD^3P^{-1}; \\ &\vdots \\ A^k &= PD^kP^{-1}. \end{aligned}$$

So for our particular example we have

$$A^k = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 12^k & 0 \\ 0 & (-4)^k \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$
$$= \frac{1}{2} \begin{bmatrix} 12^k + (-4)^k & 12^k - (-4)^k \\ 12^k - (-4)^k & 12^k + (-4)^k \end{bmatrix}.$$

Exercise 7C. Let

$$A = \begin{bmatrix} .3 & .7 \\ .7 & .3 \end{bmatrix}$$

and let

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Find scalars λ_i such that $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$. Let

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2.$$

Find a nice formula for $A^k\mathbf{x}_0$. What is the limit as k goes to infinity? Find the $\mathcal{B}$-matrix of A, where $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$.

Everything we have done so far generalizes to n dimensions with the concepts essentially unchanged.

Matrix of a linear transformation with respect to a basis

Definition 7.2. Let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis for $\mathbf{R}^n$ and let

$$P = P_{\mathcal{B}} = [\mathbf{b}_1 \ \cdots \ \mathbf{b}_n].$$

Let $T \colon \mathbf{R}^n \longrightarrow \mathbf{R}^n$ be a linear transformation with standard matrix representation A. Then the matrix $B = P^{-1}AP$ is called the **matrix for** T **with respect to** $\mathcal{B}$ or just the **$\mathcal{B}$-matrix of** T.

B maps the coordinate vector of $\mathbf{x}$ to the coordinate vector of $A\mathbf{x}$.

Theorem 7.3. *With notation as above, given any vector* $\mathbf{x}$ *in* $\mathbf{R}^n$ *we have*

$$B[\mathbf{x}]_{\mathcal{B}} = [A\mathbf{x}]_{\mathcal{B}}.$$

In addition, for all $k \geq 0$ we have

$$A^k = PB^kP^{-1}.$$

We close this section with a few comments.

- As the above example hopefully makes clear, working with respect to a useful basis makes analyzing dynamical systems much, much easier. In fact, we'll go further: *every* meaningful application of linear algebra incorporates, in some way, the choice of a useful basis. We will give several more examples as we proceed.

- We have said plenty in this chapter about how we might *use* a "good" basis; we have said *nothing* about how we might find one to begin with. We will remedy this (at least partly) over the coming chapters.
- Definition 7.2 and Theorem 7.3 are valid whether or not $\mathcal{B}$ is a particularly "good" basis. Recall that A and B are called *similar* if there is some invertible matrix P for which $A = PBP^{-1}$. In this situation, the columns of P constitute a basis $\mathcal{B}$ (whether useful or not). Therefore we see that "similar matrices represent the same linear transformation with respect to different bases."

Exercise 7D. Consider the following basis for $\mathbf{R}^2$

$$\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -4 \end{bmatrix} \right\},$$

and consider the following matrix:

$$A = \begin{bmatrix} 6 & 4 \\ -4 & -2 \end{bmatrix}.$$

Verify that

$$A\mathbf{v}_1 = 2\mathbf{v}_1 + \mathbf{v}_2 \text{ and } A\mathbf{v}_2 = 2\mathbf{v}_2.$$

Use these facts to find formulas for

$$A^k\mathbf{v}_1, \quad A^k\mathbf{v}_2, \quad \text{and} \quad A^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

(It happens to be the case that for this particular matrix A, there is *no way* to choose a basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ where $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ and $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$ for some pair of scalars λ_1, λ_2.)

Exercise 7E. This is a follow-up to Exercise 7D. Compute the $\mathcal{B}$-matrix M of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$. Show that M is the product of a scaling matrix with a shearing matrix.

Exercise 7F. Consider the matrix

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

and the following basis of $\mathbf{R}^3$:

$$\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \left\{ \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Verify that

$$A\mathbf{v}_1 = \mathbf{v}_2, \quad A\mathbf{v}_2 = \mathbf{v}_3, \quad \text{and} \quad A\mathbf{v}_3 = \mathbf{v}_1.$$

Use these facts to find (by hand)

$$A^{2025} \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}.$$

What is the $\mathcal{B}$-matrix of A?

Chapter 8 Vector spaces

Key concepts

- The definition of a vector space
- Examples: $\mathbf{R}^n$ and its subspaces, $\mathrm{M}_{n,k}(\mathbf{R})$, $\mathbf{C}$, P_n, $F(X,\mathbf{R})$
- Core linear algebra concepts in the vector space context (e.g., linearly independent sets, bases, etc.)
- Linear transformations between vector spaces
- The coordinate mapping for a vector space with a chosen basis
- Dimension
- The $\mathcal{B}$-matrix of a linear transformation

Summary. Vector spaces are sets that have addition and scalar multiplication operations which satisfy the 8 properties listed in Theorem 1.1. The most important example is $\mathbf{R}^n$ (and its subspaces). We already met the vector space of $n \times k$ matrices, but we discuss it further here. The toughest new examples involve functions, like the vector space P_n of polynomials of degree at most n, or the vector space of real valued functions $F(X,\mathbf{R})$.

Essentially all the linear algebra we did in $\mathbf{R}^n$ carries over to abstract vector spaces. The connection is made explicit by the coordinate mapping; every finite dimensional vector space is $\mathbf{R}^n$ in disguise. If V is a vector space with basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, then the coordinate mapping $V \to \mathbf{R}^n$ defined by $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ is an isomorphism. This mapping converts linear algebra questions about V into more familiar questions about $\mathbf{R}^n$. If $\mathcal{B}$ happens to be a basis for $\mathbf{R}^n$, then this is exactly the coordinate mapping from Chapter 7.

With the coordinate mapping in hand, we can represent any linear transformation $T\colon V \to V$ with a matrix (called the $\mathcal{B}$-matrix of T): it's the unique matrix M such that

$$M[\mathbf{x}]_{\mathcal{B}} = [T(\mathbf{x})]_{\mathcal{B}}.$$

For example, the $\mathcal{B}$-matrix for multiplication by i in the complex numbers (where $\mathcal{B}$ is the standard basis for $\mathbf{C}$) is the rotation matrix $\mathrm{R}_{\pi/2}$ (see §8.4.1).

Chapter 8

Vector spaces are sets that have addition and scalar multiplication operations which satisfy the 8 properties listed in Theorem 1.1. Any subspace V of $\mathbf{R}^n$ under the usual operations satisfies the 8 listed properties, so subspaces of $\mathbf{R}^n$ are all examples of vector spaces. Further, by Theorem 4.5, the set of $n \times k$ matrices $\mathrm{M}_{n,k}(\mathbf{R})$ also satisfies the 8 properties (where the zero matrix plays the role of the zero vector), so it too is an example of a vector space.

§8.1 Definition and examples

Let V be a set. A scalar multiplication operation on V is a function that takes any $s \in \mathbf{R}$ and $\mathbf{x} \in V$ and gives another element of V, denoted by $s\mathbf{x}$. An addition operation on V is a function that takes any $\mathbf{x}, \mathbf{y} \in V$ and gives another element of V, denoted by $\mathbf{x} + \mathbf{y}$. The next definition says what we mean when we say that V is a **vector space** under these operations. When V is a vector space, we will call its elements vectors. *Vector spaces are mathematical objects that have the same structural properties as* $\mathbf{R}^n$*, so we have essentially just copied the list of 8 statements from Theorem 1.1 to make this definition.*

VECTOR SPACES

Definition 8.1. Let V be a set with an addition operation and a scalar multiplication operation. The set V is a **vector space** if it satisfies the following properties for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and all $s, t \in \mathbf{R}$.

What's a vector space?

Compare this definition to Theorem 1.1 and Theorem 4.5.

① Vector addition is commutative:

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}.$$

② Vector addition is associative:

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z}).$$

③ There is an element $\mathbf{0} \in V$ that is the additive identity:

$$\mathbf{x} + \mathbf{0} = \mathbf{0} + \mathbf{x} = \mathbf{x}.$$

Part of defining a vector space is specifying what the zero vector is.

④ Every vector $\mathbf{x}$ has additive inverse $-\mathbf{x} = -1 \cdot \mathbf{x}$:

$$\mathbf{x} + (-\mathbf{x}) = -\mathbf{x} + \mathbf{x} = \mathbf{0}.$$

⑤ Scalar multiplication distributes over vector addition:

$$s(\mathbf{x} + \mathbf{y}) = s\mathbf{x} + s\mathbf{y}.$$

⑥ Scalar multiplication distributes over real number addition:

$$(s + t)\mathbf{x} = s\mathbf{x} + t\mathbf{x}.$$

⑦ The real number 1 acts as it should:

$$1\mathbf{x} = \mathbf{x}.$$

⑧ Scalar multiplication is associative:

$$s(t\mathbf{x}) = (st)\mathbf{x}.$$

To define a vector space V, we must

- define an addition operation,
- define a scalar multiplication operation,
- specify which element of V is the zero vector, and
- verify that the 8 properties hold.

We will always do the first three items above, but we will often leave all or part of the fourth step to the interested reader.

Example 8.2 (subspaces of $\mathbf{R}^n$)**.** It is straightforward to check that, under the usual operations, any subspace of $\mathbf{R}^n$ (inlucding $\mathbf{R}^n$ itself) is a vector space.

Subspaces of $\mathbf{R}^n$ are vector spaces.

$n \times k$ matrices form a vector space.

Example 8.3 ($n \times k$ matrices)**.** Theorem 4.5 says that the set of $n \times k$ matrices $M_{n,k}(\mathbf{R})$ is a vector space. The operations are matrix addition and matrix scalar multiplication, and the zero vector is the $n \times k$ zero matrix.

The complex numbers $\mathbf{C}$ *form a vector space.*

Example 8.4 (the complex numbers)**.** Some quadratic equations, like $x^2 - x + 1 = 0$, do not have real roots. If we use the quadratic formula (with $a = 1, b = -1, c = 1$), we get

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{1 \pm \sqrt{-3}}{2} = \frac{1 \pm \sqrt{3}\sqrt{-1}}{2} = \frac{1}{2} \pm \frac{\sqrt{3}}{2}\sqrt{-1}.$$

The hangup here is that pesky $\sqrt{-1}$, which can't be a real number. Let's denote this quantity by $i = \sqrt{-1}$. We will see in Chapter 10 that complex numbers are useful even if we only want to study "real" phenomena, such as dynamical systems in the plane. So let's enlarge the real numbers to a set that includes the quantity i, which you should think of as a new number that is *not* a real number and has the property that $i^2 = -1$. The **complex numbers** are the elements of the set

What's a complex number?

$$\mathbf{C} = \{a + bi \mid a, b \in \mathbf{R}\}.$$

The set of complex numbers is really just $\mathbf{R}^2$ in disguise! Points in the plane correspond to complex numbers via

$$(a, b) \leftrightarrow a + bi.$$

RQ

Reading Question 8A. Convert each of the following complex numbers into vectors using the correspondence above and plot them in the plane: $0 = 0 + 0i$, $i = 0 + 1i$, $-i$, $3i$, $-2 + 2i$.

The set $\mathbf{C}$ is a vector space. The addition and scalar multiplication operations are given by

$$s(x + yi) = (sx) + (sy)i$$
$$(x + yi) + (p + qi) = (x + p) + (y + q)i$$

and the zero vector in $\mathbf{C}$ is the complex number $0 = 0 + 0i$. You can see that, under the correspondence $(a, b) \leftrightarrow a + bi$, these operations correspond exactly with scalar multiplication and addition in $\mathbf{R}^2$, so it is no surprise that $\mathbf{C}$ satisfies the 8 properties in Definition 8.1.

Example 8.5 (real polynomials). Let

Polynomials of degree at most n form a vector space.

$$P_n = \{a_0 + a_1 t + \cdots + a_n t^n \mid a_0, \ldots, a_n \in \mathbf{R}\}$$

denote the set of all polynomials with real coefficients and degree at most n. When you scale a polynomial of degree at most n you get a polynomial of degree at most n, and when you add two polynomials of degree at most n you get a polynomial of degree at most n. The zero vector in P_n is the zero polynomial—the polynomial whose every coefficient is zero. The zero polynomial is the same as the constant function at 0; it is the function that sends every input t to zero.

For example, P_3 is the set of polynomials of degree at most 3:

$$P_3 = \{a + bt + ct^2 + dt^3 \mid a, b, c, d \in \mathbf{R}\}.$$

Any of a polynomial's coefficients might be zero, so P_3 contains P_2, P_1 and P_0. In P_3:

$$\underbrace{(1 + t + t^3)}_{\begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}} + \underbrace{(3t - t^2 + 9t^3)}_{\begin{bmatrix} 0 \\ 3 \\ -1 \\ 9 \end{bmatrix}} = \underbrace{1 + 4t - t^2 + 10t^3}_{\begin{bmatrix} 1 \\ 4 \\ -1 \\ 10 \end{bmatrix}}$$

and

$$5\,\underbrace{(3t - t^2 + 9t^3)}_{\begin{bmatrix} 0 \\ 3 \\ -1 \\ 9 \end{bmatrix}} = \underbrace{15t - 5t^2 + 45t^3}_{\begin{bmatrix} 0 \\ 15 \\ -5 \\ 45 \end{bmatrix}}$$

Above, you can see how polynomials in P_3 correspond to vectors in $\mathbf{R}^4$ in such a way that the operations on P_3 correspond to ordinary addition and scalar multiplication in $\mathbf{R}^4$. The zero polynomial corresponds to the zero vector $(0, 0, 0, 0)$. We will investigate this connection further soon with the coordinate mapping.

Example 8.6 (real-valued functions). Let X be any set (though you can just pretend $X = \mathbf{R}$ when you read this) and define $F(X, \mathbf{R})$ to be the set of all

Real-valued functions form a vector space.

functions with domain X and codomain $\mathbf{R}$:

$$F(X, \mathbf{R}) = \{f : X \to \mathbf{R}\}.$$

Define addition and scalar multiplication on $F(X, \mathbf{R})$ as follows. Given $f, g \in F(X, \mathbf{R})$, define $f + g \in F(X, \mathbf{R})$ by the rule

$$(f + g)(x) = f(x) + g(x).$$

Given $f \in F(X, \mathbf{R})$ and $s \in \mathbf{R}$, define $sf \in F(X, \mathbf{R})$ by

$$(sf)(x) = sf(x).$$

RQ

Reading Question 8B. Suppose $f, g \in F(\mathbf{R}, \mathbf{R})$ satisfy

$$f(-8) = 4$$
$$g(-8) = -3.$$

Evaluate $f + g$ at $x = -8$ and evaluate $7f$ at $x = -8$. Try to do so carefully, using the notation above.

RQ

Reading Question 8C. Suppose $f, g \in F(\mathbf{R}, \mathbf{R})$ satisfy

$$f(t) = -1 + 2t + 3t^2$$
$$g(t) = 5 - 3t + 9t^2.$$

Find rules for evaluating the functions $f + g$ and $7f$.

The zero vector in $F(X, \mathbf{R})$ is the constant function 0. More carefully, the zero vector is the function $Z : X \to \mathbf{R}$, where the rule for evaluating it is

$$Z(x) = 0$$

for all $x \in X$.

Under the operations defined above, $F(X, \mathbf{R})$ is a vector space; we will not check all the properties, but here is an example of how you'd do it. Suppose we want to verify that addition is associative. Take $f, g, h \in F(X, \mathbf{R})$. We want to prove that $(f + g) + h = f + (g + h)$. The functions on each side of this equation both have the same domain (X) and the same codomain ($\mathbf{R}$), so to prove that they are equal we must show that they have the same rule for evaluation:

$$\begin{aligned}((f+g)+h)(t) &= (f+g)(t)+h(t) && [\text{definition of } + \text{ for } F(X,\mathbf{R})]\\ &= (f(t)+g(t))+h(t) && [\text{definition of } + \text{ for } F(X,\mathbf{R})]\\ &= f(t)+(g(t)+h(t)) && [+ \text{ in } \mathbf{R} \text{ is associative}]\\ &= f(t)+(g+h)(t) && [\text{definition of } + \text{ for } F(X,\mathbf{R})]\\ &= (f+(g+h))(t) && [\text{definition of } + \text{ for } F(X,\mathbf{R})].\end{aligned}$$

You'll check commutativity below.

Reading Question 8D. Check that addition in $F(X, \mathbf{R})$ is commutative. Do it carefully, like we did in the previous example. Your proof should be shorter for this property, and you should indicate where you use a key property of + for $\mathbf{R}$. RQ

Notice that many (actually, most) of the vector spaces we have listed here have other operations *in addition* to those that make them valid vector spaces. For example, matrices (of the right sizes) can be multiplied together; complex numbers, polynomials, and real-valued functions can be multiplied together also. However, for general vector spaces there is no notion of one vector being multiplied by another. If the *only* thing you know about V is that it is a vector space, and if $\mathbf{v}$ and $\mathbf{w}$ are vectors in V, then "$\mathbf{vw}$" just doesn't make sense.

A vector space structure does not include a way to multiply two vectors together.

Key ideas revisited §8.2

If S is a set of vectors in a vector space V, then the definition of span given for vectors in $\mathbf{R}^n$ carries over perfectly. Similarly, subspaces of V are defined exactly as they're defined for $\mathbf{R}^n$, and spans are subspaces. So if you want more examples of vector spaces, you can take the span of a set of vectors in any vector space you already know about.

In fact, definitions for all of the following concepts carry over to vector spaces V without modification: linearly independent set, linearly dependent set, dependence relation, span, subspace, basis, and dimension. You should reread all these definitions (if you don't already know them by heart) to convince yourself that they make perfect sense in any vector space V.

Example 8.7 (polynomials, again)**.** By definition, P_n is a subset of $F(\mathbf{R}, \mathbf{R})$.

In fact,

$$\begin{aligned} P_n &= \{a_0 + a_1 t + \cdots + a_n t^n \mid a_0, \ldots, a_n \in \mathbf{R}\} \\ &= \operatorname{span}\{1, t, t^2, \ldots, t^n\}. \end{aligned}$$

So P_n is a subspace of $F(\mathbf{R}, \mathbf{R})$. We claim that $\mathcal{M}_n = \{1, t, t^2, \ldots, t^n\}$ is a basis for P_n. Since we already know that $\mathcal{M}_n$ is a spanning set, we need to show it is linearly independent. To do this, we need to show that the only solution to

$$a_0 + a_1 t + \cdots + a_n t^n = 0$$

is $a_0 = a_1 = \cdots = a_n = 0$. The key here is that the zero on the right hand side of the equation is the zero polynomial; in order for the above equation to be true, it has to be true *for all real numbers t*. The polynomial on the left hand side of the equation must be *the same polynomial* as the polynomial 0 on the right. That forces all the coefficients to be zero, and our set $\mathcal{M}_n$ is indeed a basis for P_n.

RQ **Reading Question 8E.** Show that $\{1 + t, 1 - t\}$ is a basis for P_1.

Example 8.8. Consider the functions $\sin t$ and $\cos t$ in $F(\mathbf{R}, \mathbf{R})$. We claim that the set $\{\sin t, \cos t\}$ is linearly independent in this vector space. To see why, take $a, b \in \mathbf{R}$ and suppose

$$a \cos t + b \sin t = 0.$$

This equation says that the function $a \cos t + b \sin t$ is the zero function; that is, it is the function from $\mathbf{R}$ to $\mathbf{R}$ that sends *every input* to zero. Plugging in $t = 0$ into the equation above, we obtain $a = 0$. Plugging $t = \pi/2$ into the equation we obtain $b = 0$. By definition, $\{\sin t, \cos t\}$ is a linearly independent set.

RQ **Reading Question 8F.** Prove that $\{e^t, e^{2t}\}$ is linearly independent in the vector space $F(\mathbf{R}, \mathbf{R})$. [Hint: use the same strategy that was used in Example 8.8 above.]

Example 8.9 (more vector spaces of functions)**.** Here are some subspaces of $F(\mathbf{R}, \mathbf{R})$ that you may already be familiar with. As you go down the list, each set contains the next set.

- The set of integrable functions.
- The set $C^0(\mathbf{R})$ of continuous functions.
- The set $C^1(\mathbf{R})$ of differentiable functions.
- The set $C^n(\mathbf{R})$ of n-times differentiable functions.
- The set $C^\infty(\mathbf{R})$ of infinitely differentiable functions.
- The set of power series with infinite radius of convergence.
- The set of series solutions to $y'' - 4y = 0$.

For more details, see the exercise below.

Exercise 8A. Looking back in a Calculus I textbook if necessary, review some of the very basic properties of the derivative and the integral. Do you see why the integrable functions form a subspace of $F(\mathbf{R}, \mathbf{R})$? Do you see why the differentiable functions form a subspace?

Exercise 8B. Determine whether each subset of P_2 is a subspace:

① $\{p \in \mathrm{P}_2 \mid p(3) = 0\}$
② $\{p \in \mathrm{P}_2 \mid p(3) = 2\}$
③ $\{p \in \mathrm{P}_2 \mid p' + 5p + 9 = 0\}$
④ $\{p \in \mathrm{P}_2 \mid p' + 5p = 0\}$

In each case where the subset is a subspace, find a basis.

Exercise 8C. A 2×2 matrix is half-magic if the sum of the numbers in each row and column is the same. Let V denote the set of half-magic squares in $\mathrm{M}_{2,2}(\mathbf{R})$. Show that V is a subspace by finding a linearly independent set in $\mathrm{M}_{2,2}(\mathbf{R})$ whose span is V (the set you found is a basis for V).

If V and W are vector spaces, then a linear transformation $T\colon V \to W$ is a function satisfying

$$T(s\mathbf{x} + t\mathbf{y}) = sT(\mathbf{x}) + tT(\mathbf{y})$$

for all $x, t \in \mathbf{R}$ and $\mathbf{x}, \mathbf{y} \in V$. This is exactly how we defined linear transformations for Euclidean spaces. The definitions of kernel and image also carry over directly.

Example 8.10. Define $T\colon \mathrm{P}_2 \to \mathrm{P}_4$ by

$$T(p(t)) = t^2 p(t).$$

This is multiplication by the fixed polynomial t^2, which takes at most degree 2 polynomials to at most degree 4 polynomials. For example,

$$T(1+3t-t^2) = t^2 + 3t^3 - t^4.$$

The function T is linear. Given $p(t), q(t) \in \mathrm{P}_2$ and $a, b \in \mathbf{R}$,

$$\begin{aligned} T(ap(t)+bq(t)) &= t^2(ap(t)+bq(t)) \\ &= a(t^2p(t)) + b(t^2q(t)) \\ &= aT(p(t)) + bT(q(t)). \end{aligned}$$

Let's compute the kernel. If $p(t) = a_0 + a_1t + a_2t^2$ is in the kernel, then that means $t^2p(t) = a_0t^2 + a_1t^3 + a_2t^4$ *is equal to the zero polynomial.* The only way for this polynomial to be the zero polynomial—the function that sends *every* input to zero—is if all the coefficients are zero: $a_0 = a_1 = a_2 = 0$. That means the kernel contains only the zero polynomial:

$$\ker T = \{0\}.$$

Note that this means that T is one-to-one, because a linear transformation is one-to-one if and only if the kernel contains only the zero vector.

How about the image of T? By the calculation above, we have

$$\begin{aligned} \operatorname{im} T &= \{a_0t^2 + a_1t^3 + a_2t^4 \mid a_0, a_1, a_2 \in \mathbf{R}\} \\ &= \operatorname{span}\{t^2, t^3, t^4\}. \end{aligned}$$

We see that T is not onto; the polynomial t, for example, is not in the image.

Exercise 8D. Show that the set of upper triangular matrices

$$\left\{ \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \mid a, b, c \in \mathbf{R} \right\}$$

is a subspace of $\mathrm{M}_2(\mathbf{R})$ by showing that it is the span of a set of three vectors (by which we mean three specific 2×2 upper triangular matrices). Show that your set of three vectors is linearly independent.

Exercise 8E. Define $T : \mathrm{P}_n \to \mathrm{P}_n$ by $T(p) = p'$ (that is, $T(p)$ is the derivative of p). Prove that T is linear. Compute the image and kernel of T.

Exercise 8F. Let F be a fixed $n \times n$ matrix. Define

$$T : \mathrm{M}_n(\mathbf{R}) \to \mathrm{M}_n(\mathbf{R})$$

by $T(X) = FX - XF$. Prove that T is linear. If X is in the kernel of T, can you say anything special about the relationship between F and X?

Exercise 8G. Suppose T is a linear transformation from V to W. Let $H = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ be a subspace of V. Write $T(H)$ as the span of a set of vectors.

Exercise 8H. The trace of a 2×2 matrix is given by

$$\operatorname{tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + d.$$

Verify that the set of matrices in $\mathrm{M}_2(\mathbf{R})$ with trace zero form a subspace. Find a basis for the subspace of 2×2 matrices with trace zero. Do the matrices with trace 1 form a subspace?

Exercise 8I. The map $T\colon \mathrm{P}_2 \to \mathbf{R}$ defined by $T(p) = p'(2) - p(1)$ is a linear transformation. Find a basis for $\ker T$.

The coordinate mapping ... again §8.3

In Chapter 7, we introduced the coordinate mapping for bases of $\mathbf{R}^n$, but we can apply the same ideas to bases of any vector space. Let V be a vector space and let

$$\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$$

be an ordered basis for V. As with $\mathbf{R}^n$, define the **$\mathcal{B}$-span mapping**

What's the span mapping?

$$\Phi : \mathbf{R}^n \to V$$

by

$$\Phi(x_1, \ldots, x_n) = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n.$$

This map takes weight vectors to linear combinations in the span of $\mathcal{B}$. The $\mathcal{B}$-coordinate mapping will be the inverse of the span mapping; we just need to know that the span mapping is a linear transformation that has an inverse (i.e., that Φ is an isomorphism). It's the coordinate mapping that we care about, but the span mapping is easier to compute and understand.

RQ **Reading Question 8G.** Write down a formula for the span mapping when $V = \mathbf{C}$ and $\mathcal{B} = \{1, i\}$. What are $\Phi(\mathbf{e}_1)$ and $\Phi(\mathbf{e}_2)$?

The span mapping is an isomorphism

Theorem 8.11. *Let V be a vector space with basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$. The map*

$$\Phi : \mathbf{R}^n \to V$$

defined by

$$\Phi(x_1, \ldots, x_n) = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$$

is an isomorphism.

Proof. We need to show that Φ is an invertible linear transformation. First, let's check that Φ is linear.

Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ be vectors in $\mathbf{R}^n$ and take $s, t \in \mathbf{R}$. Since

$$\begin{aligned}\Phi(s\mathbf{x} + t\mathbf{y}) &= \Phi(sx_1 + ty_1, \ldots, sx_n + ty_n) \\ &= (sx_1 + ty_1)\mathbf{b}_1 + \cdots + (sx_n + ty_n)\mathbf{b}_n \\ &= s(x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) + t(y_1\mathbf{b}_1 + \cdots + y_n\mathbf{b}_n) \\ &= s\Phi(\mathbf{x}) + t\Phi(\mathbf{y}),\end{aligned}$$

the map Φ is linear.

To prove that Φ is invertible, we will show that it is one-to-one and onto. To see that it is onto, simply observe that, by definition of Φ, its image is $\operatorname{span}\mathcal{B} = V$. To show that Φ is one-to-one, we need to show that its kernel is trivial ($\ker\Phi = \{\mathbf{0}\}$). Suppose $\Phi(\mathbf{x}) = \mathbf{0}$. Then,

$$x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n = \mathbf{0}.$$

Since $\mathcal{B}$ is linearly independent, $x_1 = \cdots = x_n = 0$. ■

Now we know that Φ has an inverse as a function, so we can define the $\mathcal{B}$**-coordinate mapping** to be the inverse of Φ:

$$[\mathbf{x}]_{\mathcal{B}} = \Phi^{-1}(\mathbf{x}).$$

So: *the entries of the vector $[\mathbf{x}]_{\mathcal{B}}$ are the $\mathcal{B}$-coordinates of $\mathbf{x}$:*

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

means

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n.$$

Reading Question 8H. Write down a formula for the coordinate mapping when $V = \mathbf{C}$ and $\mathcal{B} = \{1, i\}$. What are $[1]_{\mathcal{B}}$ and $[i]_{\mathcal{B}}$?

RQ

The coordinate mapping

Definition 8.12. Let V be a vector space and let $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be an ordered basis for V. The $\mathcal{B}$**-coordinate mapping**

What's the coordinate mapping?

$$[\ ]_{\mathcal{B}} : V \to \mathbf{R}^n$$

is defined by the equivalence

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \qquad \Longleftrightarrow \qquad \mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n.$$

If $V = \mathbf{R}^n$ and

$$P_{\mathcal{B}} = [\mathbf{b}_1 \cdots \mathbf{b}_n],$$

then $P_{\mathcal{B}}$ is the matrix representation of the span mapping and $P_{\mathcal{B}}^{-1}$ is the matrix representation of the coordinate mapping.

Reading Question 8I. In the context of the above discussion: what is $[\mathbf{b}_i]_{\mathcal{B}}$? What is $\Phi(\mathbf{e}_i)$?

RQ

You can use the coordinate mapping to translate essentially any linear algebra question about V into a question about $\mathbf{R}^n$. This is a really good thing, because $\mathbf{R}^n$ is the vector space that you understand best. For example, if you have a set of vectors in V and you want to know whether it's a basis, you can check whether their coordinate vectors form a basis for $\mathbf{R}^n$.

Use the coordinate mapping to trade questions about V for questions about $\mathbf{R}^n$.

Example 8.13 (standard coordinate mapping for P_n). Recall that $\mathcal{M}_n = \{1, t, t^2, \ldots, t^n\}$ is a basis for P_n. This means the coordinate mapping with respect to this basis is an isomorphism from P_n to $\mathbf{R}^{n+1}$. In particular, $\dim \mathrm{P}_n = n + 1$. If p is a polynomial, then to compute the $\mathcal{M}_n$-coordinate vector of p you need to find the weights needed to build p as a linear combination of the t^i. But these are just the coefficients of p! For example, in P_4,

$$[-2 + t^2 - 9t^3]_{\mathcal{M}_4} = (-2, 0, 1, -9, 0).$$

Example 8.14. Let $\mathcal{M}_2 = \{1, t, t^2\}$ denote the standard basis for P_2. Consider the set of polynomials $C = \{1 + t, 1 + t^2, t + t^2\}$. Their $\mathcal{M}_2$-coordinate vectors are:

$$\begin{aligned}\mathbf{u} = [1 + t]_{\mathcal{M}_2} &= (1, 1, 0)\\ \mathbf{v} = [1 + t^2]_{\mathcal{M}_2} &= (1, 0, 1)\\ \mathbf{w} = [t + t^2]_{\mathcal{M}_2} &= (0, 1, 1)\end{aligned}$$

The set C *is a basis for* P_2 *if and only if the set of coordinate vectors* $\{\mathbf{v}, \mathbf{u}, \mathbf{w}\}$ *is a basis for* $\mathbf{R}^3$. Via just two row operations,

$$A = [\mathbf{u}\ \mathbf{v}\ \mathbf{w}] = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix},$$

so $\{\mathbf{v}, \mathbf{u}, \mathbf{w}\}$ is a basis for $\mathbf{R}^3$ and C is a basis for P_2.

Exercise 8J. Find a basis for the subspace $\operatorname{span}\{1 + t, -2 - 2t, 1 + 2t\}$ of P_2.

RQ

Reading Question 8J. Following up on Example 8.14, let's find the coordinates of a particular polynomial with respect to the basis C.

- By definition, what critical equation is equivalent to

Look back at Definition 8.12.

$$[1 + 2t + 3t^2]_C = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

  and what exactly does this critical equation mean?
- Write down a linear system $A\mathbf{x} = \mathbf{b}$ whose unique solution is $[1 + 2t + 3t^2]_C$ and then solve it (you can use a computer).

Exercise 8K. Find a basis $\{\mathbf{u}, \mathbf{v}\}$ for the plane P in $\mathbf{R}^3$ with equation $y - z = 0$ such that the parallelograms in the grid determined by $\mathbf{u}$ and $\mathbf{v}$ are actually 1 by 2 unit rectangles. Find the coordinates of the point $(5, 3, 3)$ with respect to your basis. [Hint: to find the basis, note that the plane is $\ker \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$. If you use the PVF algorithm, the basis vectors you find will be perpendicular. One of the vectors will have length one, and you can scale the other so that its length is 2.]

We haven't forgotten that we owe you a proof. Theorem 6.14 says that the concept of the *dimension* of a vector space makes sense (please go back right now and review the statement). We are now in a position to prove this. In fact, the following proof shows slightly more: if V is any vector space (whether a subspace of $\mathbf{R}^n$ or not), and V has a basis with n elements, then *any* basis for V has that exact same number n of elements.

ALL BASES OF A GIVEN VECTOR SPACE HAVE THE SAME SIZE

Theorem 8.15. *Let V be a finite dimensional vector space. Any two bases for V must have the same number of vectors.*

Proof. What we need to prove is that if V is a vector space with a finite basis, then every basis for V has the same size. Suppose $\mathcal{B}$ is a basis for V with n elements and suppose $\mathcal{C}$ is a basis for V with m elements. Let

$$T_{\mathcal{B}}\colon V \longrightarrow \mathbf{R}^n$$

denote the $\mathcal{B}$-coordinate mapping and let

$$T_{\mathcal{C}}\colon V \longrightarrow \mathbf{R}^m$$

denote the $\mathcal{C}$-coordinate mapping. Both of these maps are invertible linear transformations. It is fairly straightforward to check that the inverse of a linear transformation is a linear transformation, and the composite of two linear transformations is a linear transformation (you checked the latter in Reading Question 4A). Hence, the composite

$$T\colon \mathbf{R}^n \xrightarrow{T_{\mathcal{B}}^{-1}} V \xrightarrow{T_{\mathcal{C}}} \mathbf{R}^m$$

is an isomorphism from $\mathbf{R}^n$ to $\mathbf{R}^m$. Since $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is basis for $\mathbf{R}^n$, its image set $\{T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)\}$ must be a basis for $\mathbf{R}^m$ because T is an isomorphism. But every basis for $\mathbf{R}^m$ has size m, so $n = m$ indeed! ■

Some vector spaces are *infinite-dimensional*. For these vector spaces, any basis must have infinitely many different vectors in it, and no finite subset of the vector space can span the vector space. The set of *all* polynomials, for example, is a an infinite-dimensional vector space.

Exercise 8L. Prove that the set P_∞ of all polynomials is an infinite-dimensional vector space: that is, prove that P_∞ a vector space and that no finite subset of P_∞ can span P_∞.

---

§8.4 Concrete examples in abstract vector spaces

§8.4.1 Multiplication by the complex number i

There is no doubt that $i = \sqrt{-1}$ makes for great dinner party conversation, but since it's just not a real number it can be hard to explain what it *means*. Though this number is "imaginary", there is a "real" way to think about it through the lens of geometry.

First, let's come up with a geometric interpretation for real numbers. We can view any positive real number t as S_t, the transformation of the plane that scales in both the x- and y-directions uniformly by t. What about -1? Well, scaling by -1 maps $\mathbf{e}_1$ to $-\mathbf{e}_1$ and $\mathbf{e}_2$ to $-\mathbf{e}_2$, so it has matrix representation

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \cos\pi & -\sin\pi \\ \sin\pi & \cos\pi \end{bmatrix} = \mathrm{R}_\pi$$

(rotation of the plane by π radians about the origin). So, a negative real number $t = -|t|$ corresponds to 180° rotation followed by $\mathrm{S}_{|t|}$.

Let's find a way to interpret i geometrically. Define $T\colon \mathbf{C} \to \mathbf{C}$ by $T(z) = iz$. So, T is multiplication by the complex number i. For example:

$$\begin{aligned} T(7) &= 7i \\ T(3i) &= 3i^2 = -3 \\ T(-2+5i) &= -5-2i \\ T(x+yi) &= -y+xi. \end{aligned}$$

The correspondence $(x, y) \leftrightarrow x + yi$ that we discussed when we introduced the complex numbers is nothing other than the coordinate mapping with respect to the basis $\mathcal{B} = \{1, i\}$ for $\mathbf{C}$. The span mapping takes (x, y) to $x + iy$ and the coordinate mapping takes $x + yi$ to (x, y). So, working "in coordinates", we can view our linear transformation as "the same" (given our choice of basis for $\mathbf{C}$) as the linear transformation $S\colon \mathbf{R}^2 \to \mathbf{R}^2$ defind by $S(x, y) = (-y, x)$. The matrix representation of S is called the $\mathcal{B}$-matrix of T.

RQ

Reading Question 8K. Find the $\mathcal{B}$-matrix of T by finding the matrix representation for S defined above. Then, show that the matrix you found is exactly $\mathrm{R}_{\pi/2}$ (rotation of the plane counter-clockwise $\pi/2$ radians about the origin).

Given the Reading Question above, we can say that the complex number i is 90° rotation. And, in the same way that doing 90° rotation twice is just

180° rotation, $i^2 = -1$ (and remember, -1 corresponds to 180° rotation). This connection should satisfy students and dinner party guests alike.

The $\mathcal{B}$-matrix of a linear transformation $V \to V$ §8.4.2

What we did in the last section is something we can do more generally, so let's lay it out here.

Let V be a vector space with basis $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, and let $T\colon V \to V$ be a linear transformation. When $V = \mathbf{R}^n$, we can repesent this function with an $n \times n$ matrix. It would be nice if we could do the same for any vector space, but on its face it just doesn't make sense to do so. You can multiply a matrix times a vector in $\mathbf{R}^n$, but if the vector lies in an odd-ball vector space, it's not clear what this would mean.

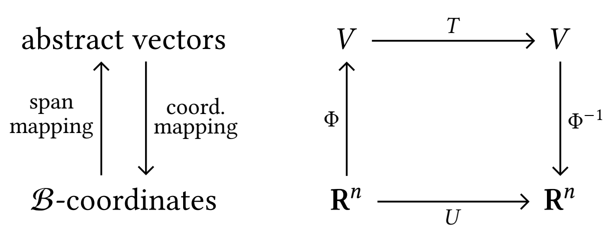
However, if we decide to work "in coordinates", then we can ask how T transforms $\mathcal{B}$-coordinate vectors. Since these lie in $\mathbf{R}^n$, we would then be able to find a matrix representation. We need a function U that that takes the *coordinate vector* of $\mathbf{v} \in V$ to the *coordinate vector* of $T(\mathbf{v})$. In symbols, we want:

$$U(\underbrace{[\mathbf{v}]_{\mathcal{B}}}_{\substack{\text{coord. vec.}\\ \text{of } \mathbf{v}}}) = \underbrace{[T(\mathbf{v})]_{\mathcal{B}}}_{\substack{\text{coord. vec.}\\ \text{of } T(\mathbf{v})}}.$$

Let $\mathbf{x} = [\mathbf{v}]_{\mathcal{B}}$; then, $\mathbf{v} = \Phi(\mathbf{x})$ and we have

$$U(\mathbf{x}) = [T(\Phi(\mathbf{x}))]_{\mathcal{B}}.$$

So, the function we care about is the bottom arrow in the following diagram:



Since $[\mathbf{b}_i]_{\mathcal{B}} = \mathbf{e}_i$, we have $\Phi(\mathbf{e}_i) = \mathbf{b}_i$. The ith column of the matrix representation of U is

$$\begin{aligned} U(\mathbf{e}_i) &= [T(\Phi(\mathbf{e}_i))]_{\mathcal{B}} \\ &= [T(\mathbf{b}_i)]_{\mathcal{B}}. \end{aligned}$$

We will call this matrix the $\mathcal{B}$-matrix of T.

What's the $\mathcal{B}$-matrix of a linear map $V \to V$?

The $\mathcal{B}$-matrix transforms $\mathcal{B}$-coordinate vectors according to the rule T.

MATRIX OF A LINEAR TRANSFORMATION WITH RESPECT TO A BASIS

Definition 8.16. Let V be a vector space with basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ and let $T\colon V \to V$ be a linear transformation. The **matrix of T with respect to the basis $\mathcal{B}$**, or just the **$\mathcal{B}$-matrix** of T, is the matrix

$$M = [\ \underbrace{[T(\mathbf{b}_1)]_{\mathcal{B}}}_{\substack{\text{first}\\ \text{column}}} \ \cdots \ \underbrace{[T(\mathbf{b}_n)]_{\mathcal{B}}}_{\substack{n\text{th}\\ \text{column}}}\ ].$$

M is the unique matrix that satisfies

$$M[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{B}}$$

for all $\mathbf{v} \in V$.

To find M, you apply T to each basis vector and put the results into coordinates.

In §7.2, we defined a basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ for $\mathbf{R}^2$ where

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix},$$

and we defined a matrix

$$A = \begin{bmatrix} 4 & 8 \\ 8 & 4 \end{bmatrix}.$$

We calculated that

$$A\mathbf{v}_1 = 12\mathbf{v}_1 = 12\mathbf{v}_1 + 0\mathbf{v}_2$$
$$A\mathbf{v}_2 = -4\mathbf{v}_2 = 0\mathbf{v}_1 - 4\mathbf{v}_2.$$

The $\mathcal{B}$-matrix of this linear transformation has a very special form. We will say more about this in Chapter 9.

RQ **Reading Question 8L.** Find the $\mathcal{B}$-matrix of $\mathbf{x} \mapsto A\mathbf{x}$ (where A is given just above).

Exercise 8M. Let $\mathbf{a}_1 = (3, 2)$ and $\mathbf{a}_2 = (1, 1)$. The set $\mathcal{B} = \{\mathbf{a}_1, \mathbf{a}_2\}$ is a basis for $\mathbf{R}^2$. Suppose $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ is a linear transformation with $T(\mathbf{a}_1) = \mathbf{a}_1 - 2\mathbf{a}_2$ and $T(\mathbf{a}_2) = \mathbf{a}_1 - \mathbf{a}_2$.

1. Find the $\mathcal{B}$-matrix of T.
2. Find the $\{\mathbf{e}_1, \mathbf{e}_2\}$-matrix of T (this is just the standard matrix representation). [Hint: if B is the matrix you found in the previous part, then the matrix you seek is PBP^{-1} for the right choice of P.]

Exercise 8N. Define $T\colon \mathrm{P}_2 \to \mathrm{P}_2$ by

$$T(p(t)) = p(2t-1).$$

So, for example,

$$T(2 - t + 3t^2) = 2 - (2t-1) + 3(2t-1)^2 = 6 - 14t + 12t^2.$$

Show that T is linear and find the $\mathcal{B}$-matrix of T, where $\mathcal{B} = \{1, t, t^2\}$ is the standard basis for P_2.

Exercise 8O. Let $a+bi$ be a fixed complex number and define $T\colon \mathbf{C} \to \mathbf{C}$ by $T(z) = (a+bi)z$. Find the $\mathcal{B}$-matrix of T, where $\mathcal{B} = \{1, i\}$ is the standard basis for $\mathbf{C}$.

Exercise 8P. Find the $\{1+i, 2-i\}$-matrix of T, where $T\colon \mathbf{C} \to \mathbf{C}$ is defined by $T(z) = iz$.

The derivative as a linear transformation §8.4.3

We can view the derivative as a linear transformation; you may have proved this in Exercise 8E. You know that the derivative of a sum is the sum of the derivatives, and when computing derivatives you know you can “pull out” scalars. In other words:

$$\frac{d}{dx}(sf(x) + tg(x)) = s\frac{d}{dx}f(x) + t\frac{d}{dx}g(x)$$

for all differentiable functions f, g and real numbers s, t. So, for example, we can define a linear transformation $D\colon \mathrm{P}_4 \to \mathrm{P}_4$ by $D(p(t)) = p'(t)$. For example,

$$D(1 + 2t + 3t^2 - 4t^3 - t^4) = 2 + 6t - 12t^2 - 4t^3.$$

Exercise 8Q. Compute the $\mathcal{B}$-matrix of D above, where $\mathcal{B} = \{1, t, t^2, t^3, t^4\}$ is the standard basis for P_4. To compute the $\mathcal{B}$-matrix for D, you must compute $D(t^k)$ for $k = 0, 1, 2, 3, 4$ and then put the results into coordinates. For example, to compute the third column:

$$[D(t^2)]_{\mathcal{B}} = [2t]_{\mathcal{B}} = (0, 2, 0, 0, 0).$$

Compute the remaining columns and find the matrix.

§8.4.4 *Fourier sine polynomials*

Trigonometric functions such as $\sin x$ and $\cos x$ can be used to model periodic phenomena like light waves or the volume of blood in the heart as a function of time. They are members of the vector space $F(\mathbf{R}, \mathbf{R})$. Let's focus on $\sin x$ for a moment. The sine function satisfies two key properties: it is an odd function (meaning $f(-x) = -f(x)$ for all x) and it is 2π-periodic (meaning $f(x + 2\pi) = f(x)$ for all x). Is there a way to use sine functions to model *any* function that is odd and periodic?

It turns out that if $f : \mathbf{R} \to \mathbf{R}$ is odd, 2π-periodic, and nice enough, then f can be approximated using sums of the form

$$a_1 \sin x + a_2 \sin 2x + \cdots + a_k \sin kx$$

Allowing infinite sums, this is called a Fourier sine series. For nice odd periodic functions f, the series coverges and is equal to f.

where $a_0, \ldots, a_k \in \mathbf{R}$. The expression above, called a Fourier sine polynomial, is a linear combination of certain functions in $F(\mathbf{R}, \mathbf{R})$!

For example, Figure 8.1 is the graph of a sawtooth wave function. It's built from copies of the graph of $y = x$ on $[-\pi, \pi]$. It is odd and 2π-periodic. It isn't continuous, but its discontinuities are not so horrible (they're "jump" discontinuities), and it can be approximated with Fourier sine polynomials. The more terms you use, the better the approximation.

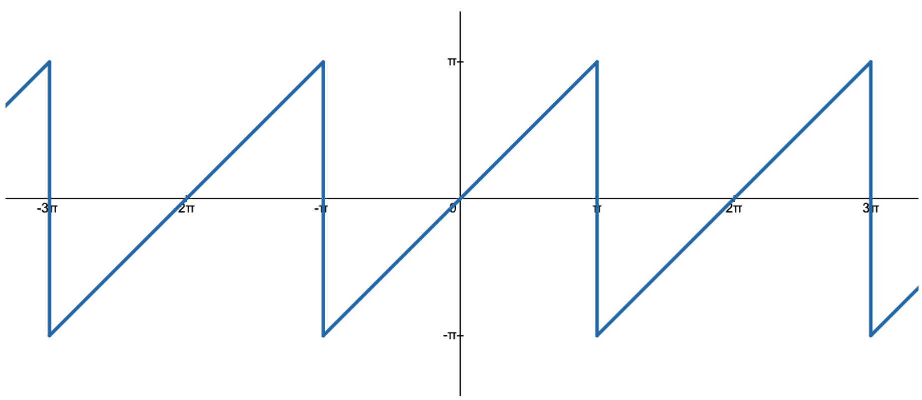


Figure 8.1: The sawtooth wave

Below is the sawtooth wave and its first Fourier sine polynomial, $2 \sin x$.

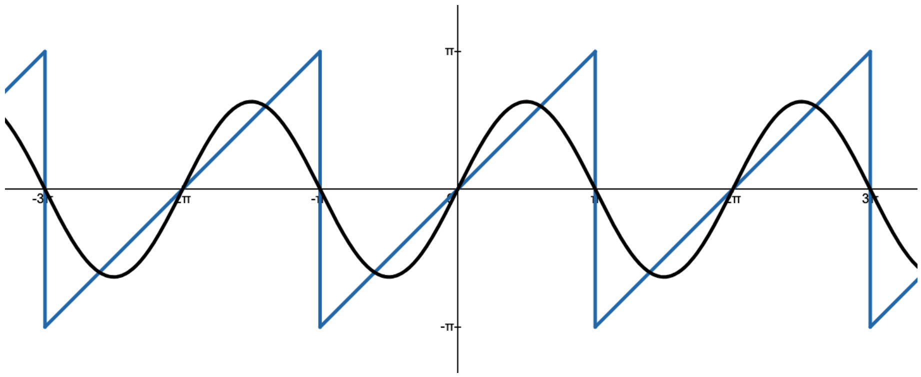


Figure 8.2: The sawtooth wave and its first Fourier sine polynomial

With its third Fourier sine polynomial, $2\sin x - \sin 2x + \frac{2}{3}\sin 3x$:

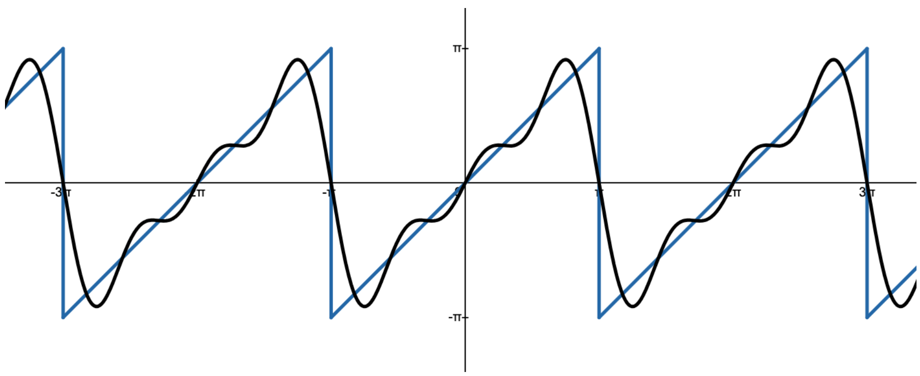


Figure 8.3: The sawtooth wave and its third Fourier sine polynomial

And, finally, with its 30th Fourier sine polynomial:

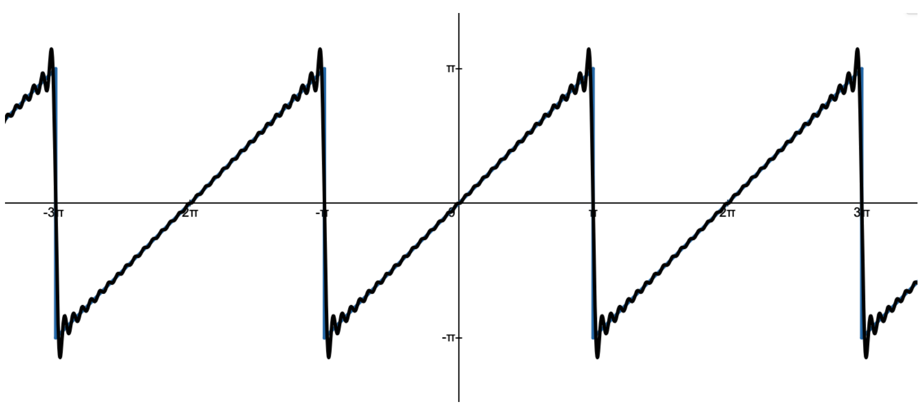


Figure 8.4: The sawtooth wave and its 30th Fourier sine polynomial

Let's define

$$V_k = \operatorname{span} \underbrace{\{\sin x, \sin 2x, \ldots, \sin kx\}}_{\mathcal{B}_k}.$$

One reason the sinusoidal functions $\sin kx$ are useful building blocks for periodic odd functions is that each set $\mathcal{B}_k$ is linearly independent and hence a basis for V_k. Let's prove this for $k = 3$.

RQ

Reading Question 8M. To show that $\{\sin x, \sin 2x, \sin 3x\}$ is linearly independent, we must show that the only solution to the equation

$$a_1 \sin x + a_2 \sin 2x + a_3 \sin 3x = 0$$

is $a_1 = a_2 = a_3 = 0$. Keep in mind that this is an equation for functions in $F(\mathbf{R}, \mathbf{R})$. So it says that the function on the left hand side *is equal to the zero function.* That means the equation must hold for all x. Get three equations involving the three unknowns a_1, a_2, a_3 by plugging in $x = \pi/4, \pi/2, 3\pi/4$. Solve these equations to prove that $a_1 = a_2 = a_3 = 0$.

For a function $f \in V_k$, writing f as a linear combination of the functions in $\mathcal{B}_k$ requires us to find the $\mathcal{B}_k$-coordinates of f. It turns out that the coefficient of $\sin mx$ is given by

$$a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin mx \, dx.$$

Exercise 8R. The sawtooth wave is not in V_3, but its 3rd Fourier sine polynomial *is* in V_3. It turns out that you can use the above integral formula to compute the weights. Verify that the 3rd Fourier sine polynomial for the sawtooth curve is

$$2 \sin x - \sin 2x + \frac{2}{3} \sin 3x$$

by computing the following integrals.

It's interesting to see that being able to integrate $x \sin mx$ has an application. Do it by parts!

$$\frac{1}{\pi} \int_{-\pi}^{\pi} x(\sin x) \, dx = 2$$
$$\frac{1}{\pi} \int_{-\pi}^{\pi} x(\sin 2x) \, dx = -1$$
$$\frac{1}{\pi} \int_{-\pi}^{\pi} x(\sin 3x) \, dx = 2/3.$$

If you're willing to work with infinite sums, then a nice enough odd 2π-periodic function f will be equal to the sum of the *series*

$$a_1 \sin x + a_2 \sin 2x + \cdots + a_k \sin kx + \cdots$$

(where the coefficients are defined using the above integral formula). Since these have infinitely many terms, this gives a concrete reason to study infinite dimensional vector spaces, though we will not do so here!

Exercise 8S. Define $D\colon V_3 \to V_3$ by $D(f) = f'$ (so D is just the derivative). Compute the $\mathcal{B}_3$-matrix of D.

Exercise 8T. Fix real nonzero constants k and r. Let $V = \operatorname{span}\{\cos kx, \sin kx, e^{rx}\}$. Show that the spanning set $\mathcal{B}$ used to define V is a basis for V. Let $D\colon V \to V$ be the derivative. Find the $\mathcal{B}$-matrix of D. When $r = 1$, this is the matrix for an affine linear transformation. What does it do, geometrically?

Chapter 9 Eigenvalues and eigenvectors

Key concepts

- For square matrices A, eigenvalues λ and eigenvectors $\mathbf{x}$ ($\mathbf{x} \neq \mathbf{0}$ with $A\mathbf{x} = \lambda\mathbf{x}$)
- Eigenspaces $E_\lambda = \{\mathbf{x} \mid A\mathbf{x} = \lambda\mathbf{x}\} = \ker(\lambda I_n - A)$ and how to compute them
- Finding eigenvalues of 2×2 matrices and triangular matrices
- Similar matrices have the same eigenvalues and isomorphic eigenspaces
- Eigenvectors with distinct eigenvalues are linearly independent
- The Eigenvector Basis Theorem
- Diagonalizability
- Using eigenvectors and diagonalizablity to study dynamical systems
- The determinant and its properties
- The characteristic polynomial and the role of algebraic multiplicity for eigenvalues

Summary. This is a crucial chapter that we have been building up to for some time. If A is a square matrix, λ is a number, and $\mathbf{x}$ is a nonzero vector with $A\mathbf{x} = \lambda\mathbf{x}$, then $\mathbf{x}$ is called an eigenvector for the eigenvalue λ. The set of all eigenvectors for λ, together with $\mathbf{0}$, is called the λ-eigenspace E_λ. The set E_λ is a subspace of $\mathbf{R}^n$.

The main point of this chapter is the Eigenvector Basis Theorem (Theorem 9.10). It's all about when, given an $n \times n$ matrix A, there is a basis for $\mathbf{R}^n$ consisting of eigenvectors for A. Such matrices are diagonalizable, and the matrix of $\mathbf{x} \mapsto A\mathbf{x}$ with respect to such a basis is diagonal. Hence, it is quite easy to compute solutions to the dynamical system $\mathbf{x} \mapsto A\mathbf{x}$ for any initial condition, and it's easy to compute the powers of A.

Associated to any $n \times n$ matrix A is a special degree n polynomial $p_A(\lambda)$ called the characteristic polynomial of A. Its roots are the eigenvalues of A. The determinant, though complicated to compute for large matrices, gives a clean definition of this polynomial: $p_A(\lambda) = \det(\lambda I_n - A)$. The determinant is also interesting because it captures invertibility ($\det A \neq 0$ if and only if A is invertible), it is well-behaved with respect to the matrix product ($\det(AB) = (\det A)(\det B)$), and it interacts nicely with row operations.

Chapter 9

In the main example of §7.2, we defined a matrix

$$A = \begin{bmatrix} 4 & 8 \\ 8 & 4 \end{bmatrix}$$

and showed that the vectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

satisfy

$$\begin{aligned} A\mathbf{v}_1 &= 12\mathbf{v}_1 = 12\mathbf{v}_1 + 0\mathbf{v}_2 \\ A\mathbf{v}_2 &= -4\mathbf{v}_2 = 0\mathbf{v}_1 - 4\mathbf{v}_2. \end{aligned}$$

We found that, if we're interested in studying the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$, $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ is a "good" basis for $\mathbf{R}^2$. Here are some reminders of what we mean by this.

Since $A\mathbf{v}_1 = 12\mathbf{v}_1$ and $A\mathbf{v}_2 = -4\mathbf{v}_2$, the $\mathcal{B}$-matrix of $\mathbf{x} \mapsto A\mathbf{x}$ is the diagonal matrix (see Reading Question 8L)

$$D = \begin{bmatrix} 12 & 0 \\ 0 & -4 \end{bmatrix}.$$

This means that $D[\mathbf{x}]_{\mathcal{B}} = [A\mathbf{x}]_{\mathcal{B}}$ for all $\mathbf{x} \in \mathbf{R}^2$ (in words: D sends the coordinate vector of $\mathbf{x}$ to the coordinate vector of $A\mathbf{x}$). Since

$$P = [\mathbf{v}_1 \ \mathbf{v}_2] = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

represents the span mapping and P^{-1} represents the coordinate mapping, we have

$$DP^{-1}\mathbf{x} = P^{-1}A\mathbf{x}$$

for all $\mathbf{x} \in \mathbf{R}^2$. From this it follows that $DP^{-1} = P^{-1}A$, or equivalently:

$$A = PDP^{-1}$$

— so A is similar to the diagonal matrix D. Finally, we can compute the powers of A easily via the formula $A^k = PD^kP^{-1}$.

Since A is similar to a diagonal matrix, it is called **diagonalizable**. The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are examples of **eigenvectors** with corresponding **eigenvalues** 12 and -4. In this chapter we will prove that an $n \times n$ matrix A is diagonalizable if and only if there is a basis for $\mathbf{R}^n$ consisting entirely of eigenvectors for A. Having such a basis enables us to easily compute solutions to the dynamical system $\mathbf{x} \mapsto A\mathbf{x}$ for any initial condition, and it allows us to easily compute the powers of A.

§9.1 Definitions and basic properties

What's an eigenvector? What's an eigenvalue?

The zero vector is not an eigenvector.

EIGENVALUES AND EIGENVECTORS

Definition 9.1. Let A be an $n \times n$ matrix. If $\mathbf{v}$ is a nonzero vector and λ is a number such that $A\mathbf{v} = \lambda\mathbf{v}$, then λ is called an **eigenvalue** for A and $\mathbf{v}$ is called an **eigenvector** for A (corresponding to the eigenvalue λ).

RQ

Reading Question 9A. Let

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

and let

$$\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Use the definition above to show that each of these vectors is an eigenvector (find the corresponding eigenvalue). For example, compute $A\mathbf{u}$ and you should be able to tell that the result is a multiple of $\mathbf{u}$. Use a computer to check that $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a linearly independent set. Conclude that, in this case, $\mathbf{R}^3$ has a basis consisting entirely of eigenvectors for A.

In general, given an $n \times n$ matrix A, $\mathbf{R}^n$ might or might not have a basis consisting entirely of eigenvectors for A. We repeat that (as we asserted above and will prove below), this happens if and only if A is diagonalizable.

Note that, in Definition 9.1, we do not allow zero vectors to be eigenvectors. The reason is that, if $\mathbf{v} = \mathbf{0}$, then the equation $A\mathbf{v} = \lambda\mathbf{v}$ is too trivial:

$$A\mathbf{0} = \lambda\mathbf{0}$$

holds for *every* λ! Please, though, do not get confused: $\lambda = 0$ is a perfectly good eigen*value*.

Reading Question 9B. For which matrices A is $\lambda = 0$ an eigenvalue? To answer this question, plug $\lambda = 0$ into the definition above and then review The Isomorphism Theorem. RQ

Reading Question 9C. Show that any number λ is an eigenvalue for some $n \times n$ matrix. [Hint: given λ, find a matrix A such $A\mathbf{v} = \lambda\mathbf{v}$ for *all* vectors $\mathbf{v}$!] RQ

Exercise 9A. Suppose that A is an $n \times n$ matrix.

(1) Show that, if λ is an eigenvalue for A, then $c\lambda^k$ is an eigenvalue for cA^k.
(2) Suppose that $A^3 = 0$. What are the possible eigenvalues for A?

Exercise 9B. Suppose T is the linear transformation from the plane to itself that first reflects across the line $y = 10x$ and then scales by a factor of 3. Find an eigenvector and corresponding eigenvalue for the matrix representation A of T without actually computing A. (Think geometrically.) Can you find a different eigenvalue than the one you just found (and an associated eigenvector)?

Given a matrix A, how can we find its eigenvalues and eigenvectors? Let's begin with the question of how to find eigenvectors (assuming the eigenvalues are known); we'll talk about how to find eigenvalues shortly. Suppose $\mathbf{x}$ is an eigenvector with eigenvalue λ. Then, by definition,

$$A\mathbf{x} = \lambda\mathbf{x}.$$

Let's do some algebra and try to isolate $\mathbf{x}$:

$$\mathbf{0} = \lambda\mathbf{x} - A\mathbf{x}.$$

We can't factor $\mathbf{x}$ out of the expression on the right because "$\lambda - A$" doesn't make sense. We can only subtract A from another $n \times n$ matrix. Well, $\mathbf{x} = I_n\mathbf{x}$, so

$$\begin{aligned} \mathbf{0} &= \lambda\mathbf{x} - A\mathbf{x} \\ \mathbf{0} &= \lambda I_n\mathbf{x} - A\mathbf{x} \\ \mathbf{0} &= (\lambda I_n - A)\mathbf{x}. \end{aligned}$$

This makes sense; $\lambda I_n - A$ is an $n \times n$ matrix.

This means that, excluding the zero vector, the eigenvectors for the eigenvalue λ are exactly the nonzero vectors in $\ker(\lambda I_n - A)$.

What's an eigenspace?

Eigenspaces

Definition 9.2. Suppose A is an $n \times n$ matrix with eigenvalue λ. The subspace

$$E_\lambda = \ker(\lambda I_n - A)$$

is called the **eigenspace** of A for the eigenvalue λ. The space E_λ contains exactly the eigenvectors of A for the eigenvalue λ, together with the zero vector. If we need to keep track of A, we will write E_λ^A to denote this eigenspace.

RQ

Reading Question 9D. The fact that the eigenvectors for λ (along with the zero vector) constitute a subspace of $\mathbf{R}^n$ yields the important observations that any nonzero sum of two eigenvectors is an eigenvector (for the same eigenvalue), and any nonzero scalar multiple of an eigenvector is an eigenvector (for the same eigenvalue). Verify these observations directly, using Definition 9.1.

Since you know how to find a basis for the kernel of a matrix, *given an eigenvalue λ, you can find all the eigenvectors associated with λ.* Note also that since an eigenspace must contain a nonzero vector by definition, its dimension must be at least 1. For example, it is a fact that the matrix

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

has eigenvalues 2 and -3. As explained above, E_{-3} is exactly the kernel of the matrix

$$-3I_2 - A = \begin{bmatrix} -6 & -2 \\ 3 & 1 \end{bmatrix}.$$

A one-step row reduction shows that this kernel is

$$E_{-3} = \operatorname{span}\left\{\begin{bmatrix} 1 \\ -3 \end{bmatrix}\right\}.$$

We can also find a basis for E_{-3} by using the definition of eigenvector more directly. Saying that $\mathbf{v} = (v_1, v_2)$ is an eigenvector for A with eigenvalue -3 *means* that $A\mathbf{v} = -3\mathbf{v}$:

$$\begin{bmatrix} 3 & 2 \\ -3 & -4 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -3v_1 \\ -3v_2 \end{bmatrix}.$$

Writing all this out explicitly yields

$$3v_1 + 2v_2 = -3v_1;$$
$$-3v_1 - 4v_2 = -3v_2.$$

Both of these equations yield the same information: namely, that $v_2 = -3v_1$. So again we see that the eigenvectors for -3 are precisely the (nonzero) scalar multiples of $(1, -3)$. (The "redundancy" in the two eigenvector equations above reflects the fact that the solution set is at least one-dimensional! If you are looking for eigenvectors, by whatever method, and you arrive at a conclusion that there is only one eigenvector, you know you've made a mistake — either in your algebra or in what the eigenvalues actually were to begin with.)

Reading Question 9E. Given that $\lambda = 2$ is an eigenvalue for RQ

$$A = \begin{bmatrix} 0 & -6 \\ 1 & 5 \end{bmatrix},$$

find a basis for E_2 (that is, find a basis for the kernel of $2I_2 - A$).

Exercise 9C. Given that -1 and 5 are eigenvalues for

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

find bases for the eigenspaces E_{-1} and E_5.

We now understand how to find eigenvectors if we know a matrix's eigenvalues; but how can we find the eigenvalues to begin with? Since λ is an eigenvalue exactly when there is a *nonzero* $\mathbf{x}$ such that $(\lambda I_n - A)\mathbf{x} = \mathbf{0}$, the Isomorphism Theorem tells us that λ is an eigenvalue if and only if the $n \times n$ matrix $\lambda I_n - A$ is singular.

In the case of 2×2 matrices, we have the determinant as a way to check singularity. If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then

$$\begin{aligned} \det(\lambda I_2 - A) &= \det\left(\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - \begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) \\ &= \det \begin{bmatrix} \lambda - a & -b \\ -c & \lambda - d \end{bmatrix} \\ &= (\lambda - a)(\lambda - d) - bc \\ &= \lambda^2 - (a + d)\lambda + ad - bc. \end{aligned}$$

Thus the eigenvalues of the 2×2 matrix A are the solutions to the quadratic equation

$$\lambda^2 - (a + d)\lambda + ad - bc = 0.$$

What's the characteristic polynomial of a 2×2 matrix?

The quadratic polynomial above is called the **characteristic polynomial of** A.

RQ

Reading Question 9F. Find the eigenvalues of the matrix A in Reading Question 9E. Find a basis for the eigenspace you didn't already find a basis for.

Any $n \times n$ matrix M has a determinant $\det M$; we will say a little more about determinants of matrices larger than 2×2 in §9.3.1. It turns out that, as for the 2×2 case, $\det M = 0$ if and only if M is singular; therefore the eigenvalues of an $n \times n$ matrix A are precisely the numbers λ that satisfy the characteristic equation $\det(\lambda I - A) = 0$. This criterion is useful in understanding some of the theory underlying eigenvalues, and can potentially make explicit calculations of eigenvalues practical in certain cases (especially if n is pretty small). The fact is, however, that for large n we usually rely on computers to find (or approximate) eigenvalues.

We can also, in principle, use row reduction to find eigenvalues. Here is an example.

Example 9.3. To find the eigenvalues of

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

we must find the values of λ where the matrix

$$\lambda I_3 - A = \begin{bmatrix} \lambda - 1 & -1 & 1 \\ -5 & \lambda - 2 & -4 \\ 0 & -1 & \lambda - 3 \end{bmatrix}$$

has fewer than three pivots. Let's row reduce. Start with

$$R_1 \mapsto 5R_1 + (\lambda - 1)R_2$$

to get:

$$\begin{bmatrix} 0 & \lambda^2 - 3\lambda - 3 & -4\lambda + 9 \\ -5 & \lambda - 2 & -4 \\ 0 & -1 & \lambda - 3 \end{bmatrix}.$$

Then, swap the top row to the bottom and the middle row to the top:

$$\begin{bmatrix} -5 & \lambda-2 & -4 \\ 0 & -1 & \lambda-3 \\ 0 & \lambda^2-3\lambda-3 & -4\lambda+9 \end{bmatrix}.$$

Finally, do $R_3 \mapsto R_3 + (\lambda^2 - 3\lambda - 3)R_2$:

Note that every row operation we did was valid, no matter what the value of λ is.

$$\begin{bmatrix} -5 & \lambda-2 & -4 \\ 0 & -1 & \lambda-3 \\ 0 & 0 & \lambda^3-6\lambda^2+2\lambda+18 \end{bmatrix}.$$

This matrix has will have three pivot positions unless

$$\lambda^3 - 6\lambda^2 + 2\lambda + 18 = 0.$$

The eigenvalues of A are the roots of this polynomial, which you could use a computer to find.

Let's look at a special situation where it's easy to spot the eigenvalues. An $n \times n$ matrix is called **upper triangular** if every entry below the main diagonal is zero. (It's perfectly fine if there are *more* zeros than this; having zeros below the main diagonal is the only requirement.) An $n \times n$ matrix is called **lower triangular** if every entry above the main diagonal is zero. A **triangular** matrix is a matrix that is either upper or lower triangular (note that a diagonal matrix is both). For such matrices, we can read off the eigenvalues.

What're upper triangular, lower triangular, and triangular matrices?

EIGENVALUES OF TRIANGULAR MATRICES

Theorem 9.4. *The eigenvalues of a triangular matrix are exactly the diagonal entries.*

Proof. Let's prove the theorem for upper triangular matrices; the proof for lower triangular matrices is almost identical. If A is upper triangular, then it has the form

$$\begin{bmatrix} a_1 & * & * & \cdots & * \\ 0 & a_2 & * & \cdots & * \\ 0 & 0 & a_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_n \end{bmatrix}.$$

So, $\lambda I_n - A$ has the form

$$\begin{bmatrix} \lambda - a_1 & * & * & \cdots & * \\ 0 & \lambda - a_2 & * & \cdots & * \\ 0 & 0 & \lambda - a_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda - a_n \end{bmatrix}.$$

This matrix will have exactly n pivots unless $\lambda = a_i$ for some i. Thus the a_i are exactly the eigenvalues of A. ■

RQ

Reading Question 9G. In each case, write down an example of a non-diagonal 3×3 matrix whose ...

- ... only eigenvalue is 0.
- ... only eigenvalues are 0 and 2.
- ... only eigenvalues are 1, 2, and 3.

RQ

Reading Question 9H. Find a 2×2 matrix that does not have any real eigenvalues. [Perhaps you would like to answer this question by writing down a triangular matrix with complex numbers on the main diagonal. Nice job, smarty-pants — you're totally correct. But now try to find a 2×2 matrix with *real* entries that does not have any real eigenvalues.]

The next theorem implies that if a matrix has k distinct eigenvalues, then it has a set of k linearly independent eigenvectors. In particular, if an $n \times n$ real matrix A has n distinct real eigenvalues, then there is a basis for $\mathbf{R}^n$ consisting entirely of eigenvectors of A.

When the eigenvalues are not distinct, the set may or may not be LI. It just depends.

EIGENVECTORS DRAWN FROM DIFFERENT EIGENSPACES ARE LI

Theorem 9.5. *Let A be an $n \times n$ matrix and suppose $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are eigenvectors with corresponding* distinct *eigenvalues $\lambda_1, \ldots, \lambda_k$. Then, the set*

$$\mathcal{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$$

is linearly independent. Thus,

$$\dim \operatorname{span} \mathcal{V} = k.$$

Proof. This proof is neat. Let's write

$$\begin{aligned}\mathcal{V}_1 &= \{\mathbf{v}_1\}\\ \mathcal{V}_2 &= \{\mathbf{v}_1, \mathbf{v}_2\}\\ &\vdots\\ \mathcal{V}_k &= \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \mathcal{V}.\end{aligned}$$

Since eigenvectors are nonzero, the set $\mathcal{V}_1$ is linearly independent because it contains only one vector. Imagine that $\mathcal{V}_k$ is NOT linearly independent. Then, there has to be *first* index i (it might be any index between 2 and k) such that $\mathcal{V}_i$ is linearly dependent and $\mathcal{V}_{i-1}$ is linearly independent.

Since $\mathcal{V}_i$ is linearly dependent, there is some nontrivial linear combination of the $\mathbf{v}_i$ that yields the zero vector: that is,

$$d_1\mathbf{v}_1 + \cdots + d_i\mathbf{v}_i = \mathbf{0}$$

where not all the d_j are zero. Since $\mathcal{V}_{i-1}$ is linearly independent, in the equation above we see that d_i in particular must be nonzero. Dividing through by d_i and moving $\mathbf{v}_i$ to the other side shows that $\mathbf{v}_i$ lies in the span of the vectors that came before it: so for some scalars c_j we have

You'll be asked to look back at this equation when we get close to the punchline.

$$\mathbf{v}_i = c_1\mathbf{v}_1 + \cdots + c_{i-1}\mathbf{v}_{i-1}.$$

Let's get eigenvalues involved in two different ways. On the one hand, we can multiply both sides of the equation above by λ_i to obtain

$$\lambda_i\mathbf{v}_i = \lambda_i c_1\mathbf{v}_1 + \cdots + \lambda_i c_{i-1}\mathbf{v}_i. \quad (\star)$$

On the other hand, we can left-multiply both sides of the equation by A and use the fact that the $\mathbf{v}_j$ are eigenvectors:

$$\begin{aligned}A\mathbf{v}_i &= A(c_1\mathbf{v}_1 + \cdots + c_{i-1}\mathbf{v}_i)\\ A\mathbf{v}_i &= c_1 A\mathbf{v}_1 + \cdots + c_{i-1}A\mathbf{v}_i\\ \lambda_i\mathbf{v}_i &= c_1\lambda_1\mathbf{v}_1 + \cdots + c_{i-1}\lambda_i\mathbf{v}_i. \quad (\dagger)\end{aligned}$$

Subtracting equation ($\dagger$) from equation ($\star$), we get

$$\mathbf{0} = c_1(\lambda_i - \lambda_1)\mathbf{v}_1 + \cdots c_{i-1}(\lambda_i - \lambda_{i-1})\mathbf{v}_{i-1}.$$

But the set $\mathcal{V}_{i-1}$ is linearly independent, so all the coefficients above must be zero. Since the eigenvalues are all distinct, the only way this can happen is if $c_1 = \cdots = c_{i-1} = 0$. Now go back through the proof and you'll see a line that now says $\mathbf{v}_i = \mathbf{0}$. This is impossible since $\mathbf{v}_i$ is an eigenvector, so it must be the case that $\mathcal{V}_k$ is linearly independent. ■

Reading Question 9I. Explain how you know that there is a basis for $\mathbf{R}^3$ consisting of

RQ

eigenvectors for

$$\begin{bmatrix} 1 & 0 & 0 \\ 4 & -3 & 0 \\ 3 & -9 & 12 \end{bmatrix}$$

without doing any computations.

Example 9.6. In Example 9.3, we looked at a 3×3 matrix and found that its eigenvalues are exactly the solutions to the equation

$$\lambda^3 - 6\lambda^2 + 2\lambda + 18 = 0.$$

Below is a plot of this polynomial:

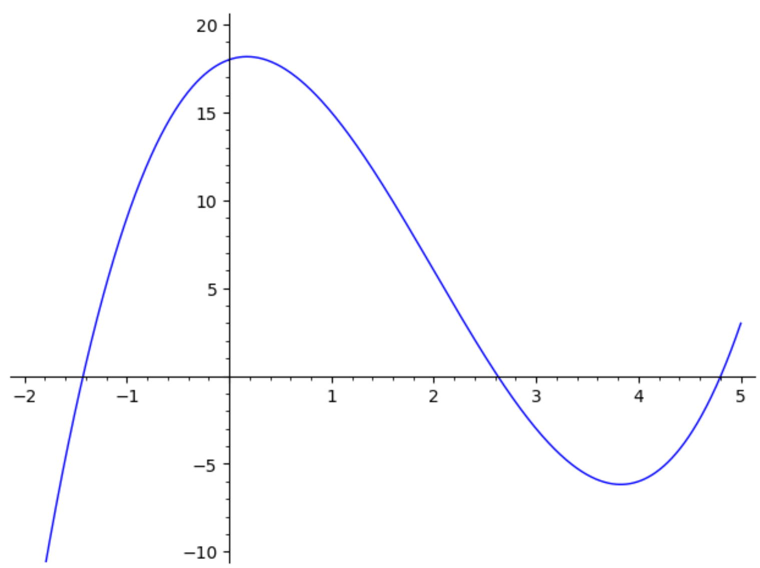


Figure 9.1: Plot of $\lambda^3 - 6\lambda^2 + 2\lambda + 18$

It's clear from the graph that the equation has three distinct solutions. Using a computer, we found their approximate values: $\lambda_1 = -1.4, \lambda_2 = 2.6$, and $\lambda_3 = 4.8$. We also found an approximate eigenvector associated to each eigenvalue:

$$\mathbf{v}_1 = \begin{bmatrix} 1.00 \\ -1.98 \\ 0.447 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1.00 \\ 0.444 \\ -1.18 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1.00 \\ 8.54 \\ 4.73 \end{bmatrix}.$$

The previous theorem tells us that these three vectors give a basis for $\mathbf{R}^3$ of eigenvectors for the matrix in Example 9.3.

If a matrix has fewer than n distinct real eigenvalues, then whether $\mathbf{R}^n$ has a basis consisting of eigenvectors for A hinges on whether the eigenspace dimensions for each real eigenvaluue are large enough. For example, if a 3×3 matrix A has only the eigenvalues 1 and 2, then there will be an eigenvector basis for $\mathbf{R}^3$ if and only if one of the eigenspaces (E_1 or E_2) has dimension 2. The reason is that Theorem 9.5 can be generalized slightly: suppose, for each i, we have a set $\mathcal{B}_i$ of linearly independent vectors in E_{λ_i} (where, as above, the λ_i are distinct eigenvalues). Then, the set

$$\mathcal{V} = \mathcal{B}_1 \cup \cdots \cup \mathcal{B}_k$$

is linearly independent. Thus, the only way to find a basis for $\mathbf{R}^n$ consisting of eigenvectors for A is if the eigenspace dimensions for each real eigenvalue sum to n. Otherwise, you just can't find enough eigenvectors!

Eigenvector Bases and Eigenspace Dimensions

Theorem 9.7. *Let A be a real $n \times n$ matrix and suppose $\lambda_1, \ldots, \lambda_k$ is a complete list of the distinct real eigenvalues of A. Then, there is a basis for $\mathbf{R}^n$ consisting of eigenvectors for A if and only if*

$$\dim E_{\lambda_1} + \cdots + \dim E_{\lambda_k} = n.$$

Example 9.8. Here is a matrix that doesn't yield an eigenvector basis for $\mathbf{R}^3$:

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

The only eigenvalues of this upper triangular matrix are $\lambda = 2$ and $\lambda = 3$. We have

$$\begin{aligned} E_2 &= \ker(2I_3 - A) \\ &= \ker \begin{bmatrix} 0 & -1 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \end{bmatrix} \\ &= \ker \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

So, $\dim E_2 = 1$. Similarly:

$$E_3 = \ker(3I_3 - A)$$

$$= \ker \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

So, $\dim E_3 = 1$. Since there are no other eigenvalues, we simply cannot find three linearly independent eigenvectors for A (two is the best we can do, plucking one eigenvector from each eigenspace).

Later, we will give an example of a 3×3 matrix with exactly two real eigenvalues that *does* yield an eigenvector basis for $\mathbf{R}^3$ (see Example 9.11).

We conclude this section with a theorem that says similar matrices have the same eigenvalues and isomorphic eigenspaces. There are important iterative computational methods for finding eigenvalues that use Theorem 9.9 item (1), but we will not discuss them in this book.

Similar matrices and eigendata

Similar matrices have the same eigenvalues and isomorphic eigenspaces.

Theorem 9.9. *Let A and B be similar $n \times n$ matrices.*

(1) *The number λ is an eigenvalue of A if and only if λ is an eigenvalue of B.*

(2) *If λ is an eigenvalue for A and B, then E_λ^A is isomorphic to E_λ^B. In particular,*

$$\dim E_\lambda^A = \dim E_\lambda^B.$$

Proof. Suppose A and B are simliar. Then, there is an invertible $n \times n$ matrix P such that $A = PBP^{-1}$. You can rearrange this equation in two useful ways:

$$P^{-1}A = BP^{-1}, \text{ and}$$
$$AP = PB.$$

First, let's show that if $\mathbf{v} \in E_\lambda^A$, then $P^{-1}\mathbf{v} \in E_\lambda^B$. Take $\mathbf{v} \in E_\lambda^A$. Then, $A\mathbf{v} = \lambda\mathbf{v}$ and

$$\begin{aligned} B(P^{-1}\mathbf{v}) &= (BP^{-1})\mathbf{v} \\ &= (P^{-1}A)\mathbf{v} \\ &= P^{-1}(A\mathbf{v}) \\ &= P^{-1}(\lambda\mathbf{v}) \\ &= \lambda(P^{-1}\mathbf{v}). \end{aligned}$$

This proves that $P^{-1}\mathbf{v} \in E_\lambda^B$.

We now have a linear transformation $T\colon E_\lambda^A \longrightarrow E_\lambda^B$ defined by $T(\mathbf{x}) = P^{-1}\mathbf{x}$. This function is one-to-one since P^{-1} is invertible, so it takes nonzero vectors to nonzero vectors. Hence, it takes eigenvectors to eigenvectors and if λ is an eigenvalue for A, then it is also an eigenvalue for B.

Conversely, an almost identical computation shows that if $\mathbf{w} \in E_\lambda^B$, then $P\mathbf{w} \in E_\lambda^A$. So, T has inverse $S\colon E_\lambda^B \longrightarrow E_\lambda^A$ defined by $S(\mathbf{x}) = P\mathbf{x}$. Again, S is one-to-one since P is invertible, so it takes eigenvectors to eigenvectors. So if λ is an eigenvalue for B, then it is also an eigenvalue for A.

Since we have established that T is an invertible linear transformation, the eigenspaces E_λ^A and E_λ^B are isomorphic and therefore have the same dimension. ∎

The Eigenvector Basis Theorem §9.2

What's a diagonalizable matrix?

An $n \times n$ matrix A is **diagonalizable** if A is similar to a diagonal matrix. Recall that this means $A = PDP^{-1}$ for some invertible matrix P and diagonal matrix D. If we want to refer to the specific matrices P and D that accomplish the diagonalization, then we will say that P **and** D **diagonalize** A. Recall also that when $A = PDP^{-1}$, we have $A^k = PD^kP^{-1}$ for all $k \geq 0$ (see Theorem 7.3).

For any sequence of numbers $d_1, \ldots, d_n$, let

$$\operatorname{diag}(d_1, \ldots, d_n) = [d_1\mathbf{e}_1 \ \cdots \ d_n\mathbf{e}_n] = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & d_n \end{bmatrix}$$

denote the diagonal matrix whose ith diagonal entry is d_i.

This theorem is actually true if you do linear algebra over $\mathbf{C}$ *instead of* $\mathbf{R}$*; you can replace* $\mathbf{R}$ *with* $\mathbf{C}$ *everywhere.*

$\mathcal{B}$ is a basis for $\mathbf{R}^n$ since P is invertible.

The Eigenvector Basis Theorem

Theorem 9.10. *Take $A, P \in \mathrm{M}_n(\mathbf{R})$ with P invertible and take $\lambda_1, \ldots, \lambda_n \in \mathbf{R}$. Write*

- $P = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n]$,
- $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, *and*
- $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

The following statements are equivalent.

① *$\mathcal{B}$ is a basis for $\mathbf{R}^n$ consisting of eigenvectors for A, with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$.*

② *P and D diagonalize A.*

③ *For all $k \geq 0$, the $\mathcal{B}$-matrix of A^k is D^k.*

④ *For all $k \geq 0$ and all scalars $c_1, \ldots, c_n$,*

$$A^k(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_n\lambda_n^k\mathbf{v}_n.$$

Proof. First, let's prove ① implies ②. Compute:

$$\begin{aligned} AP &= A[\mathbf{v}_1 \ \cdots \ \mathbf{v}_n] \\ &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_n] \\ &= [\lambda_1\mathbf{v}_1 \ \cdots \ \lambda_n\mathbf{v}_n] \\ &= [\lambda_1 P\mathbf{e}_1 \ \cdots \ \lambda_n P\mathbf{e}_n] \\ &= [P\lambda_1\mathbf{e}_1 \ \cdots \ P\lambda_n\mathbf{e}_n] \\ &= P[\lambda_1\mathbf{e}_1 \ \cdots \ \lambda_n\mathbf{e}_n] \\ &= PD. \end{aligned}$$

Since P is invertible, $A = PDP^{-1}$.

Next, we prove ② implies ③. Since $A = PDP^{-1}$, we have $A^k = PD^kP^{-1}$ for any k. Rearranging, we get $P^{-1}A^k = D^kP^{-1}$. Since P^{-1} is the $\mathcal{B}$-coordinate mapping, for all $\mathbf{x} \in \mathbf{R}^n$,

$$[A^k\mathbf{x}]_{\mathcal{B}} = D^k[\mathbf{x}]_{\mathcal{B}}.$$

Thus, the $\mathcal{B}$-matrix of A^k is D^k; see Definition 8.16.

Now we'll prove ③ implies ④. As we saw above, if the $\mathcal{B}$-matrix of A^k is

D^k, then for all $\mathbf{x} \in \mathbf{R}^n$,

$$\begin{aligned} [A^k\mathbf{x}]_{\mathcal{B}} &= D^k[\mathbf{x}]_{\mathcal{B}} \\ \implies P^{-1}A^k\mathbf{x} &= D^kP^{-1}\mathbf{x} \\ \implies A^k\mathbf{x} &= PD^kP^{-1}\mathbf{x}. \end{aligned}$$

Let $\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$. We have:

$$\begin{aligned} A^k(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) &= A^k\mathbf{x} \\ &= PD^kP^{-1}\mathbf{x} \\ &= PD^k[x]_{\mathcal{B}} \\ &= PD^k\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \\ &= P\begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \\ &= P\begin{bmatrix} c_1\lambda_1^k \\ \vdots \\ c_n\lambda_n^k \end{bmatrix} \\ &= c_1\lambda_1^k\mathbf{v}_1 + \cdots + c_n\lambda_n^k\mathbf{v}_n. \end{aligned}$$

The last equality follows from the fact that P represents the span mapping.

Finally, to prove ④ implies ①, simply take $k = 1$ and $\mathbf{c} = \mathbf{e}_i$ to obtain $A\mathbf{v}_i = \lambda_i$ for each i. ■

Note that point ④ of Theorem 9.10 shows us how to forecast the long-term behavior of dynamical systems in the diagonalizable case. In particular, since $\lambda_i^k \to 0$ if $|\lambda_i| < 1$, solutions will tend toward linear combinations of vectors in eigenspaces cooresponding to eigenvalues with absolute value greater than or equal to 1. We will see several examples of this observation in action below.

Exercise 9D. Construct a 2×2 matrix that is invertible but not diagonalizable. Construct a two by two matrix that is diagonalizable but not invertible.

Example 9.11. Let's study the matrix

$$A = \begin{bmatrix} 0.6 & -0.4 & -0.6 \\ 0.8 & 1.8 & 1.2 \\ -0.4 & -0.4 & 0.4 \end{bmatrix}.$$

Using a computer, we find that

$$\begin{aligned} xI_3 - A &= \begin{bmatrix} x - 0.60 & 0.40 & 0.60 \\ -0.80 & x - 1.8 & -1.2 \\ 0.40 & 0.40 & x - 0.40 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & -\frac{5}{4}x + \frac{9}{4} & \frac{3}{2} \\ 0 & x - 1 & 2x - 2 \\ 0 & 0 & x^2 - \frac{9}{5}x + \frac{4}{5} \end{bmatrix}. \end{aligned}$$

This matrix will be singular if $x - 1 = 0$ or $x^2 - (9/5)x + (4/5) = 0$. Using the quadratic formula or by just factoring, the roots of the quadratic are $x = 1$ and $x = 4/5$. So, there are exactly two eigenvalues.

Verifying these eigenspace computatations by finding the kernels in PVF would be good practice.

For $x = 1$, we have

$$E_1 = \ker(I_3 - A) = \ker \begin{bmatrix} 1 & 1 & 3/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \operatorname{span}\left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3/2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

For $x = 4/5$, we have

$$\begin{aligned} E_{4/5} &= \ker((4/5)I_3 - A) \\ &= \ker \begin{bmatrix} 1 & 5/4 & 3/2 \\ 0 & -1/5 & -2/5 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \ker \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \operatorname{span}\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}. \end{aligned}$$

We have found a basis for $\mathbf{R}^3$ of eigenvectors of A:

$$\mathcal{B} = \left\{ \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -3/2 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \right\}.$$

The Eigenvector Basis Theorem tells us:

- The matrices

$$P = \begin{bmatrix} -1 & -3/2 & 1 \\ 1 & 0 & 2 \\ 0 & 1 & -2 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4/5 \end{bmatrix}$$

diagonalize A.

You might directly check that $A = PDP^{-1}$, but the theorem tells you this equation really does hold.

- The $\mathcal{B}$-matrix of A^k is

$$D^k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (4/5)^k \end{bmatrix}.$$

Using a $\mathcal{B}$-coordinate system in $\mathbf{R}^3$, D^k is exactly what A^k does to $\mathcal{B}$-coordinate vectors.

- Any $\mathbf{x} \in \mathbf{R}^3$ has the form $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$ and

$$A^k(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3) = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3(4/5)^k\mathbf{v}_3.$$

To understand the long-run behavior of the dynamical system $\mathbf{x} \mapsto A\mathbf{x}$, note that $(4/5)^k \to 0$ as $k \to \infty$, so the last point above tells us that $A^k\mathbf{x} \to c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. The vectors in the eigenspace E_1 are fixed by the powers of A.

Exercise 9E. Do Exercise 9C if you haven't done so. What are all the things the Eigenvector Basis Theorem tells you in this situation?

Example 9.12. In Example 9.8 we gave an example of a matrix that does not yield an eigenvector basis for $\mathbf{R}^3$. By the Eigenvector Basis Theorem, the matrix is not diagonalizable.

Example 9.13. Here is another matrix that is not diagonalizable:

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

This columns of this matrix are clearly independent, and so 0 is not an eigenvalue. Assuming, then, that $\lambda \neq 0$ and row-reducing, we get

$$\lambda I_3 - A = \begin{bmatrix} \lambda + 1 & 0 & 0 \\ 0 & \lambda & 2 \\ 0 & -2 & \lambda \end{bmatrix} \sim \begin{bmatrix} \lambda + 1 & 0 & 0 \\ 0 & \lambda & 2 \\ 0 & 0 & \lambda^2 + 4 \end{bmatrix}.$$

We therefore see that $\lambda I_3 - A$ is singular if $\lambda + 1 = 0$ or if $\lambda^2 + 4 = 0$. The first condition tells us that -1 is an eigenvalue, and you can check that $E_1 = \text{span}\{\mathbf{e}_1\}$.

$\lambda^2 + 4$, however, is only equal to 0 if λ is equal to one of the two *complex* numbers $\pm 2i$. There is certainly no vector $\mathbf{v}$ in $\mathbf{R}^3$ for which $A\mathbf{v} = \pm 2i\mathbf{v}$ (since all the entries of $A\mathbf{v}$ are real!) Thus $\mathbf{R}^3$ has no basis consisting of eigenvectors for A, and A is not diagonalizable in the sense of this section.

Diagonalizability fails in very different ways in the last two examples. In Example 9.12 (which is really Example 9.8), all the eigenvalues were real but it was impossible to come up with three linearly independent eigenvectors. In Example 9.13, the problem stemmed instead from the fact that we had *complex* numbers that were apparently eigenvalues, for which we couldn't hope to find real eigenvectors.

It turns out that, for the matrix A in Example 9.13, the numbers $\pm 2i$ really are eigenvalues in the sense of Definition 9.1 *if we consider ourselves as working in* $\mathbf{C}^n$ *rather than* $\mathbf{R}^n$. Moreover, there is in an invertible matrix P (with complex entries) such that $A = P\text{diag}(-1, 2i, -2i)P^{-1}$, and so A is "diagonalizable over the complex numbers." Put another way, Theorem 9.10 is true if you replace $\mathbf{R}$ with $\mathbf{C}$.

In this book we have worked in the real setting — that is, the entries in our vectors and matrices are real numbers. This setting is far more familiar to most of us. It may seem unfair to spring this on you now, but nearly **everything** we have done so far (row-reduction, solution sets, spanning, independence, properties of linear transformations, invertibility, subspaces, dimension, similarity, and diagonalizability) works in the same way, with the same big theorems, if we allow our vectors and matrices to have complex entries. (An exception is the geometric material we have studied in the plane, where the fact that we were working in $\mathbf{R}^2$ definitely made a difference.) In addition, there are some aspects of the complex setting that make things easier, and the theorems stronger. For example, in the complex setting it turns out that

- *every* matrix has at least one eigenvalue, with a corresponding complex eigenspace of dimension at least one (this is a consequence of the so-called *fundamental theorem of algebra*);
- the collection of all eigenvalues yields excellent long-run information about how fast the entries of A^k grow or shrink;
- *every* matrix is similar to an upper-triangular matrix.

All this said, with the exception of a brief (and delightful) foray into some complex-valued computations in Chapter 10, we will actually continue to work in the real setting for the rest of this this book. However, you should be aware that **for many mathematicians linear algebra is presumptively carried out in the complex setting**. So if you are reading other sources and someone talks about a "diagonalizable" matrix A, they may mean that there are **complex** matrices P and D for which $A = PDP^{-1}$ — and this is *not* what we mean in this course.

Exercise 9F. Let

$$A = \begin{bmatrix} 3 & 5 & 1 \\ 0 & 5 & c \\ 0 & 0 & 3 \end{bmatrix}$$

For which values of c is A diagonalizable?

Exercise 9G. Let

$$A = \begin{bmatrix} r & s & t \\ 0 & a & b \\ 0 & c & d \end{bmatrix}.$$

Prove that the eigenvalues of A are $\lambda = r$ and the roots of the characteristic polynomial for the 2×2 matrix in the bottom right corner of A.

Exercise 9H (from Lay et. al.'s linear algebra book)**.** Let

$$A = \begin{bmatrix} .5 & .2 & .3 \\ .3 & .8 & .3 \\ .2 & 0 & .4 \end{bmatrix}.$$

This matrix is an example of a **stochastic matrix**: its column sums are all equal to 1. The vectors

$$\mathbf{v}_1 = \begin{bmatrix} .3 \\ .6 \\ .1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

are all eigenvectors of A. Find their associated eigenvalues. Now, let $\mathbf{x}_0$ be a vector with non-negative entries that sum to 1 (such a vector is called a probability vector). Explain why there are constants c_1, c_2, c_3 such that

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3.$$

Compute $\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \cdot \mathbf{x}_0$ and deduce that $c_1 = 1$. Finally, let $\mathbf{x}_k = A^k\mathbf{x}_0$. Show that $\mathbf{x}_k \to \mathbf{v}_1$ as k goes to infinity. (The vector $\mathbf{v}_1$ is called a **steady-state vector** for A.)

---

§9.3 The characteristic polynomial

Though it is possible to do otherwise, we think it's best to use the determinant to define the characteristic polynomial of a matrix A. The most important thing about this polynomial is that its roots are the eigenvalues of A, and the multiplicity of an eigenvalue λ as a root of the characteristic polynomial bounds the dimension of E_λ.

§9.3.1 *The definition of the determinant*

What are the key properties defining the determinant?

The determinant

Theorem 9.14. *There is a unique function*

$$\det\colon \mathrm{M}_n(\mathbf{R}) \to \mathbf{R}$$

satisfying the following properties.

1. *The determinant of I_n is 1; i.e.,* $\det I_n = 1$.
2. *If $A \sim B$ via a single row swap, then* $\det A = -\det B$.
3. *If $A \sim B$ via a single row replacement, then* $\det A = \det B$.
4. *If $A \sim B$, where A is obtained from B via a single row scaling by r (here, r can be* any *real number), then* $\det A = r \det B$.

Row equivalence does not allow scaling by 0, but we need this property even when $r = 0$ in the definition of the determinant.

Where did this list come from? We already have a determinant for 2×2 matrices, and we don't have to convince you that row operations are important. You can check (see the reading question below) that all of the items in the above theorem are true for 2×2 matrices. It turns out that if you want a determinant for $n \times n$ matrices that satisfies the same list of properties, then there is only one such function. It's complicated to compute, but take comfort in the fact that there just isn't any alternative. For our present work, the key thing about the determinant is that it detects whether a matrix is singular (see Theorem 9.16 below).

RQ **Reading Question 9J.** Using the formula we already found for it, verify the above properties for determinants of 2×2 matrices. That is, first show that $\det(I_2) = 1$. Then, start with

$$B = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

apply a single row operation, and compare $\det B$ to the determinant of the result (do this for each of the three row operations).

In the next example, we present an algorithm for computing 3×3 determinants. This strategy can be justified using Theorem 9.14, but it's a bit irritating to write it up. The procedure generalizes to the $n \times n$ case; we hope the method will be clear from the example.

To save space, we will enclose a matrix in vertical bars to indicate its determinant instead of using det. For example,

$$\det \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}.$$

To compute a determinant, you need to know what a matrix minor is and what the sign pattern matrix is.

A matrix minor is obtained from any matrix entry by blocking out the row and column the entry is in. For example, the minor of

$$\begin{bmatrix} 2 & \boxed{3} & -1 \\ 4 & -2 & 0 \\ 0 & 1 & 6 \end{bmatrix}$$

corresponding to the entry 3 is

$$\begin{bmatrix} \blacksquare & \blacksquare & \blacksquare \\ 4 & \blacksquare & 0 \\ 0 & \blacksquare & 6 \end{bmatrix}.$$

Note that a minor is smaller than the original matrix (ignoring the black boxes, the above minor is 2×2). The algorithm below will show you how to compute determinants in terms of minor determinants, so it's a recursive procedure that "ends" once you hit minors of size 2×2, because you already know how to compute those determinants.

The sign pattern matrix has a + in the upper left corner and the entries alternate between + and − across each row and down each column. You can write it down next to your matrix if you can't remember it. The 3×3 sign pattern matrix is

$$\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}.$$

Example 9.15 (3 × 3 determinants)**.** Here's how you compute det A when A is 3 × 3.

1. Pick *any* row or column in the matrix. *Usually, you should pick the row or column with the most zeros.*
2. The determinant is the sum of three terms, one for each entry in the row or column you chose. Each term has the form:

$$(\text{sign})(\text{entry})|\text{minor}|,$$

   where the sign comes from the entry's corresponding sign in the sign pattern matrix.
3. Compute the minor determinants and simplify the result.

Let's compute the determinant of

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 4 & -2 & 0 \\ 0 & 1 & 6 \end{bmatrix}$$

Here's what you get if you do it across the top row:

$$\begin{aligned} &+ (2) \begin{vmatrix} \blacksquare & \blacksquare & \blacksquare \\ \blacksquare & -2 & 0 \\ \blacksquare & 1 & 6 \end{vmatrix} - (3) \begin{vmatrix} \blacksquare & \blacksquare & \blacksquare \\ 4 & \blacksquare & 0 \\ 0 & \blacksquare & 6 \end{vmatrix} + (-1) \begin{vmatrix} \blacksquare & \blacksquare & \blacksquare \\ 4 & -1 & \blacksquare \\ 0 & 1 & \blacksquare \end{vmatrix} \\ &= (2)(-2 \cdot 6 - 1 \cdot 0) - (3)(4 \cdot 6 - 0 \cdot 0) + (-1)(4 \cdot 1 - 0 \cdot -1) \\ &= -100 \end{aligned}$$

Here's what happens if you do it down the last column:

$$\begin{aligned} &+ (-1) \begin{vmatrix} \blacksquare & \blacksquare & \blacksquare \\ 4 & -2 & \blacksquare \\ 0 & 1 & \blacksquare \end{vmatrix} - (0)|\text{doesn't matter}| + (6) \begin{vmatrix} 2 & 3 & \blacksquare \\ 4 & -2 & \blacksquare \\ \blacksquare & \blacksquare & \blacksquare \end{vmatrix} \\ &= (-1)(4) - 0 + 6(-16) \\ &= -100 \end{aligned}$$

You can follow exactly the same procedure for, say, a 4×4 matrix. The difference is that there will be four terms and you will have to compute the determinants of four 3×3 minors. But you know how to do that now!

Reading Question 9K. Pick a random 3×3 matrix with no zeros in it and compute its determinant. Check your answer with a computer. Now pick a random matrix with a column with exactly two nonzero entries and use that column to compute the determinant. Again, check your answer with a computer. The second case should be easy.

RQ

Properties of the determinant §9.3.2

When we studied row operations, we noted that starting with any matrix A, you can put it into a REF B using only swaps and replacements; in this case, $|\det A| = |\det B|$. Further, if B is in REF and has a row of all zeroes, then Theorem 9.14 item ④ implies

$$\det B = 0 \det B = 0$$

since you can multiply the row of all zeros in B by zero and not change B. Since a square matrix in REF has a row of all zeros if and only if it is singular, we have proved the following.

A MATRIX IS SINGULAR IFF ITS DETERMINANT IS ZERO

Theorem 9.16. *An $n \times n$ matrix A is singular if and only if $\det A = 0$.*

You can add to the Isomorphism Theorem that A is invertible iff $\det A \neq 0$.

Now suppose A is an invertible (nonsingular) triangular matrix with diagonal entries $s_1, \ldots, s_n$. Using only replacement operations, A is row equivalent to the diagonal matrix $D = \operatorname{diag}(s_1, \ldots, s_n)$. So, $\det A = \det D$. Further, D is obtained from the identity matrix from n scaling operations: scale the top row of I_n by s_1, then the next row by s_2, etc. By Theorem 9.14 item ④,

$$\det \operatorname{diag}(s_1, \ldots, s_n) = s_1 s_2 \cdots s_n.$$

The determinant and REF

How do you compute the determinant of a triangular matrix?

Theorem 9.17. *Suppose A is an $n \times n$ matrix and B is an REF of A obtained using any number of row replacements and k row swaps. If $s_1, \ldots, s_n$ are the diagonal entries in B, then*

$$\det A = (-1)^k s_1 \cdots s_n.$$

In particular, the determinant of a triangular matrix is the product of its diagonal entries.

The next property of the determinant is one of the things that makes it most useful; we omit the proof (it follows from Theorem 9.14 and a technical, fiddly argument).

The determinant is multiplicative

The determinant of a product is the product of the determinants.

Theorem 9.18. *If A and B are $n \times n$ matrices, then* $\det AB = (\det A)(\det B)$.

RQ

Reading Question 9L. Use the definition of invertible, the defining properties of the determinant, and Theorem 9.18 to prove that

$$\det A^{-1} = \frac{1}{\det A}$$

for any $n \times n$ matrix A.

Exercise 9I. Prove Theorem 9.18 in the following special cases: (1) when A or B is not invertible; (2) when A and B are both upper triangular.

Exercise 9J. Though AB is not always the same as BA, what can you say about $\det(AB)$ and $\det(BA)$?

Exercise 9K. Suppose $\det(A) = 2$ and $\det(B) = -3$. Using the fact that a matrix and its transpose have the same determinant, compute: $\det(A^{-1})$, $\det(AB)$, $\det(AB^{-1})$, $\det(B^{\mathrm{T}}A)$, $\det(A^4)$, and $\det(B^{-1}AB)$.

Exercise 9L. True or false: $\det(A + B) = \det(A) + \det(B)$. Give a proof or provide a specific counter-example.

Exercise 9M. Let X be a square matrix with real entries such that $X^{\mathrm{T}}X = I$. Argue that $\det(X) = \pm 1$.

Exercise 9N. Is there a square matrix X with real entries such that $\det(X^{\mathrm{T}}X) = -3$?

Exercise 9O. What's the procedure for using the determinant to decide whether a set of three vectors in $\mathbf{R}^3$ is linearly independent?

Exercise 9P. What must be true about n so that an $n \times n$ matrix A satisfies $\det(-A) = -\det(A)$?

The characteristic polynomial §9.3.3

Recall that λ is an eigenvalue for an $n \times n$ matrix A if and only if $\lambda I_n - A$ is singular. By Theorem 9.16, λ is an eigenvalue if and only if $\det(\lambda I_n - A) = 0$.

It is not too difficult to see that

$$p_A(\lambda) = \det(\lambda I_n - A)$$

is a polynomial of degree n in the variable λ; this is called the **characteristic polynomial of** A. The equation

How are the characteristic polynomial and equation of a matrix defined?

$$\det(\lambda I_n - A) = 0$$

is called the **characteristic equation**. We already know that the characteristic polynomial of

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is

$$p(\lambda) = \lambda^2 - (a + d)\lambda + ad - bc.$$

Here is an example in the 3×3 case.

Example 9.19. Let

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

Then

$$\lambda I_3 - A = \begin{bmatrix} \lambda - 1 & 0 & -3 \\ -4 & \lambda - 7 & 0 \\ -5 & 0 & \lambda + 1 \end{bmatrix}$$

Using the middle row to compute the determinant, there is only one nonzero term corresponding to the entry $\lambda - 7$ whose associated sign from the sign pattern matrix is +. So

$$\begin{aligned} \det(\lambda I_3 - A) &= +(\lambda - 7) \begin{vmatrix} \lambda - 1 & \blacksquare & -3 \\ \blacksquare & \blacksquare & \blacksquare \\ -5 & \blacksquare & \lambda + 1 \end{vmatrix} \\ &= (\lambda - 7)((\lambda - 1)(\lambda + 1) - (-3)(-5)). \\ &= (\lambda - 7)(\lambda^2 - 16) \\ &= (\lambda - 7)(\lambda - \sqrt{16})(\lambda + \sqrt{16}). \end{aligned}$$

The three roots of our characteristic polynomial are $\lambda = 7, \pm 16$. Since they are distinct, there is a basis for $\mathbf{R}^3$ consisting of eigenvectors for A (put another way, A is diagonalizable).

As we have already mentioned, for large matrices there are more sophisticated techniques for finding eigenvalues and eigenvectors. But the mere fact that the eigenvalues are the roots of a degree n polynomial with real coefficients gives us some information.

What's the multiplicity of a root?

A root r of a polnomial p has **multiplicity** m if $p(\lambda) = (\lambda - r)^m q(\lambda)$ for some polynomial q that does *not* have r as a root. The multiplicity is the number of "times" r appears as a root if you factor the polynomial as much as you can. A real polynomial of degree n can have anywhere between 0 and n real roots counting multiplicity. It is a fact (called the fundamental theorem of algebra) that if you work over the complex numbers, then every polynomial of degree n has exactly n roots (counting multiplicity). Therefore we see that *every matrix has eigenvalues*, but those eigenvalues might be complex. For example, if the characteristic polynomial factors as

$$(\lambda + 5)^2(\lambda - 7)(\lambda^2 + 4),$$

then the only real roots are $\lambda = -5$ (with multiplicity 2) and $\lambda = 7$ (with multiplicity 1). This polynomial factors further over $\mathbf{C}$ as

$$(\lambda+5)^2(\lambda-7)(\lambda-2i)(\lambda+2i).$$

The complex roots $2i, -2i$, each with multiplicity 1, were found using the quadratic formula.

It turns out that, just like in the example in the previous paragraph, every polynomial with real coefficients factors as a product of linear factors and quadratic factors, where the quadratic factors have no real roots. (Factors might be repeated many times.) For example,

$$x^5 - 5x^4 - x + 5 = (x-5)(x^4+1) = (x-5)(x^2+\sqrt{2}x+1)(x^2-\sqrt{2}x+1).$$

You can check using the quadratic formula that the two quadratics above don't have real roots. Let's call a number **purely complex** if it is a complex number that's not a real number. According to the quadratic formula, a quadratic with purely complex roots has roots of the form

What's a purely complex number?

$$x = \frac{b}{2a} \pm i\frac{\sqrt{|b^2-4ac|}}{2a}.$$

So, purely complex roots to real polynomials always come in pairs (called conjugate pairs), just like in the example above.

The complex numbers $a+bi$ and $a-bi$ are called conjugate.

Let's think about what can happen with the characteristic polynomial p for a 3×3 real matrix A. It either has three real roots counting multiplicity or it has exactly one real root r and a pair of (distinct) purely complex roots.

- If p has three distinct real roots, then A is diagonalizable.
- If p has two distinct real roots, then one must have multiplicity 2. To check whether A is diagonalizable, you have to compute the dimensions of the eigenspaces (they need to sum to 3).
- If p has one real root r of multipliclity 3, then it's diagonaliable if and only if $\dim E_r = 3$.
- If p has one real root and a conjugate pair of purely complex roots, then it is not diagonalizable (over $\mathbf{R}$). However, it is diagonalizable over $\mathbf{C}$ (though this is not what we mean by "diagonalizable" in this book).

We already know that similar matrices have the same eigenvalues, but similar matrices also have exactly the same characterisitic polynomial, so they have the same eigenvalues *counting multiplicity*. Further, $\dim E_\lambda$ cannot exceed the multiplicity of λ as a root of the characteristic polynomial, as we prove below.

Similar matrices have the same eigenvalues counting multiplicity, and the dimension of an eigenspace can't exceed the multiplicity of the eigenvalue.

Characteristic polynomial facts

Theorem 9.20. *Let A and B be $n \times n$ matrices with characteristic polynomials p_A and p_B.*

1. *If A is similar to B, then $p_A = p_B$, so A and B have exactly the same eigenvalues counting multiplicity.*
2. *If λ is an eigenvalue of A, then $\dim E_\lambda \leq m_\lambda$, where m_λ is the multiplicity of λ as a root of p_A.*

Proof. Suppose A is similar to B, so that there is an invertible matrix P such that $A = PBP^{-1}$. Using the fact that the determinant is multiplicative and Reading Question 9L, we have

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I_n - A) \\ &= \det(\lambda I_n - PBP^{-1}) \\ &= \det(\lambda P I_n P^{-1} - PBP^{-1}) \\ &= \det(P(\lambda I_n - B)P^{-1}) \\ &= (\det P) p_B(\lambda) \frac{1}{\det P} \\ &= p_B(\lambda). \end{aligned}$$

Since A and B have exactly the same characteristic polynomial, they have exactly the same eigenvalues counting multiplicity.

Next, suppose λ_0 is an eigenvalue of A. Let $\mathcal{B} = \{\mathbf{v}_1, \cdots \mathbf{v}_k, \mathbf{v}_{k+1}, \cdots \mathbf{v}_n\}$ be a basis for $\mathbf{R}^n$ where the first k vectors form a basis for E_{λ_0} (we know there is such a basis, because we can start with a basis for E_{λ_0} and then enlarge it by Theorem 6.9). Let

$$P = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_k \ \mathbf{v}_{k+1} \ \cdots \ \mathbf{v}_n].$$

Check, for example, that $(P^{-1}AP)\mathbf{e}_1 = \lambda_0 \mathbf{e}_1$.

Now A has the same eigenvalues counting multiplicity as

$$B = P^{-1}AP = [\lambda_0 \mathbf{e}_1 \ \cdots \ \lambda_0 \mathbf{e}_k \ \underbrace{* \ \cdots \ *}_{n-k \text{ cols}}]$$

(it doesn't matter what the last $n-k$ columns look like). More concisely, B has the form

$$B = \begin{bmatrix} \lambda_0 I_k & X \\ 0 & Y \end{bmatrix},$$

where X is $k \times (n-k)$ and Y is $(n-k) \times (n-k)$. So

$$\lambda I_n - B = \begin{bmatrix} (\lambda - \lambda_0) I_k & -X \\ 0 & \lambda I_{n-k} - Y \end{bmatrix}.$$

Using the general version of the algorithm for computing the determinant that we described for 3×3 matrices, one can show that the determinant of the matrix above is the product of the determinants of the matrices on the main diagonal (the matrices $(\lambda - \lambda_0)I_k$ and $\lambda I_{n-k} - Y$). So

$$\begin{aligned} p_A(\lambda) &= p_B(\lambda) \\ &= \det(\lambda I_n - B) \\ &= (\det(\lambda - \lambda_0)I_k)(\det(\lambda I_{n-k} - Y)) \\ &= (\lambda - \lambda_0)^k p_Y(\lambda). \end{aligned}$$

This proves that $\dim E_{\lambda_0} = k$ less than or equal to the multiplicity of λ_0 as a root of p_A. These numbers might not be equal because λ_0 might also be a root of p_Y. ■

Most authors define the characteristic polynomial of an $n \times n$ matrix to be $\det(A - \lambda I_n)$. This definition changes nothing about the characteristic polynomial except its sign (if n is odd). In partciular, whether you define the characteristic polynomial as $\det(\lambda I_n - A)$ or $\det(A - \lambda_n I)$ does not change its roots or the multiplicities of those roots. Under our definition, the "leading term" of the characteristic polynomial is always λ^n (which Professors Lockridge and Kennedy find soothing) instead of being $(-1)^n \lambda^n$ (which Professors Lockridge and Kennedy find irritating). But be alert when consulting other sources.

Exercise 9Q. Explain why an $n \times n$ matrix can have at most n distinct eigenvalues.

Exercise 9R. Suppose the characteristic polynomial of a real 4×4 matrix A has a purely complex root. What is the largest size a set of linearly independent eigenvectors in $\mathbf{R}^4$ for A can be, and what is the smallest? Answer the same question for a 5×5 matrix.

Exercise 9S. In each case below, determine whether the real matrix A must be diagonalizable, must not be diagonalizable, or there is not enough information given to tell.

① A is 3×3 and $p_A(\lambda) = (\lambda - 7)(\lambda^2 - 4)$

② A is 3×3 and $p_A(\lambda) = (\lambda - 7)(\lambda^2 + 4)$

③ A is 3×3 and $p_A(\lambda) = (\lambda - 7)(\lambda - 4)^2$

④ A is 7×7, $\lambda = 5$ is a root of p_A with multiplicity 4, and $\dim E_5 = 3$

§9.4 Dynamical systems revisited

Theorem 9.10 shows that the analysis of dynamical systems $\mathbf{x} \mapsto A\mathbf{x}$ is straightforward in the case where A is diagonalizable. Here we circle back to two examples.

§9.4.1 *Lionfish*

This is a follow-up to Example 1.2, Example 2.7, and §4.2.1. The matrix we used to model lionfish subpopulations is

$$A = \begin{bmatrix} 0.00000 & 0.00000 & 35315. \\ 0.000030000 & 0.77700 & 0.00000 \\ 0.00000 & 0.071000 & 0.94900 \end{bmatrix}.$$

Using a computer, the eigenvalues and some associated eigenvectors are:

These are not actually equalities, they are approximations.

$$\begin{aligned} \lambda_1 &= 0.15026 \\ \mathbf{v}_1 &= \left(1.0000,\ -0.000047867,\ 4.2549 \times 10^{-6}\right) \\ \lambda_2 &= 0.44126 \\ \mathbf{v}_2 &= (1.0000,\ -0.000089355,\ 0.000012495) \\ \lambda_3 &= 1.1345 \\ \mathbf{v}_3 &= (1.0000,\ 0.000083921,\ 0.000032125) \end{aligned}$$

So A has three distinct eigenvalues and the eigenvectors above form a basis for $\mathbf{R}^3$. Given any initial state vector

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3,$$

the Eigenvector Basis Theorem tells us that

$$\mathbf{x}_k = A^k\mathbf{x}_0 = c_1(0.15026)^k\mathbf{v}_1 + c_2(0.44126)^k\mathbf{v}_2 + c_3(1.1345)^k\mathbf{v}_3.$$

As $k \longrightarrow \infty$, the first two terms go to zero. So, for large k,

$$\mathbf{x}_k \approx c_3(1.1345)^k\mathbf{v}_3.$$

Now you can see why all three subpopulations tend to infinity in this model. Further, for *any* of the subpopulations, the ratio of the population at month $k+1$ to the population at month k is tending to the eigenvalue 1.1345; so, this eigenvalue of largest magnitude captures the long-term growth rate of the subpopulations!

The entry 0.071 in the matrix A is the rate at which juveniles become adults. The lionfish is an invasive species; suppose we are interested in population control. What would happen if we were able to decrease this rate significantly, to say 0.01? The new matrix model would be:

$$B = \begin{bmatrix} 0.00000 & 0.00000 & 35315. \\ 0.000030000 & 0.77700 & 0.00000 \\ 0.00000 & 0.010000 & 0.94900 \end{bmatrix}.$$

This time, the eigenvalues and eigenvectors are:

$$\begin{aligned} \lambda_1 &= 0.014882 \\ \mathbf{v}_1 &= \left(1.0000,\ -0.000039364,\ 4.2140 \times 10^{-7}\right) \\ \lambda_2 &= 0.71388 \\ \mathbf{v}_2 &= (1.0000,\ -0.00047529,\ 0.000020215) \\ \lambda_3 &= 0.99724 \\ \mathbf{v}_3 &= (1.0000,\ 0.00013622,\ 0.000028238) \end{aligned}$$

The Eigenvector Basis Theorem tells us that the new system evolves according to the formula

$$\mathbf{x}_k = B^k \mathbf{x}_0 = c_1(0.014882)^k \mathbf{v}_1 + c_2(0.71388)^k \mathbf{v}_2 + c_3(0.99724)^k \mathbf{v}_3.$$

Now all the terms drive to zero, and so the long-run behavior suggests a declining lionfish population.

The Fibonacci sequence §9.4.2

This is a follow-up to §4.2.2. Let's find a closed formula for the Fibonacci sequence

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots$$

defined by the recursive formula:

$$F_0 = 0,\ F_1 = 1,\ \text{and } F_n = F_{n-1} + F_{n-2} \text{ for } n \geq 2.$$

In §4.2.2, we defined a matrix

$$M = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

and observed that

$$M^{n-1} \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix}.$$

Let's compute the eigenvalues of M. The characteristic equation is

$$\lambda^2 - \lambda - 1 = 0,$$

and using the quadratic formula we find that its roots are $\phi = (1 + \sqrt{5})/2$ and

The number ϕ is called the golden ratio.

$\psi = (1 - \sqrt{5})/2$. Here are their associated eigenspaces:

You know *the spaces have dimension 1, so you don't need to actually do row replacement to zero out the bottom row!*

$$E_\phi = \ker \begin{bmatrix} \phi - 1 & -1 \\ -1 & \phi \end{bmatrix} = \ker \begin{bmatrix} 1 & -\phi \\ 0 & 0 \end{bmatrix} = \operatorname{span}\left\{ \begin{bmatrix} \phi \\ 1 \end{bmatrix} \right\}$$
$$E_\psi = \operatorname{span}\left\{ \begin{bmatrix} \psi \\ 1 \end{bmatrix} \right\}.$$

We can write the seed (F_1, F_0) as a linear combination of the eigenvectors we found (check that $\phi - \psi = \sqrt{5}$):

$$\begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \phi \\ 1 \end{bmatrix} - \frac{1}{\sqrt{5}} \begin{bmatrix} \psi \\ 1 \end{bmatrix}.$$

So,

$$\begin{aligned} \begin{bmatrix} F_n \\ F_{n-1} \end{bmatrix} &= M^{n-1} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= M^{n-1} \left(\frac{1}{\sqrt{5}} \begin{bmatrix} \phi \\ 1 \end{bmatrix} - \frac{1}{\sqrt{5}} \begin{bmatrix} \psi \\ 1 \end{bmatrix} \right) \\ &= \frac{1}{\sqrt{5}} M^{n-1} \begin{bmatrix} \phi \\ 1 \end{bmatrix} - \frac{1}{\sqrt{5}} M^{n-1} \begin{bmatrix} \psi \\ 1 \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \phi^{n-1} \begin{bmatrix} \phi \\ 1 \end{bmatrix} - \frac{1}{\sqrt{5}} \psi^{n-1} \begin{bmatrix} \psi \\ 1 \end{bmatrix}. \end{aligned}$$

Take the top entry in the equation above to get

$$\begin{aligned} F_n &= \frac{1}{\sqrt{5}} \phi^n - \frac{1}{\sqrt{5}} \psi^n \\ &= \frac{1}{\sqrt{5}} \left[\left(\frac{1 + \sqrt{5}}{2} \right)^n - \left(\frac{1 - \sqrt{5}}{2} \right)^n \right]. \end{aligned}$$

(It is striking that the above expression always evaluates to be an integer!)

Please take moment to be impressed by this. You may have seen this result before; maybe you even proved it by induction in Math 215 or a similar course. But for inductive proofs to work, you have to already suspect what the answer is; here we set out to find a formula for F_n and *discovered* one. Freakin' amazing.

Exercise 9T. Find a formula for the kth term of the following "Fibonacci-like" sequence:

$$S_0 = 0, \; S_1 = 1, \; S_n = 2S_{n-2} + S_{n-1} \text{ for } n \geq 2.$$

Exercise 9U. It is Senior Week in Ocean City. There are five parties along the boardwalk, and every hour the population of partying graduates moves among the various parties. Parties 1 and 5 are ragers, so all students at parties 1 or 5 stay there. Students at the other parties move to either neighboring party with probability 1/2. See Figure 9.2.

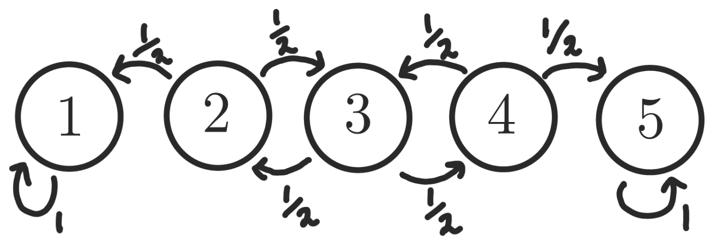


Figure 9.2: Party time

Consider the following matrix:

$$T = \begin{bmatrix} 1 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 1 \end{bmatrix}$$

Suppose that the ith entry of the vector $\mathbf{p} = (p_1, p_2, p_3, p_4, p_5)$ gives the proportion of students at party i in the current time period. Then the ith entry of the vector $T\mathbf{p}$ gives the proportion of students at party i in the following time period.

[This is an example of a type of dynamical system called a *Markov chain* (Exercise 9H is another example). In this context, T is called a *transition matrix* and p is called a *probability vector*. This particular Markov chain is actually called a *drunkard's walk*, but Professors Lockridge and Kennedy both advocate responsible fun during Senior Week.]

(i) Let $\mathbf{p} = (1/10, 2/10, 3/10, 4/10, 0/10)$. By hand, compute $T\mathbf{p}$, and $T^2\mathbf{p}$. Do these computations both by actually performing the matrix multiplications and just by using the illustration above. Make sure you believe and understand the above interpretation of $T\mathbf{p}$ (that it gives the population in each site next time period if $\mathbf{p}$ gives the population in each site in the current time period).

(ii) Suppose a large group of friends begins at party number 2. After a long time, what percentage of these friends will wind up where? Use a computer to help you do and check your calculations, but write up an analytic proof of your answer.

Key concepts

- The similarity classification of 2×2 matrices according to the nature of their eigenvalues, and how to interpret the results geometrically
- Manipulating complex eigenvalues and eigenvectors for 2×2 matrices
- The Spectral Theorem for 2×2 matrices and what it means to be orthogonally diagonalizable
- The matrix of a linear transformation $V \longrightarrow W$ with repect to chosen bases for the domain and codomain

Summary. Recall that if a matrix A is similar to a matrix B, then $A = PBP^{-1}$, and if you let $\mathcal{B}$ be the set of columns of P, then the $\mathcal{B}$-matrix of A is B. So, in order to gain qualitative understanding of the dynamical system $\mathbf{x} \mapsto A\mathbf{x}$, you can work in $\mathcal{B}$-coordinates and study the dynamical system $\mathbf{x} \mapsto B\mathbf{x}$.

In this chapter, we completely classify dynamical systems for 2×2 matrices up to similarity in the sense of the previous paragraph. There are three cases: (I) the matrix has two real eigenvalues s, t (counting multiplicity) and is diagonalizable; (II) the matrix has one real eigenvalue λ of multiplicity 2 and is not diagonalizable; and (III) the matrix has two purely complex eigenvalues $\lambda = a \pm bi$ where θ is the angle (a, b) makes with the positive x-axis (CCW). In each case, the matrix is similar to something nice:

I. $A \sim \mathrm{S}_{s,t} = \begin{bmatrix} s & 0 \\ 0 & t \end{bmatrix}$;

II. $A \sim J_\lambda = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} = \mathrm{S}_\lambda \, \mathrm{H}^x_{1/\lambda}$ (the second equality holds when $\lambda \neq 0$);

III. $A \sim C_\lambda = \mathrm{S}_{|\lambda|} \, \mathrm{R}_\theta$.

We also prove the Spectral Theorem for 2×2 matrices, which says symmetric matrices are orthogonally diagalizable. Finally, since this chapter is about similarity, we discuss a version of similarity for abstract vectors spaces by introducing the matrix of a linear transformation $V \longrightarrow W$ with repsect to chosen bases for the domain and codomain.

Chapter 10

Our goal in this chapter is to classify and visualize discrete dynamical systems $\mathbf{x} \mapsto A\mathbf{x}$ when A is a 2×2 matrix. In particular, we will classify 2×2 matrices "up to similarity".

Remember, if A is similar to B, then there is an invertible matrix P such that $A = PBP^{-1}$. If we let $\mathcal{B}$ be the set of columns of P, then P^{-1} represents the $\mathcal{B}$-coordinate mapping and

$$B[\mathbf{x}]_{\mathcal{B}} = [A\mathbf{x}]_{\mathcal{B}}$$

for all $\mathbf{x}$. This means the $\mathcal{B}$-matrix of $\mathbf{x} \mapsto A\mathbf{x}$ is B, and we think about the difference between these two matrices as follows: *B is just another way to represent the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ if you're willing to work entirely in a $\mathcal{B}$-coordinate system.* If you change your basis from the standard basis to $\mathcal{B}$, the matrix A "becomes" the matrix B. Similar comments apply to the powers of A since $A^k = PB^kP^{-1}$.

Our main question for this chapter is the following: given a 2×2 matrix A, is there a way to pick a different basis so that the behavior of the dynamical system $\mathbf{x} \mapsto A\mathbf{x}$ is more transparent? In other words: *can we find a matrix B that is similar to A, and whose powers B^k are easy to compute and understand?* The answer is **yes**. These matrices B come in three essential types, depending on the eigenvalues and diagonalizability of A.

As you know, the matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

has characteristic polynomial

$$\lambda^2 - (a + d)\lambda + ad - bc.$$

You probably noticed that $a + d$ is the **trace** of A (that is, the sum of the diagonal entries), written $\operatorname{tr} A$, and $ad - bc$ is the determinant of A, $\det A$. So, more concisely,

$$p_A(\lambda) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A.$$

The roots of this polynomial, according to the quadratic formula, are

$$\lambda = \frac{1}{2}\left(\operatorname{tr} A \pm \sqrt{(\operatorname{tr} A)^2 - 4\det A}\right).$$

There are three interesting cases here; they will correspond to the three types of "nice" matrices to which A can be similar.

I. A has two real eigenvalues (counting multiplicity) and is diagonalizable. This certainly holds when $(\operatorname{tr} A)^2 > 4\det A$ because then the eigenvalues are distinct, and it may or may not hold when $(\operatorname{tr} A)^2 = 4\det A$.

II. A has two real eigenvalues (counting multiplicity) and is *not* diagonalizable. This happens only when $(\operatorname{tr} A)^2 = 4\det A$ (though again, $(\operatorname{tr} A)^2 = 4\det A$ does not alone tell you whether A is diagonalizable).

III. A has a pair of purely complex (conjugate) eigenvalues. In this case, A is not diagonalizable (over $\mathbf{R}$) and $(\operatorname{tr} A)^2 < 4\det A$.

Remember, when we say that a real matrix A isn't diagonalizable, we mean that there is no invertible matrix P with *real* entries such that $P^{-1}AP$ is diagonal. The matrices in case II are not diagonalizable over the real numbers *and* they are not diagonalizable over the complex numbers. The matrices in case III, on the other hand, are not diagonalizable over the real numbers but they are diagonalizable over the complex numbers. In this chapter, we are only interested in which *real* matrices a real matrix can be similar to because when we change coordinates we want the resulting matrix to still determine a transformation from $\mathbf{R}^2$ to $\mathbf{R}^2$.

RQ **Reading Question 10A.** Explain why a 2×2 matrix with negative determinant is diagonalizable.

Exercise 10A. The **trace** of an $n \times n$ matrix A, denoted $\operatorname{tr} A$, is the sum of the diagonal entries of A. Here are two facts about the trace: (1) $\operatorname{tr} XY = \operatorname{tr} YX$ for any pair of $n \times n$ matrices X and Y; and (2) the trace of a matrix is the sum of its eigenvalues (counting multiplicity). Use fact (1) to argue that similar matrices have the same trace. Verify fact (2) for matrices that are diagonalizable.

Diagonalizable matrices §10.1

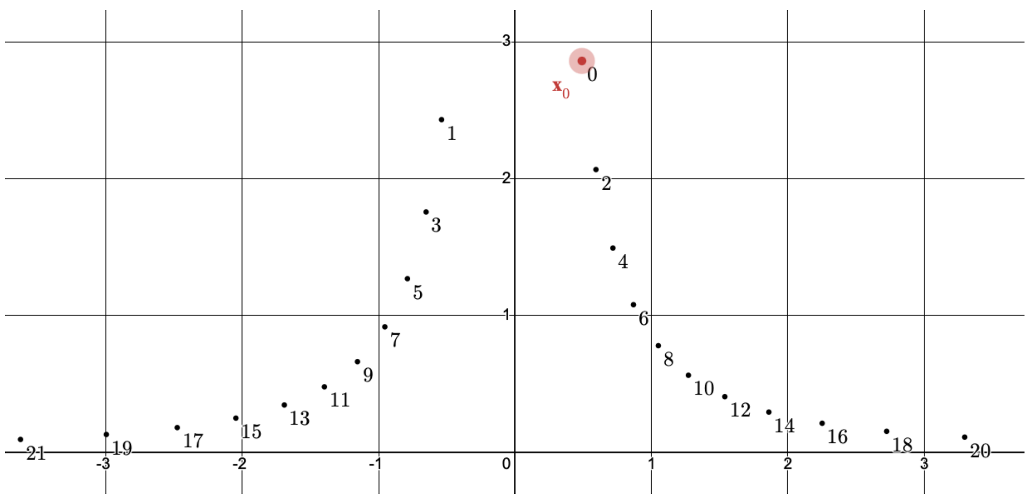
If A is diagonalizable, then it is similar to a diagonal matrix

$$S_{s,t} = \begin{bmatrix} s & 0 \\ 0 & t \end{bmatrix}$$

(where s, t are the (not necessarily distinct) eigenvalues of A). To see how $S_{s,t}$ behaves when viewed as a discrete dynamical system, we can pick an initial condition $\mathbf{x}_0 = (q_1, q_2)$ and plot

$$\mathbf{x}_k = S_{s,t}^k \mathbf{x}_0 = \begin{bmatrix} s^k & 0 \\ 0 & t^k \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} s^k q_1 \\ t^k q_2 \end{bmatrix}.$$

It helps to consider various initial conditions and values of s, t, with differing magnitudes and signs. For example, when $s = -1.1$ and $t = 0.85$, the solution to the dynamical system looks like Figure 10.1 below. The point $\mathbf{x}_k$ is labeled simply as k.



www.desmos.com/calculator/0hof7vdk23

Figure 10.1: Dynamical system with matrix $S_{-1.1, 0.85}$ (Desmos)

Since the sign of s^k alternates and $|s|^k \to \infty$ as $k \to \infty$, and since t^k is always positive and $t^k \to 0$, the solution $\mathbf{x}_k$ alternates between the first and second quadrant while asymptotically approaching the x-axis.

Reading Question 10B. Go to the above Desmos illustration and experiment with various values of s and t. Find specific values where the behavior of the dynamical system is qualitatively different from the example above.

RQ

§10.2 Non-diagonalizable matrices with a real eigenvalue

Suppose A has a real eigenvalue λ of multiplicity two. If A is not diagonalizable, then the dimension of the eigenspace E_λ is 1. Let $\{\mathbf{v}\}$ be a basis for this eigenspace (so, $\mathbf{v}$ is an eigenvector with eigenvalue λ). The set $\{\mathbf{v}\}$ may be completed to a basis $\mathcal{B} = \{\mathbf{v}, \mathbf{w}\}$ for $\mathbf{R}^2$ (just pick any vector $\mathbf{w}$ that's not a multiple of $\mathbf{v}$).

RQ

Reading Question 10C. Show that the $\mathcal{B}$-matrix for $\mathbf{x} \mapsto A\mathbf{x}$ has the form

$$B = \begin{bmatrix} \lambda & s \\ 0 & t \end{bmatrix}$$

where s and t are real numbers with $s \neq 0$.

From the reading question above, we have:

$$A\mathbf{v} = \lambda\mathbf{v} + 0\mathbf{w}$$
$$A\mathbf{w} = s\mathbf{v} + t\mathbf{w}.$$

We can make B look a little nicer by adjusting our basis. The above equations imply

$$A\mathbf{v} = \lambda\mathbf{v} + 0 \cdot \frac{1}{s}\mathbf{w}$$
$$A\left(\frac{1}{s}\mathbf{w}\right) = \mathbf{v} + t\left(\frac{1}{s}\mathbf{w}\right).$$

So if we replace $\mathcal{B}$ with $\{\mathbf{v}, (1/s)\mathbf{w}\}$, the $\mathcal{B}$-matrix is now

$$\begin{bmatrix} \lambda & 1 \\ 0 & t \end{bmatrix}.$$

But it gets even better: since this matrix is upper triangular, its diagonal entries are its eigenvalues, so we must have $t = \lambda$ (otherwise, the eigenvalues would be distinct and the matrix would be diagonalizable). Now we have that the $\mathcal{B}$-matrix is actually

$$J_\lambda = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}.$$

RQ

Reading Question 10D. If $\lambda \neq 0$, show that the above matrix J_λ is a shearing transformation followed by a uniform scaling transformation. What is the scale factor, and what is the shear factor? If $\lambda = 0$, what does J_0 do geometrically?

What does the dynamical system $\mathbf{x} \mapsto J_\lambda \mathbf{x}$ look like? Well, when $\lambda \neq 0$,

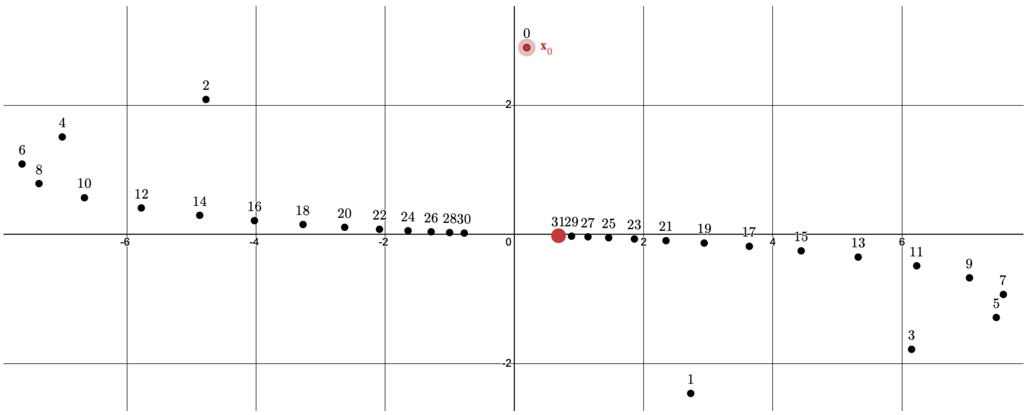
$$J_\lambda^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{bmatrix} = \lambda^k \begin{bmatrix} 1 & k/\lambda \\ 0 & 1 \end{bmatrix}.$$

Compute the first few powers of J_λ to check this formula.

So

$$\mathbf{x}_k = \lambda^k \begin{bmatrix} q_1 + (k/\lambda) q_2 \\ q_2 \end{bmatrix}.$$

Here's an example where $\lambda = -0.85$.



www.desmos.com/calculator/xjf6zrdpg3

Figure 10.2: Dynamical system with matrix $J_{-0.85}$ (Desmos)

Reading Question 10E. Go to the above Desmos illustration and experiment with various values of λ (called t in Desmos). Find specific values where the behavior of the dynamical system is qualitatively different from the example above.

RQ

Exercise 10B. Prove the above formula for J_λ^k: that is, prove that

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix}^k = \begin{bmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{bmatrix}$$

for all positive integers k.

Exercise 10C. Prove that the only 2×2 matrices with a single eigenvalue that *are* diagonalizable are of the form

$$\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

Can you generalize?

§10.3 Matrices with purely complex eigenvalues

When a matrix has a purely complex eigenvalue, it still makes sense to talk about eigenvectors, you just have to be willing to work over the complex numbers instead of the real numbers. Row reduction works just as well for complex matrices, and complex matrices can be reduced to a REF and have a unique RREF. If A is an $n \times n$ matrix over $\mathbf{C}$, then it remains the case that $A\mathbf{z} = \mathbf{0}$ has a unique solution (the zero vector) if and only if the RREF of A is I_n, if and only if $\det A \neq 0$.

What's the conjugate of a complex number?

The **complex conjugate** of $a + bi$ is $\overline{a + bi} = a - bi$.

Example 10.1. Consider the matrix

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

Using the formula we found earlier, the eigenvalues of this matrix are

$$\lambda = \frac{3}{2}(1 + i\sqrt{3})$$

and its conjugate $\overline{\lambda}$. Let's compute the eigenspaces:

$$E_\lambda = \ker(\lambda I_2 - A) = \ker \underbrace{\begin{bmatrix} \lambda - 2 & -1 \\ 1 & \lambda - 1 \end{bmatrix}}_{B}$$

Now, the top left entry of B is a pivot position because $\lambda \neq 2$. But there cannot be a pivot position on the bottom row, because if there were, then B would be row equivalent to I_2 and therefore invertible. But then λ could not be an eigenvalue! So, without doing any calculations, we *know*

$$\ker(\lambda I_2 - A) = \ker \begin{bmatrix} \lambda - 2 & -1 \\ 0 & 0 \end{bmatrix}.$$

There is only one free variable needed to describe the kernel, though it is a complex variable. By inspection, you can see that the vector $\mathbf{v} = (1, \lambda - 2)$ lies in the kernel (check!), so we must have

$$E_\lambda = \left\{ z \begin{bmatrix} 1 \\ \lambda - 2 \end{bmatrix} \mid z \in \mathbf{C} \right\} = \operatorname{span}_{\mathbf{C}} \left\{ \begin{bmatrix} 1 \\ \lambda - 2 \end{bmatrix} \right\} = \operatorname{span}_{\mathbf{C}}\{\mathbf{v}\}.$$

What we just did works equally well for $\overline{\lambda}$, and we obtain that $\overline{\mathbf{v}}$ is an eigenvector for $\overline{\lambda}$.

One key lesson to take away from the previous example is that if a real matrix has a pair of purely complex eigenvalues λ and $\overline{\lambda}$, then if $\mathbf{v}$ is any eigenvector for λ, $\overline{\mathbf{v}}$ is an eigenvector for $\overline{\lambda}$.

Let's go over a little more terminology related to complex numbers and vectors. A complex number $z = a+bi$ has a **real part** $\operatorname{Re} z = a$ and **imaginary part** $\operatorname{Im} z = b$. (Note that the imaginary part is just the real coefficient of i—it does *not* include i!) Similarly, we can define the real part and the imaginary part of a complex vector. For example, if $\mathbf{z} = (1 - 2i, 7i, 6)$, then

What are the real and imaginary parts of a complex number?

$$\operatorname{Re} \mathbf{z} = \begin{bmatrix} 1 \\ 0 \\ 6 \end{bmatrix}$$

$$\operatorname{Im} \mathbf{z} = \begin{bmatrix} -2 \\ 7 \\ 0 \end{bmatrix}.$$

To compute the complex conjugate of a vector, just compute the conjugates of its entries:

$$\overline{z} = \begin{bmatrix} 1 + 2i \\ -7i \\ 6 \end{bmatrix}.$$

So far in this chapter we have found matrices similar to A, when A has real eigenvalues, whose powers are easy to compute and understand (the matrix $S_{s,t}$ in the diagonalizable case, and the matrix J_λ in the non-diagonalizable case). Our goal in this section is to do the same thing in the situation that the eigenvalues of A are purely complex.

Let $\lambda = a - bi$. (Note that we chose to use a minus sign in our definition of λ; this turns out to make the nice matrix we eventually find look even nicer. This doesn't mean that λ has to have negative real part: for example, if you really want to think about $\lambda = 1 + i$, just take $a = 1$ and $b = -1$.)

Let $\mathbf{v}$ be an eigenvector for $\lambda = a - bi$ and define two matrices:

$$C_\lambda = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

$$P = [\operatorname{Re} \mathbf{v} \;\; \operatorname{Im} \mathbf{v}].$$

Let's compute $\lambda \mathbf{v}$ in two ways. (In the calculations below, just multiply out all the terms and remember that $i^2 = -1$.)

First, since $\mathbf{v} = \operatorname{Re} \mathbf{v} + i \operatorname{Im} \mathbf{v}$:

$$\lambda \mathbf{v} = (a - bi)(\operatorname{Re} \mathbf{v} + i \operatorname{Im} \mathbf{v})$$

$$
\begin{aligned}
&= (a \operatorname{Re}\mathbf{v} + b \operatorname{Im}\mathbf{v}) + i(-b \operatorname{Re}\mathbf{v} + a \operatorname{Im}\mathbf{v}) \\
&= P \begin{bmatrix} a \\ b \end{bmatrix} + iP \begin{bmatrix} -b \\ a \end{bmatrix} \\
&= PC_\lambda \mathbf{e}_1 + iPC_\lambda \mathbf{e}_2.
\end{aligned}
$$

This tells you that:

$$
\begin{aligned}
\operatorname{Re}(\lambda\mathbf{v}) &= PC_\lambda \mathbf{e}_1 \\
\operatorname{Im}(\lambda\mathbf{v}) &= PC_\lambda \mathbf{e}_2.
\end{aligned}
$$

Next, we use the fact that $A\mathbf{v} = \lambda\mathbf{v}$:

$$
\begin{aligned}
\lambda\mathbf{v} &= A\mathbf{v} \\
&= A(\operatorname{Re}\mathbf{v} + i \operatorname{Im}\mathbf{v}) \\
&= A(\operatorname{Re}\mathbf{v}) + iA(\operatorname{Im}\mathbf{v}) \\
&= AP\mathbf{e}_1 + iAP\mathbf{e}_2.
\end{aligned}
$$

Putting both computations together, we get

$$
\begin{aligned}
PC_\lambda \mathbf{e}_1 &= AP\mathbf{e}_1 \\
PC_\lambda \mathbf{e}_2 &= AP\mathbf{e}_2.
\end{aligned}
$$

This says that $PC_\lambda = AP$ or, equivalently,

$$A = PC_\lambda P^{-1}.$$

So if we use the basis $\mathcal{B} = \{\operatorname{Re}\mathbf{v}, \operatorname{Im}\mathbf{v}\}$ for $\mathbf{R}^2$, then the $\mathcal{B}$-matrix of A is C_λ.

What does C_λ do, geometrically? Define the magnitude of a complex number $a + bi$ to be its distance to the origin when viewed as a point in the plane:

$$|a + bi| = \sqrt{a^2 + b^2}.$$

Note that $|\lambda| = |\bar{\lambda}|$. We can factor this quantity out of C_λ:

$$C_\lambda = |\lambda| \begin{bmatrix} \underbrace{\begin{matrix} a/|\lambda| \\ b/|\lambda| \end{matrix}}_{\mathbf{u}} & \underbrace{\begin{matrix} -b/|\lambda| \\ a/|\lambda| \end{matrix}}_{\mathbf{v}} \end{bmatrix}.$$

We did this because now the vectors $\mathbf{u}$ and $\mathbf{v}$ lie on the unit circle! So, there is an angle θ such that

$$\left(\frac{a}{|\lambda|}, \frac{b}{|\lambda|} \right) = (\cos\theta, \sin\theta).$$

We now see that

$$C_\lambda = \mathrm{S}_{|\lambda|}\,\mathrm{R}_\theta$$

is rotation by θ radians counter-clockwise about the origin, followed by uniform scaling of the plane by the magnitude of the eigenvalue λ.

Similarity for 2 by 2 matrices with complex eigenvalues

Theorem 10.2. *Suppose A is a 2×2 matrix with a purely complex eigenvalue $\lambda = a - bi$. Let $\mathbf{v}$ be a (complex) eigenvector for λ. Then,*

$$A = PC_\lambda P^{-1},$$

where

$$P = [\operatorname{Re}\mathbf{v} \;\; \operatorname{Im}\mathbf{v}]$$

and

$$C_\lambda = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Further, C acts as a rotation followed by a uniform scaling operation. We may write

$$C_\lambda = |\lambda|\,\mathrm{R}_\theta = \mathrm{S}_{|\lambda|}\,\mathrm{R}_\theta$$

where

$$(\cos\theta, \sin\theta) = \left(\frac{a}{|\lambda|}, \frac{b}{|\lambda|}\right), \quad |\lambda| = \sqrt{a^2 + b^2}$$

and R_θ denotes rotation by θ radians counter-clockwise about the origin.

You can certainly apply this theorem when the eigenvalue has positive imaginary part. For example, if you want to use $\lambda = 1 + i$, take $a = 1$ and $b = -1$.

Note that this theorem uses a complex eigenvalue as an auxiliarly tool; other than the ingredients λ and $\mathbf{v}$ in the setup of the theorem, all the other quantities involved are exclusively real. We used complex numbers to draw conclusions about real matrices! This isn't too much of a surprise: in §8.4.1, we saw that the complex number i has a "real" interpretation as 90° rotation of the plane.

Example 10.3. Let's go back to Example 10.1. We'll use the eigenvalue

$$\lambda = \frac{3}{2} - i\frac{3\sqrt{3}}{2}$$

with associated eigenvector

$$\mathbf{v} = \begin{bmatrix} 1 - \lambda \\ 1 \end{bmatrix} = \begin{bmatrix} -1/2 + i3\sqrt{3}/2 \\ 1 \end{bmatrix}.$$

So

$$P = \begin{bmatrix} -1/2 & 3\sqrt{3}/2 \\ 1 & 0 \end{bmatrix}$$

$$a = 3/2$$

$$b = 3\sqrt{3}/2$$

$$C_\lambda = \begin{bmatrix} 3/2 & -3\sqrt{3}/2 \\ 3\sqrt{3}/2 & 3/2 \end{bmatrix}$$

$$|a - bi| = \sqrt{(3/2)^2 + (3\sqrt{3}/2)^2} = 3.$$

To find the rotation angle, we need to solve for θ:

$$(\cos\theta, \sin\theta) = (1/2, \sqrt{3}/2).$$

The solution is $\theta = \pi/3$. So,

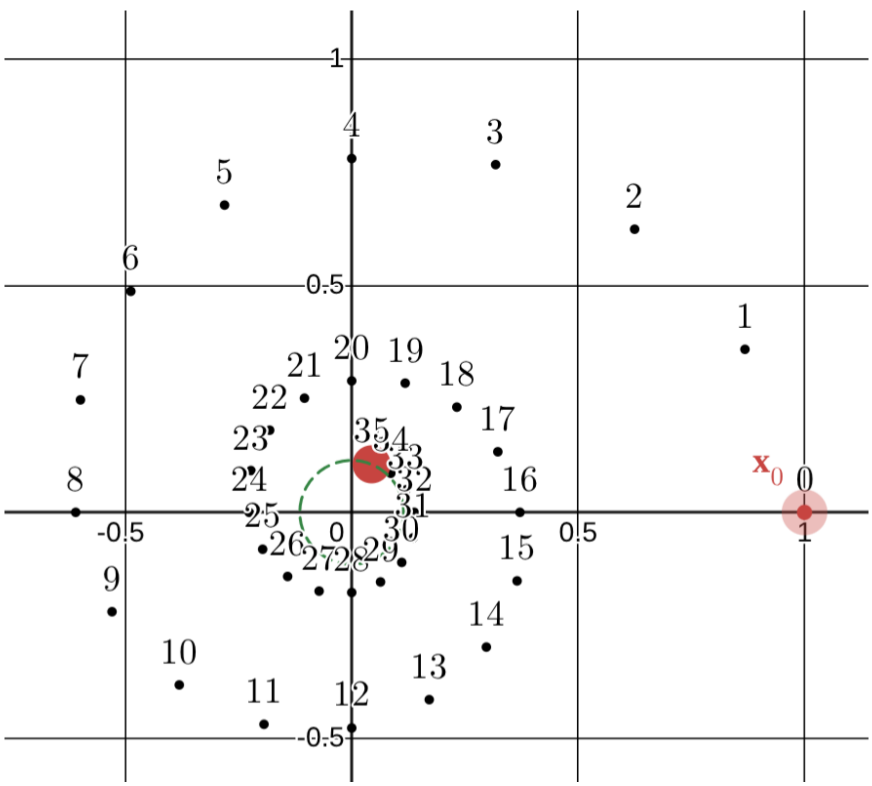
$$C_\lambda = \mathrm{S}_3\,\mathrm{R}_{\pi/3}$$

(C_λ rotates by $\pi/3$ and then scales by 3).

To understand the dynamical system $\mathbf{x} \mapsto C_\lambda \mathbf{x}$, observe that

$$C_\lambda^k = |\lambda|^k \mathrm{R}_{k\theta}.$$

For example, if $|\lambda| = 0.94$, $\theta = \pi/8$ and $\mathbf{x}_0 = (1, 0)$, then the dynamical system starts at $(1, 0)$, and each iteration rotates counter-clockwise by $\pi/8$ while moving inward toward the origin (hence the spiral effect).



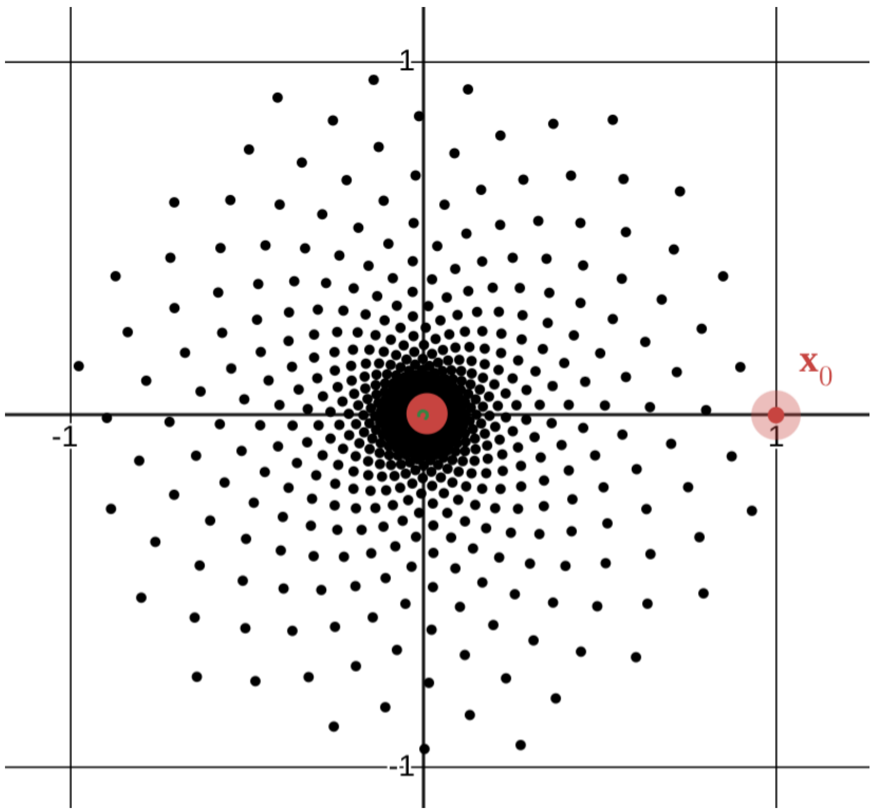
www.desmos.com/calculator/2gburodlqx

Figure 10.3: Dynamical system with matrix $\mathrm{S}_{0.94}\,\mathrm{R}_{\pi/8}$ (Desmos)

Reading Question 10F. Go to the above Desmos illustration and experiment with various values of $|\lambda|$ (s in Desmos) and θ, where θ is a rational multiple of π. In Desmos, $\theta = w\pi$ and you can experiment with w. Find specific values where the behavior of the dynamical system is qualitatively different from the example above.

RQ

Here's an example where $|\lambda| = .995$ and $\theta = 1$. The pattern is a bit more interesting because the rotation angles $k\theta$ do not repeat mod 2π. The system starts at $(1, 0)$ and spirals inward toward the origin.



www.desmos.com/calculator/n4zglocy4x

Figure 10.4: Dynamical system with matrix $S_{0.995}\,R_1$ (Desmos)

Reading Question 10G. Go to the above Desmos illustration and experiment with various values of $|\lambda|$ (s in Desmos) and θ (w in Desmos). When you use the slider to choose the angle, it will never be a rational multiple of π. Find specific values where the behavior of the dynamical system is qualitatively different from the example above.

RQ

Exercise 10D. Compute the characteristic polynomial for the 2×2 matrix R_θ that represents rotation by θ radians counter-clockwise about the origin. For which values of θ does the matrix have a real eigenvalue? When R_θ has purely complex eigenvalues, is there a basis $\mathcal{B}$ for $\mathbf{R}^2$ such that the $\mathcal{B}$-matrix for R_θ is a (real) diagonal matrix?

Exercise 10E. For each matrix A below, find a very nice matrix that it's similar to (of the form $S_{s,t}$, J_λ, or C_λ) and then describe what the similar matrix does geometrically.

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

② $\begin{bmatrix} 6 & -1 \\ 9 & 0 \end{bmatrix}$

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

In each case, find a formula for A^k.

Exercise 10F. Find a friend who has taken a differential equations course, and show them the pictures in this chapter, with an explanation. They should tell you something interesting in return.

We will close this section with one last return to the lionfish; it's a situation where we have a 3×3 matrix with one real eigenvalue and two (distinct, conjugate) purely complex eigenvalues. So though there is no basis of real eigenvectors, there *is* a basis of complex eigenvectors. The matrix is not diagonalizable if you want to work only over $\mathbf{R}$, but it is diagonalizable if you're willing to work over $\mathbf{C}$. We just want you to see that you can still get good information about the underlying real model even though complex numbers are involved.

Example 10.4 (lionfish, for the last time)**.** This is a follow-up to Example 1.2, Example 2.7, §4.2.1, and §9.4.1. This was the original matrix used to model lionfish subpopulation growth:

$$\begin{bmatrix} 0.00000 & 0.00000 & 35315. \\ 0.000030000 & 0.77700 & 0.00000 \\ 0.00000 & 0.071000 & 0.94900 \end{bmatrix}.$$

This time, we'll suppose that, to control this invasive species, adults are culled. Suppose this changes the original rate at which adults stay adults from 0.949 to 0.6. Our new model is:

$$C = \begin{bmatrix} 0.00000 & 0.00000 & 35315. \\ 0.000030000 & 0.77700 & 0.00000 \\ 0.00000 & 0.071000 & 0.60000 \end{bmatrix}.$$

This time, the (approximate) eigenvalues and eigenvectors, found with a computer, are:

$$\begin{aligned}
\lambda_1 &= 0.97942 \\
\mathbf{v}_1 &= (1.0000,\ 0.00014821,\ 0.000027734) \\
\lambda_2 &= 0.19879 - 0.19309i \\
\mathbf{v}_2 &= (1.0000,\ -0.000046679 + 0.000015588i,\ 0.00000) \\
\lambda_3 &= 0.19879 + 0.19309i \\
\mathbf{v}_3 &= (1.0000,\ -0.000046679 - 0.000015588i,\ 0.00000)
\end{aligned}$$

The three eigenvectors above (associated to the three distinct eigenvalues of the matrix C) can't form a basis for $\mathbf{R}^3$ since they don't have all real entries. However, they *do* form a basis for $\mathbf{C}^3$. The Eigenvector Basis Theorem still works over $\mathbf{C}$, and we can still conclude that, for any initial state vector

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3,$$

we have

$$\begin{aligned}
\mathbf{x}_k = A^k\mathbf{x}_0 = c_1(0.97942)^k\mathbf{v}_1 &+ c_2(0.19879 - 0.19309i)^k\mathbf{v}_2 \\
&+ c_3(0.19879 + 0.19309i)^k\mathbf{v}_3.
\end{aligned}$$

The magnitude of each complex eigenvalue is 0.27713, so we have $\lambda_i^k \to 0$ as $k \to \infty$ for each i. This means $\mathbf{x}_k$ tends to zero in the long-run.

The Spectral Theorem (2 by 2 case) §10.4

Given two vectors $\mathbf{v}$ and $\mathbf{w}$, we define their **dot product** to be the number $\mathbf{v}^T\mathbf{w}$. (We are being a teeny bit sloppy here — $\mathbf{v}^T\mathbf{w}$ is technically speaking a 1×1 matrix; not a number. But we will conflate the two and consider $\mathbf{v}^T\mathbf{w}$ a number in this course.) You can check that

What's the dot product of two 2-vectors?

$$\mathbf{v}^T\mathbf{w} = \mathbf{w}^T\mathbf{v}$$

and that

$$\mathbf{v}^T\mathbf{v} = |\mathbf{v}|^2.$$

In particular, a vector $\mathbf{v}$ has unit length if and only if $\mathbf{v}^T\mathbf{v} = 1$.

We say that the vectors $\mathbf{v}$ and $\mathbf{w}$ are **orthogonal** if

What're orthogonal vectors?

$$\mathbf{v}^T\mathbf{w} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = v_1w_1 + v_2w_2 = 0.$$

Observe that the zero vector is orthogonal to every vector.

Now, given any nonzero vector $\mathbf{v} = (v_1, v_2)$ in $\mathbf{R}^2$, the line through the origin perpendicular to $\mathbf{v}$ has equation

$$v_1 x + v_2 y = 0.$$

We see from this equation that a vector $\mathbf{w}$ is on the line through the origin perpendicular to $\mathbf{v}$ if and only if it is orthogonal to $\mathbf{v}$.

Look back at Theorem 5.5. If you revisit the proof, you'll see that we showed that if A is the matrix of a linear isometry, then its columns have unit length and are orthogonal. In fact, that theorem is a classification of exactly the 2×2 matrices whose columns are orthogonal and have unit length. Such matrices are called **orthogonal matrices**.

An orthogonal matrix has orthogonal columns **AND** *its columns have unit length.*

There is a neat way to check that a matrix is orthogonal. Let $A = [\mathbf{a}_1 \ \mathbf{a}_2]$ be a 2×2 matrix. Then

The off-diagonal entries are equal because $\mathbf{v}^{\mathrm{T}}\mathbf{w} = \mathbf{w}^{\mathrm{T}}\mathbf{v}$.

$$A^{\mathrm{T}}A = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}} \\ \mathbf{a}_2^{\mathrm{T}} \end{bmatrix} [\mathbf{a}_1 \ \mathbf{a}_2] = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}}\mathbf{a}_1 & \mathbf{a}_1^{\mathrm{T}}\mathbf{a}_2 \\ \mathbf{a}_2^{\mathrm{T}}\mathbf{a}_1 & \mathbf{a}_2^{\mathrm{T}}\mathbf{a}_2 \end{bmatrix} = \begin{bmatrix} |\mathbf{a}_1|^2 & \mathbf{a}_1^{\mathrm{T}}\mathbf{a}_2 \\ \mathbf{a}_2^{\mathrm{T}}\mathbf{a}_1 & |\mathbf{a}_2|^2 \end{bmatrix}.$$

In particular, $A^{\mathrm{T}}A = I_2$ if and only if the columns of A are orthogonal and have unit length; i.e., A is an orthogonal matrix if and only if $A^{\mathrm{T}}A = I_2$. This equation implies that $A^{-1} = A^{\mathrm{T}}$; so, it's easy to compute the inverse of an orthogonal matrix!

What's a symmetric matrix?

A matrix A is called **symmetric** if $A = A^{\mathrm{T}}$. A 2×2 symmetric matrix looks like

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

and has characteristic polynomial

$$\lambda^2 - (a + d)\lambda + ad - b^2.$$

Every symmetric matrix is diagonalizable; you'll prove this for 2×2 matrices in the next Reading Question.

RQ

Reading Question 10H. Show that a symmetric matrix is always diagonalizable because either (1) the characteristic polynomial has two distinct real roots or (2) $b = 0$ and so A is itself diagonal.

It turns out we can do better. The above exercise says that if A is symmetric, then there is an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$. In fact, we may choose P to be an orthogonal matrix. When ever this is possible, we say that the matrix is **orthogonally diagonalizable**.

What's an orthogonally diagonaizable matrix?

The Spectral Theorem (2 by 2 case)

Theorem 10.5. *A 2×2 matrix A is symmetric if and only if it is orthogonally diagonalizable.*

Proof. We leave it to you to prove that an orthogonally diagonalizable matrix is symmetric (see the next exercise).

Suppose A is a symmetric 2×2 matrix (so, $A = A^{\mathrm{T}}$). Then, as noted above, either A is diagonal (in which case you can take $P = I_2$ and $D = A$) or it has two distinct real eigenvalues λ_1 and λ_2. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be corresponding eigenvectors. Any nonzero scalar multiple of an eigenvector is also an eigenvector, so we may as well assume that the eigenvectors have unit length. Since the eigenvalues are distinct, they form a linearly independent set. We'll next prove that they are orthogonal.

Track carefully through the following computation:

$$\begin{aligned}\lambda_1 \mathbf{v}_1^{\mathrm{T}} \mathbf{v}_2 &= (\lambda_1 \mathbf{v}_1)^{\mathrm{T}} \mathbf{v}_2 \\ &= (A\mathbf{v}_1)^{\mathrm{T}} \mathbf{v}_2 \\ &= \mathbf{v}_1^{\mathrm{T}} A^{\mathrm{T}} \mathbf{v}_2 \\ &= \mathbf{v}_1^{\mathrm{T}} A \mathbf{v}_2 \\ &= \mathbf{v}_1^{\mathrm{T}} (\lambda_2 \mathbf{v}_2) \\ &= \lambda_2 \mathbf{v}_1^{\mathrm{T}} \mathbf{v}_2.\end{aligned}$$

But since $\lambda_1 \neq \lambda_2$, the only way we can have $\lambda_1 \mathbf{v}_1^{\mathrm{T}} \mathbf{v}_2 = \lambda_2 \mathbf{v}_1^{\mathrm{T}} \mathbf{v}_2$ is if $\mathbf{v}_1^{\mathrm{T}} \mathbf{v}_2 = 0$. This implies that the eigenvectors are orthogonal.

Now, according to the Eigenvector Basis Theorem, $A = PDP^{-1}$ where $P = [\mathbf{v}_1 \ \mathbf{v}_2]$ and $D = \operatorname{diag}(\lambda_1, \lambda_2)$. Above, we proved that P is an orthogonal matrix. ■

Exercise 10G. Prove that if A is orthogonally diagonalizable, then A is symmetric. [Hint: $A = PDP^{-1}$ where P is orthogonal. So what's P^{-1}? What do you have to check to verify that A is symmetric? You'll also need to use the fact that the transpose is product-reversing.]

For a symmetric matrix A, this theorem says that you can choose a basis $\mathcal{B}$ so that the $\mathcal{B}$-matrix of $\mathbf{x} \mapsto A\mathbf{x}$ is diagonal, and both the span mapping (whose matrix is the orthogonal matrix P) and the coordinate mapping (whose matrix is the orthogonal matrix P^{T}) are linear isometries. So the $\mathcal{B}$-coordinate system is a mere rotation or reflection (over some line through the origin) of the standard coordinate system.

Exercise 10H. Explain why you immediately know that

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

is orthogonally diagonalizable, and then find the relevant matrices P and D that accomplish the orthogonal diagonalization.

Exercise 10I. Suppose A is a 2×2 matrix that is orthogonally diagonalized by the orthogonal matrix $P = [\mathbf{v}_1\ \mathbf{v}_2]$ and diagonal matrix $D = \operatorname{diag}(\lambda_1, \lambda_2)$. Let $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$. For $i = 1, 2$, let $\pi_i = \mathbf{v}_i\mathbf{v}_i^{\mathrm{T}}$. Note that π_i is a 2×2 matrix.

① Show that $\pi_1(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\mathbf{v}_1$ and $\pi_2(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_2\mathbf{v}_2$. What are the $\mathcal{B}$-matrices of π_1 and π_2?

② Show that $\pi_1\pi_2 = \pi_2\pi_1 = 0$ and $\pi_i^2 = \pi_i$ for $i = 1, 2$.

③ Show that $A = \lambda_1\pi_1 + \lambda_2\pi_2$. [Hint: apply each side of this equation to $\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ and show that the results are the same.]

④ Show that $A^k = \lambda_1^k\pi_1 + \lambda_2^k\pi_2$ for all integers $k \geq 0$.

⑤ If you did Exercise 10H, then compute π_1 and π_2 for the matrix in that problem and then check the identity in ③.

---

§10.5 Similarity, generalized

We know how to construct the $\mathcal{B}$-matrix of a linear transformation from V to itself, and we used this idea repeatedly in the previous sections of this chapter. You can use exactly the same technique discussed in §8.4.2 to construct the matrix of a linear transformation between *any* two vector spaces, as long as you pick a basis for each.

Matrix of a linear transformation with respect to 2 bases

Definition 10.6. Let V be a vector space with basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$, let W be a vector space with basis $C = \{\mathbf{c}_1, \dots, \mathbf{c}_k\}$, and let $T\colon V \to W$ be a linear transformation. The **matrix of** T **with respect to the bases** $\mathcal{B}$ **and** C, or just the **$(\mathcal{B}, C)$-matrix** of T, is the $k \times n$ matrix

$$M = [\ \underbrace{[T(\mathbf{b}_1)]_C}_{\substack{\text{first}\\ \text{column}}} \ \cdots \ \underbrace{[T(\mathbf{b}_n)]_C}_{\substack{n\text{th}\\ \text{column}}}\ ].$$

M is the unique matrix that satisfies

$$M[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_C$$

for all $\mathbf{v} \in V$.

What's the $(\mathcal{B}, C)$-matrix of a linear map $V \to W$?

The $(\mathcal{B}, C)$-matrix transforms $\mathcal{B}$-coordinate vectors into C-coordinate vectors according to the rule T.

Example 10.7. Let's find the matrix representation for the derivative, as a function from P_2 to P_2, with respect to the basis $\mathcal{B} = \{1 + t, 1 - t\}$ for the domain and the standard basis $C = \{1, t\}$ for the codomain. To do this, we use the formula above, applying the derivative to each basis vector in $\mathcal{B}$ and then putting the result into C-coordinates:

$$1 + t \mapsto 1 \mapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$1 - t \mapsto -1 \mapsto \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

The matrix we seek is

$$M = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}.$$

Just to check:

$$\frac{d}{dt}(c_1(1 + t) + c_2(1 - t)) = c_1 - c_2$$

and

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} c_1 - c_2 \\ 0 \end{bmatrix}$$

is indeed the C-coordinate vector of the constant polynomial $c_1 - c_2$.

Exercise 10J. Do Example 10.7 again, but use the basis $\mathcal{B}$ for both the domain and the codomain.

Exercise 10K. Find the $\mathcal{B}$-matrix for T, where $\mathcal{B} = \{1, t, t^2\}$ and $T : \mathrm{P}_2 \to \mathrm{P}_2$ is defined by $T(a_0 + a_1 t + a_2 t^2) = 3a_0 + (5a_0 - 2a_1)t + (4a_1 + a_2)t^2$.

Exercise 10L. Figure out how to view integration as a linear transformation from P_n to P_{n+1} (Hint: when you compute an anti-derivative, you must make a choice). Compute the kernel and image of this transformation and give the matrix relative to the bases $\{1, t, \ldots, t^n\}$ and $\{1, t, \ldots, t^{n+1}\}$.

What matrix translates between two coordinate systems in the same vector space?

Example 10.8 (change of basis)**.** Suppose V is a vector space with two bases, $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$. If we take T to be the identity map ($T(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$) in Definition 10.6, then the matrix M is sometimes called a **change of basis matrix** or a **change of coordinates matrix** because, in this case,

$$M[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{C}}$$

for all $\mathbf{v} \in V$. That is, the matrix M converts $\mathcal{B}$-coordinate vectors to $\mathcal{C}$-coordinate vectors. To compute the columns of M, we must compute the $\mathcal{C}$-coordinate vector of each vector in $\mathcal{B}$.

For example, consider the following two bases for $\mathbf{R}^2$:

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

and

$$\mathcal{C} = \left\{ \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \end{bmatrix}. \right\}$$

If we as usual let $P_{\mathcal{C}}$ denote the matrix whose columns are the basis vectors in $\mathcal{C}$, then to find the columns of M we must solve $P_{\mathcal{C}}\mathbf{x} = \mathbf{b}_1$ and $P_{\mathcal{C}}\mathbf{x} = \mathbf{b}_2$. This can be done with one computation by computing the RREF of the augmented matrix

$$\begin{aligned} [P_{\mathcal{C}} \ P_{\mathcal{B}}] &= \begin{bmatrix} 2 & 1 & 1 & 1 \\ -3 & 5 & 1 & -1 \end{bmatrix} \\ &\sim \begin{bmatrix} 1 & 0 & \frac{4}{13} & \frac{6}{13} \\ 0 & 1 & \frac{5}{13} & \frac{1}{13} \end{bmatrix}. \end{aligned}$$

Thus, multiplication by the matrix

$$M = \frac{1}{13} \begin{bmatrix} 4 & 6 \\ 5 & 1 \end{bmatrix}$$

converts $\mathcal{B}$-coordinate vectors in $\mathbf{R}^2$ to C-coordinate vectors in $\mathbf{R}^2$. In this case, you may have noticed that $M = P_C^{-1} P_{\mathcal{B}}$. However, the method used above (where we row reduced the matrix $[P_C \; P_{\mathcal{B}}]$) is the more computationally efficient one, especially for larger bases.

Exercise 10M. Let's consider two different bases for P_2:

$$\mathcal{B} = \{t + t^2, t^2, 1 - t^2\}$$
$$C = \{2 + t + t^2, -1 - t^2, 3 + t^2\}.$$

Find the $(\mathcal{B}, C)$-matrix M of the identity map on P_2 (the identity map is defined by $p \mapsto p$). This matrix transforms $\mathcal{B}$-coordinate vectors into C-coordinate vectors, so it's a change-of-basis matrix for this particular pair of bases for P_2. [Hint: try to completely translate this into a question about bases in $\mathbf{R}^3$ using the *standard* coordinate mapping for P_2. There should be a way for you to cook up a 3×6 matrix (a 3×3 matrix augmented by three column vectors) whose RREF is $[I_3 \; M]$.]

Example 10.9. The set of matrices

$$\mathrm{M}_2(\mathbf{R}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \;\middle|\; a, b, c, d \in \mathbf{R} \right\}$$

is a 4 dimensional vector space; its standard basis is the set

$$\mathcal{E} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

The coordinate mapping

$$[\;]_{\mathcal{E}} : \mathrm{M}_2(\mathbf{R}) \to \mathbf{R}^4$$

is just

$$\left[\begin{bmatrix} a & b \\ c & d \end{bmatrix} \right]_{\mathcal{E}} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

Let

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

and consider the linear transformation

$$T : \mathrm{M}_2(\mathbf{R}) \to \mathrm{M}_2(\mathbf{R})$$

defined by

$$X \longmapsto AX.$$

Since

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

the $\mathcal{E}$-matrix of T is

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

Exercise 10N. The set

$$\mathcal{B}=\left\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix},\begin{bmatrix}1 & 1\\0 & 0\end{bmatrix},\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix},\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\right\}$$

is also a basis for $M_{2\times 2}$ because the corresponding $\mathcal{E}$-coordinate vectors in $\mathbf{R}^4$ are linearly independent. Let

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

and consider the linear transformation

$$T : M_2(\mathbf{R}) \longrightarrow M_2(\mathbf{R})$$

defined by

$$X \longmapsto AX.$$

We want you to compute the $\mathcal{B}$-matrix of T. We'll compute the first column. To compute the first column, we need to compute the $\mathcal{B}$-coordinate vector of the matrix

$$T\left(\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\right)=\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}=\begin{bmatrix}1 & 0\\3 & 0\end{bmatrix}.$$

We seek scalars s, t, u, v such that

$$\begin{aligned}\begin{bmatrix}1 & 0\\3 & 0\end{bmatrix} &= s\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}+t\begin{bmatrix}1 & 1\\0 & 0\end{bmatrix}+u\begin{bmatrix}1 & 1\\1 & 0\end{bmatrix}+v\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\\ &= \begin{bmatrix}s+t+u+v & t+u+v\\u+v & v\end{bmatrix}.\end{aligned}$$

We can solve this system by visual inspection; we must have $v = 0$ and $u = 3$. This forces $t = -3$ and $s = 1$. So, the first column of the $\mathcal{B}$-matrix is

$$\begin{bmatrix} 1 \\ -3 \\ 3 \\ 0 \end{bmatrix}.$$

Find the remaining 3 columns.

Chapter 11 Orthogonal and orthonormal bases

Key concepts

- The dot product and its properties
- Vector projection (onto another vector or line)
- Orthogonal and orthonormal sets
- The matrix $A^{\mathrm{T}}A$
- Orthogonal sets of nonzero vectors are linearly independent
- Linear isometries and their properties
- Isometric and orthogonal matrices
- Coordinates with respect to an orthogonal or orthonormal basis
- Orthogonal complements
- Orthogonal projection as a linear transformation
- Least squares solutions to linear systems
- Applications to data compression

Summary. Two major things happen in this chapter. First, we define orthogonal and orthonormal bases and show how to compute coordinates with respect to such bases. At this point, you know how important it is to pick a good basis to study problems using linear algebra, and you also know how tough it is to compute the coordinates of a vector with respect to a basis. When the basis is orthogonal or orthonormal, this is quite easy, and we give applications to data compression.

The other important thing we discuss is how to find least-squares solutions to linear systems. The idea is that a linear system $A\mathbf{x} = \mathbf{b}$ may be the right model for a given situation even when the system is inconsistent. We show how to replace $\mathbf{b}$ with the vector $\widehat{\mathbf{b}}$ in $\operatorname{im} A$ that is closest to the original vector $\mathbf{b}$. This new system $A\mathbf{x} = \widehat{\mathbf{b}}$ is guaranteed to be consistent, and its solutions are "best approximate solutions" to the original. A key application is finding the a best-fit line (or other curve or surface or higher-dimensional object) for a given data set.

At the end, as we say goodbye, we also suggest a few further mathematics courses that you might consider taking as a follow up to linear algebra!

Chapter 11

Let's return to something we studied very early on in this book (and quite a bit since): the solution set to the matrix-vector equation $A\mathbf{x} = \mathbf{b}$. We learned a lot about this solution set when the equation is consistent, but when the equation is inconsistent, we had nothing more to say about it. You might think that an inconsistent equation has nothing to offer, but it turns out that there are situations where an inconsistent equation is the right model and a "best approximate" solution is actually what you're looking for.

For example, suppose we've collected data in the form of a sequence of points in the plane:

$$(p_1, q_1), \dots, (p_n, q_n).$$

Maybe we plot the points and they look like this:

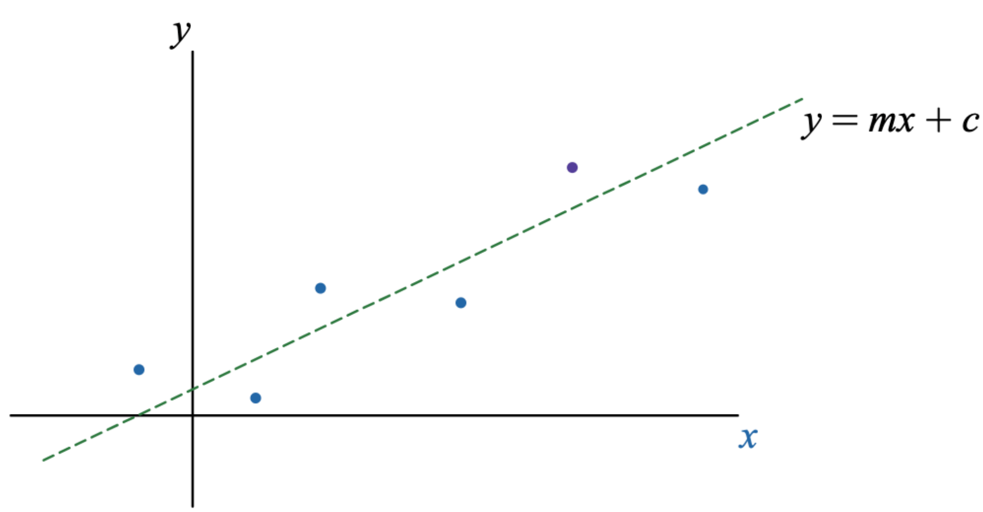


Figure 11.1: Data with a "best fit" line

The points definitely do not all lie on a single line, but it does look like they might be *close* to some single line — a "line of best fit" that describes a linear relationship between the variables x and y that holds, in some sense, on average. How might we use linear algebra to find such a line?

Consider a line with equation $y = mx + c$. It is possible, of course, that $q_i = mp_i + c$ for every point (p_i, q_i) – in other words, all of our data points lie exactly on the line (they are collinear). If this happens, the following equation holds:

$$\begin{bmatrix} c + mp_1 \\ \vdots \\ c + mp_n \end{bmatrix} = \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix};$$

$$\underbrace{\begin{bmatrix} 1 & p_1 \\ \vdots & \vdots \\ 1 & p_n \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} c \\ m \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix}}_{\mathbf{b}}.$$

So if our data points are collinear, then the solution vector (c, m) to $A\mathbf{x} = \mathbf{b}$ determines the line on which the points lie. In most situations involving real data, however, we do not expect the data points to be perfectly collinear. When the points are not collinear, the equation $A\mathbf{x} = \mathbf{b}$ is inconsistent. It turns out, however, that $A\mathbf{x} = \mathbf{b}$ has an *approximate* solution that is best possible in a sense that we can make precise. The line described by this solution won't go exactly through all the points, but it will be the line that "best fits" our data, as in Figure 11.1. One of our objectives in this chapter is to figure out how to find this approximate solution to $A\mathbf{x} = \mathbf{b}$.

How should we proceed? Our best bet is to change the inconsistent linear system just a little bit so that it *is* consistent. Since the only way to get a consistent linear system is for the augmentation vector to be in the image of the coefficient matrix, we will replace $\mathbf{b}$ with the vector in $\operatorname{im} A$ that is closest to $\mathbf{b}$. This vector is labeled $\widehat{\mathbf{b}}$ in Figure 11.2 below; it's is called the projection of $\mathbf{b}$ onto $\operatorname{im} A$.

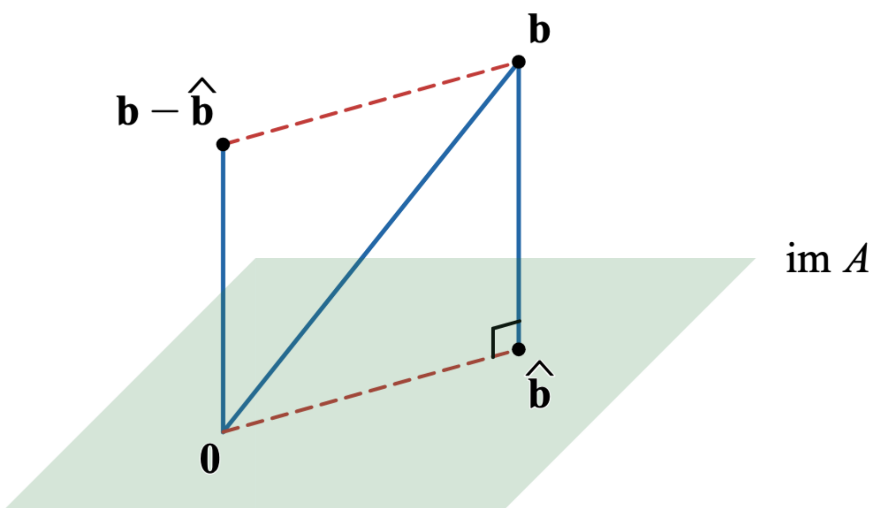


Figure 11.2: Projecting $\mathbf{b}$ onto im A

A solution to the guaranteed-to-be-consistent system $A\mathbf{x} = \widehat{\mathbf{b}}$ is called a **least-squares solution** to $A\mathbf{x} = \mathbf{b}$. We need to lay a bit of groundwork before we go any further. Let's pick up the thread from §5.4 and §10.4 where we discussed distance preserving linear transformations.

Reading Question 11A. To slightly generalize the example above, suppose A is $n \times k$. What vector space is im A a subset of? In what vector spaces do $\mathbf{b}$, $\widehat{\mathbf{b}}$, and $\mathbf{x}$ lie? RQ

More vector geometry §11.1

The dot product

Definition 11.1. Given $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{w} = (w_1, \ldots, w_n)$ in $\mathbf{R}^n$, we define the **dot product** of $\mathbf{v}$ and $\mathbf{w}$ as

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^{\mathrm{T}}\mathbf{w} = v_1 w_1 + \cdots + v_n w_n.$$

What is the dot product?

The dot product is also commonly called the *inner product*, and is sometimes denoted $\langle \mathbf{v}, \mathbf{w} \rangle$. The dot product satisfies several important properties, listed below. For any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{R}^n$ and $t \in \mathbf{R}$:

- $\mathbf{v} \cdot \mathbf{v} = |\mathbf{v}|^2$;

The dot product is commutative, associates with scalar multiplication, and distributes over addition.

- $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$;
- $(t\mathbf{v}) \cdot \mathbf{w} = \mathbf{v} \cdot (t\mathbf{w}) = t(\mathbf{v} \cdot \mathbf{w})$; and
- $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$.

The first two properties are easy to verify directly using the formula, and the last two are an immediate consequence of the fact that the dot product is defined via matrix multiplication.

RQ **Reading Question 11B.** What is $\mathbf{e}_i \cdot \mathbf{e}_j$? Consider the cases where $i = j$ and $i \neq j$ separately.

Though $\mathbf{v}$ and $\mathbf{w}$ lie in n-dimensional Euclidean space, when they're linearly independent they span a plane, which is 2-dimensional. We will therefore be able to use geometry in the plane to understand what it is that the dot product measures.

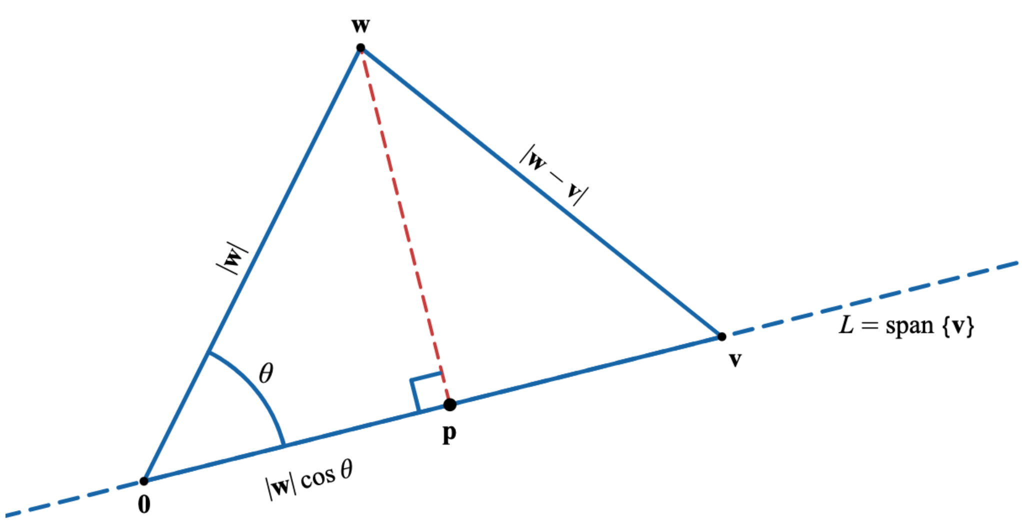


Figure 11.3: $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{p} = \text{proj}_{\mathbf{v}}\, \mathbf{w}$

In Figure 11.3, notice that the vectors $\mathbf{v}$ and $\mathbf{w}$, as the bottom and top-left sides of the triangle with vertices $\mathbf{0}, \mathbf{v}$, and $\mathbf{w}$, are perpendicular if and only if the triangle is a right triangle (with $\theta = \pi/2$). This happens if and only if the sides of the triangle obey the Pythagorean Theorem:

$$|\mathbf{v}|^2 + |\mathbf{w}|^2 = |\mathbf{v} - \mathbf{w}|^2$$

$$\begin{aligned}
&\iff \sum_{i=1}^{n} v_i^2 + \sum_{i=1}^{n} w_i^2 = \sum_{i=1}^{n} (v_i - w_i)^2 \\
&\iff \sum_{i=1}^{n} v_i^2 + \sum_{i=1}^{n} w_i^2 = \sum_{i=1}^{n} v_i^2 + \sum_{i=1}^{n} w_i^2 - 2\sum_{i=1}^{n} v_i w_i \\
&\iff 0 = \sum_{i=1}^{n} v_i w_i \\
&\iff 0 = \mathbf{v} \cdot \mathbf{w}.
\end{aligned}$$

Thus, we will say that $\mathbf{v}$ and $\mathbf{w}$ are **perpendicular** or **orthogonal**, written $\mathbf{v} \perp \mathbf{w}$, if $\mathbf{v} \cdot \mathbf{w} = 0$.

What're orthogonal or perpendicular vectors?

Exercise 11A. Let $\mathbf{v} = (1, 0, 2)$. Find a basis for the space of vectors perpendicular to $\mathbf{v}$, and describe this space geometrically. Later, we will refer to the space you just found as the **orthogonal complement** of $V = \operatorname{span}\{\mathbf{v}\}$, and we will denote it as $V^{\perp}$.

There is also a nice relationship between the dot product and the angle between $\mathbf{v}$ and $\mathbf{w}$ (labled θ in Figure 11.3, with $0 \leq \theta \leq \pi$). By the law of cosines,

$$|\mathbf{w} - \mathbf{v}|^2 = |\mathbf{v}|^2 + |\mathbf{w}|^2 - 2|\mathbf{v}||\mathbf{w}| \cos\theta.$$

Simplifying this expression using algebra (much like what was done above), we obtain

$$\mathbf{v} \cdot \mathbf{w} = |\mathbf{v}||\mathbf{w}| \cos\theta.$$

The vector labeled $\mathbf{p}$ in Figure 11.3 will play a special role in this chapter, so we'll discuss it now. What's special about $\mathbf{p}$ is that it is the point on the line $L = \operatorname{span}\{\mathbf{v}\}$ that is closest to $\mathbf{w}$. It is called the **projection of w onto** L. (We — and others — sometimes also call $\mathbf{p}$ the **projection of w onto v**; but this name is not quite as good since $\mathbf{p}$ really depends only on $\mathbf{w}$ and L, and not on the particular vector $\mathbf{v}$ that we choose to span L.) We'll denote $\mathbf{p}$ by $\operatorname{proj}_L \mathbf{w}$ or $\operatorname{proj}_{\mathbf{v}} \mathbf{w}$. Let's find a nice formula for $\mathbf{p}$.

First, note that $\mathbf{p}$ needs to be parallel to $\mathbf{v}$ and point in the same direction as $\mathbf{v}$. The *unit* length vector with this property is $\mathbf{v}/|\mathbf{v}|$. So, we just need to scale this unit vector so that it has the correct length. Using trigonometry, the length of $\mathbf{p}$ should be $|\mathbf{w}| \cos\theta$ (again: check Figure 11.3). So

$$\begin{aligned}
\operatorname{proj}_L \mathbf{w} &= \operatorname{proj}_{\mathbf{v}} \mathbf{w} \\
&= |\mathbf{w}|(\cos\theta)\frac{\mathbf{v}}{|\mathbf{v}|} \\
&= \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}|}\frac{\mathbf{v}}{|\mathbf{v}|}
\end{aligned}$$

$$= \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}|^2}\mathbf{v}$$
$$= \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v}.$$

What is the projection of a vector onto a line (or onto another vector)?

VECTOR PROJECTION

Definition 11.2. Take $\mathbf{v}, \mathbf{w} \in \mathbf{R}^n$ with $\mathbf{v}$ nonzero and let $L = \text{span}\{\mathbf{v}\}$. Define the **projection of w onto** L **(or onto v)** by

$$\text{proj}_L \mathbf{w} = \text{proj}_{\mathbf{v}} \mathbf{w} = \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{v} \cdot \mathbf{v}}\mathbf{v}.$$

The magnitude of this vector is called the **scalar projection of w onto** L **(or onto v)**:

$$|\text{proj}_L \mathbf{w}| = |\text{proj}_{\mathbf{v}} \mathbf{w}| = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}|} = |\mathbf{w}| \cos\theta,$$

where $0 \le \theta \le \pi$ is the angle between $\mathbf{v}$ and $\mathbf{w}$.

If $\mathbf{w} = \mathbf{0}$ then θ is undefined and the scalar projection is just 0.

RQ

Reading Question 11C. With notation as in Definition 11.2, what is $\text{proj}_L \mathbf{w}$ if $\mathbf{w}$ and $\mathbf{v}$ are orthogonal? What is $\text{proj}_L \mathbf{w}$ if $\mathbf{w} = s\mathbf{v}$? Work out your answers both algebraically and pictorially.

RQ

Reading Question 11D. Let $\mathbf{y} = (3, 1)$ and $\mathbf{u} = (8, 6)$. Draw a picture containing the following vectors: $\mathbf{y}$, $\mathbf{u}$, $\text{proj}_{\mathbf{u}} \mathbf{y}$. The distance from $\mathbf{y}$ to the line spanned by $\mathbf{u}$ is exactly the distance between $\mathbf{y}$ and $\text{proj}_{\mathbf{u}} \mathbf{y}$. This distance is the length of the vector $\mathbf{y} - \text{proj}_{\mathbf{u}} \mathbf{y}$. Compute this length.

What is an orthogonal set? what is an orthonormal set?

ORTHOGONAL AND ORTHONORMAL SETS

Definition 11.3. Suppose $S = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ is a set of vectors in $\mathbf{R}^k$. We call S an **orthogonal set** if $\mathbf{a}_i \cdot \mathbf{a}_j = 0$ for $i \neq j$. If S is an orthogonal set and, furthermore, each vector in S has unit length (i.e. $\mathbf{a}_i \cdot \mathbf{a}_i = 1$ for all i), then we call S an **orthonormal set**.

If S is any set of vectors, there is a neat way to compute a single matrix that contains all possible dot products of pairs of vectors from S. This matrix lets us easily detect whether S is an orthogonal or orthonormal set.

The matrix A^TA

Theorem 11.4. *Let $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$ be a $k \times n$ matrix. Then,*

$$A^{\mathrm{T}}A = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{a}_n^{\mathrm{T}} \end{bmatrix} [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$$

$$= \text{ the matrix whose } (i, j) \text{ entry is } \mathbf{a}_i^{\mathrm{T}}\mathbf{a}_j = \mathbf{a}_i \cdot \mathbf{a}_j.$$

In particular:

① *The columns of A form an orthogonal set if and only if*

$$A^{\mathrm{T}}A = \operatorname{diag}(|a_1|^2, \ldots, |a_n|^2).$$

② *The columns of A form an orthonormal set if and only if $A^{\mathrm{T}}A = I_n$.*

We saw the 2×2 version of $A^{\mathrm{T}}A$ in §10.4.

Reading Question 11E. In the notation of Theorem 11.4, what are the dimensions of the matrix $A^{\mathrm{T}}A$? RQ

Reading Question 11F. Let RQ

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

Compute $A^{\mathrm{T}}A$. What can you now say about the columns of A?

Exercise 11B. Let A be a $k \times n$ matrix such that $A^{\mathrm{T}}A$ is invertible. Prove that the columns of A are linearly independent. Be careful! You may not assume A is invertible (it may not even be a square matrix!).

We end this section with a crucial fact about sets of nonzero orthogonal vectors.

An orthogonal set of nonzero vectors is LI

Theorem 11.5. *An orthogonal set of nonzero vectors is linearly independent.*

Exercise 11C. Prove Theorem 11.5 in the following steps. Let S be an orthogonal set of nonzero vectors and let A be a matrix whose columns are the vectors in S. To prove that S is linearly independent, it suffices to prove that the zero vector is the only solution to $A\mathbf{x} = \mathbf{0}$. So let's suppose

$$A\mathbf{v} = \mathbf{0}.$$

Left-multiply the above equation by A^{T}. What does Theorem 11.4 tell you?

Exercise 11D. Find an orthonormal basis for the row space of the matrix

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

§11.2 Linear isometries

What are linear isometries, isometric matrices, and orthogonal matrices?

Linear isometry, isometric matrix, orthogonal matrix

Definition 11.6. A **linear isometry** is a linear transformation $T : \mathbf{R}^n \to \mathbf{R}^k$ that preserves distances:

$$|T(\mathbf{x}) - T(\mathbf{y})| = |\mathbf{x} - \mathbf{y}|$$

for all $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$. If U is the matrix representation of T, then we will call U an **isometric matrix.** A *square* isometric matrix is called an **orthogonal matrix**.

In §5.4, we found the linear isometries from $\mathbf{R}^2$ to $\mathbf{R}^2$. In doing so, we proved that a 2×2 matrix is orthogonal if and only if it is either R_θ or $\mathrm{R}_\theta\, \mathrm{M}_{y=0}$ (a rotation or a reflection; see Exercise 5H).

RQ

Reading Question 11G. In the notation of Definition 11.6, assume that T is a linear isometry.

- Show that $k \geq n$. Hint: if $k < n$, then what can you say about $\ker T$? Can T be distance preserving?
- Give an explicit example showing that $k = n$ is possible.
- Give an explicit example showing that $k > n$ is possible.

Isometric matrices

Theorem 11.7. *Let $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_n]$ be a $k \times n$ matrix. The following statements are equivalent.*

① *U is isometric.*

② *U preserves magnitudes: for all $\mathbf{x} \in \mathbf{R}^n$,*

$$|U\mathbf{x}| = |\mathbf{x}|.$$

③ *U preserves dot products: for all $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$,*

$$(U\mathbf{x}) \cdot (U\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}.$$

④ *The columns of U form an orthonormal set; equivalently,*

$$U^{\mathrm{T}} U = I_n.$$

Proof. First, observe that ① implies ②; just take $\mathbf{y} = \mathbf{0}$ in the definition of linear isometry. Conversely, suppose U is magnitude preserving and take $\mathbf{x}, \mathbf{y} \in \mathbf{R}^n$. Then,

$$|\mathbf{x} - \mathbf{y}| = |U(\mathbf{x} - \mathbf{y})| = |U\mathbf{x} - U\mathbf{y}|$$

and so U is an isometric matrix. We've shown that ① and ② are equivalent.

Next, let's prove that ② implies ③. Starting with the fact that U preserves the magnitude of $\mathbf{x} - \mathbf{y}$, we compute:

$$\begin{aligned} |U(\mathbf{x} - \mathbf{y})| &= |\mathbf{x} - \mathbf{y}| \\ |U(\mathbf{x} - \mathbf{y})|^2 &= |\mathbf{x} - \mathbf{y}|^2 \\ |U\mathbf{x} - U\mathbf{y}|^2 &= |\mathbf{x} - \mathbf{y}|^2 \\ (U\mathbf{x} - U\mathbf{y}) \cdot (U\mathbf{x} - U\mathbf{y}) &= (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) \\ |U\mathbf{x}|^2 - 2(U\mathbf{x} \cdot U\mathbf{y}) + |U\mathbf{y}|^2 &= |\mathbf{x}|^2 - 2\mathbf{x} \cdot \mathbf{y} + |\mathbf{y}|^2. \end{aligned}$$

Since we are assuming that U is magnitude-preserving, $|U\mathbf{x}|^2 = |\mathbf{x}|^2$ and $|U\mathbf{y}|^2 = |\mathbf{y}|^2$ and so the last line above yields

$$U\mathbf{x} \cdot U\mathbf{y} = \mathbf{x} \cdot \mathbf{y}.$$

To see that ③ implies ④:

$$\begin{aligned} \mathbf{u}_i \cdot \mathbf{u}_j &= (U\mathbf{e}_i) \cdot (U\mathbf{e}_j) \\ &= \mathbf{e}_i \cdot \mathbf{e}_j \end{aligned}$$

$$= \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

Thus $U^{\mathrm{T}}U = I_n$ (see Theorem 11.4).

To finish, we'll show (4) implies (2). Given $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbf{R}^n$, we have

Multiply this product out and use the fact that $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ when $i \neq j$ and 1 otherwise.

$$\begin{aligned} |U\mathbf{x}|^2 &= (U\mathbf{x}) \cdot (U\mathbf{x}) \\ &= (x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n) \cdot (x_1\mathbf{u}_1 + \cdots + x_n\mathbf{u}_n) \\ &= x_1^2 + \cdots + x_n^2 \\ &= \mathbf{x} \cdot \mathbf{x} \\ &= |\mathbf{x}|^2. \end{aligned}$$

From this it follows that $|U\mathbf{x}| = |\mathbf{x}|$. ■

What are orthogonal and orthonormal bases?

We are primarily interested in orthogonal matrices (the square isometric matrices). The next theorem follows from the previous one. An orthogonal set that is also a basis is called an **orthogonal basis**. Similarly, an orthonormal set that is also a basis is called an **orthonormal basis**.

ORTHOGONAL MATRICES

Theorem 11.8. *Let $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_n]$ be an $n \times n$ matrix. The following statements are equivalent.*

(1) *U is orthogonal.*
(2) *The columns of U form an orthonormal basis for $\mathbf{R}^n$.*
(3) *$U^{\mathrm{T}}U = I_n$ (i.e. U is an invertible matrix and $U^{-1} = U^{\mathrm{T}}$).*

You may find it irritating that for a matrix to be *orthogonal*, its columns have to be *orthonormal*. Matrices whose columns are merely orthogonal have no standard name.

RQ **Reading Question 11H.** Let U and V be $n \times n$ orthogonal matrices. Show that UV is an orthogonal matrix. [Hint: use Theorem 11.8 (3).]

RQ **Reading Question 11I.** What are the possible values for the determinant of an orthogonal matrix? Use the fact that the determinant is multiplicative and the determinant of a matrix is the same as the determinant of its transpose. Can you use your answer as a way to distinguish between the two cases in Theorem 5.5?

Exercise 11E. The special orthogonal matrices are the orthogonal matrices with determinant 1. Explain why I_n is a special orthogonal matrix and the set of special orthogonal matrices is closed under matrix products and matrix inversion.

Exercise 11F. Find all vectors (x, y, z) such that the columns of the matrix

$$A = \begin{bmatrix} 1 & 0 & x \\ 0 & \cos\theta & y \\ 0 & \sin\theta & z \end{bmatrix}$$

are orthogonal. Find all vectors (x, y, z) such that the columns are orthonormal (or, put another way, such that A is an orthogonal matrix).

Exercise 11G. Let

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Find an *orthonormal* basis for $\mathbf{R}^3$ consisting of eigenvectors for A. (If you let P be the matrix whose columns are these basis vectors, and if you let D be the diagonal matrix whose diagonal entries are the corresponding eigenvalues, then $A = PDP^{-1}$, and since P is an orthogonal matrix, A is orthogonally diagonalizable.)

Coordinates with respect to orthogonal bases §11.3

In this section, we will show you something truly remarkable. If there's one thing in this course that is a pain to compute (and explain, and understand!), it's the coordinate mapping. Underneath it all, finding the coordinates of a vector with respect to a basis required us, in general, to solve a linear system. There is no nice general formula for how to do this. There is, however, a nice formula when the basis is orthogonal. And if the basis is orthonormal, it gets even better!

Suppose $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthogonal basis for a subspace V of $\mathbf{R}^n$, and write $U = [\mathbf{u}_1 \ \ldots \ \mathbf{u}_k]$. Recall what the coordinate vector is:

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} \iff \mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = U[\mathbf{x}]_{\mathcal{B}}.$$

We can pick off the coordinates c_i using the dot product since $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ when $i \neq j$:

$$\begin{aligned} \mathbf{u}_i \cdot \mathbf{x} &= \mathbf{u}_i \cdot (c_1\mathbf{u}_1 + \cdots + c_i\mathbf{u}_i + \cdots + c_k\mathbf{u}_k) \\ &= c_1\mathbf{u}_i \cdot \mathbf{u}_1 + \cdots + c_i\mathbf{u}_i \cdot \mathbf{u}_i + \cdots + c_k\mathbf{u}_i \cdot \mathbf{u}_k \\ &= c_i\mathbf{u}_i \cdot \mathbf{u}_i. \end{aligned}$$

Solving for c_i, we obtain

$$c_i = \frac{\mathbf{u}_i \cdot \mathbf{x}}{\mathbf{u}_i \cdot \mathbf{u}_i}.$$

If the basis is orthonormal, the denominator in the expression above is just 1. So with a few simple dot products, you can compute the coordinate mapping!

How can we compute coordinate vectors when the basis is orthogonal or orthonormal?

Coordinates for orthogonal and orthonormal bases

Theorem 11.9. *Suppose $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthogonal basis for a subspace V of $\mathbf{R}^n$. Then for all $\mathbf{x} \in V$,*

$$[\mathbf{x}]_{\mathcal{B}} = \left(\frac{\mathbf{u}_1 \cdot \mathbf{x}}{\mathbf{u}_1 \cdot \mathbf{u}_1}, \cdots, \frac{\mathbf{u}_k \cdot \mathbf{x}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right).$$

If the basis is orthonormal, then for all $\mathbf{x} \in V$,

$$[\mathbf{x}]_{\mathcal{B}} = (\mathbf{u}_1 \cdot \mathbf{x}, \ldots, \mathbf{u}_k \cdot \mathbf{x}).$$

Further, when the basis is orthonormal, the matrix $U = [\mathbf{u}_1 \ \ldots \ \mathbf{u}_k]$ is isometric, so

$$|[\mathbf{x}]_{\mathcal{B}}| = |U[\mathbf{x}]_{\mathcal{B}}| = |\mathbf{x}|.$$

RQ

Reading Question 11J. Let $\mathbf{u}_1 = (1, 0, 1)$, $\mathbf{u}_2 = (-1, 4, 1)$, $\mathbf{u}_3 = (2, 1, -2)$, and $\mathbf{x} = (8, -4, -3)$. Verify that $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal basis for $\mathbf{R}^3$ and then compute $[\mathbf{x}]_{\mathcal{B}}$.

RQ

Reading Question 11K. Let

$$\mathbf{u}_1 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}$$

and let $U = [\mathbf{u}_1 \ \ \mathbf{u}_2]$. Compute $U^{\mathrm{T}}U$. Explain why this computation proves that the columns of U form an orthonormal basis for $\operatorname{im} U$. Find the $\mathcal{B}$-coordinates of $(20, 1, 11)$, where $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$.

Exercise 11H. We proved Theorem 11.9 by computing the dot products $\mathbf{u}_i \cdot \mathbf{x}$ directly. Here's another way: start with the equation $\mathbf{x} = U[\mathbf{x}]_{\mathcal{B}}$. Use the fact that $U^{\mathrm{T}}U = \mathrm{diag}(|\mathbf{u}_1|^2, \ldots, |\mathbf{u}_k|^2)$ to find the matrix A such that $[\mathbf{x}]_{\mathcal{B}} = A\mathbf{x}$. Now, when you compute $A\mathbf{x}$, you should get exactly the formula given in Theorem 11.9.

Subspace projections §11.4

In the introduction to this chapter we saw that, given a vector $\mathbf{b}$ and a subspace V of $\mathbf{R}^n$, we would like to be able to find the vector $\widehat{\mathbf{b}} \in V$ that is closest to $\mathbf{b}$. Let's work on that now.

Let V be a subspace of $\mathbf{R}^n$ and take $\mathbf{x} \in \mathbf{R}^n$. We'll say that $\mathbf{x}$ is perpendicular to V, written $\mathbf{x} \perp V$, if $\mathbf{x} \perp \mathbf{v}$ for all $\mathbf{v} \in V$. (Put another way, $\mathbf{x} \perp V$ means that $\mathbf{x}$ is perpendicular to all the vectors in $\mathbf{v}$.) Recall that you can check perpendicularity using the dot product: $\mathbf{x} \perp \mathbf{v} \iff \mathbf{x} \cdot \mathbf{v} = 0$. Define the **orthogonal complement** of V, written $V^\perp$, to be the set of all vectors perpendicular to V:

What's the orthogonal complement of a subspace of $\mathbf{R}^n$?

$$V^\perp = \{\mathbf{x} \in \mathbf{R}^n \mid \mathbf{x} \perp V\}.$$

If V is a plane in $\mathbf{R}^3$, then the orthogonal complement is the line through the origin that's perpendicular to V.

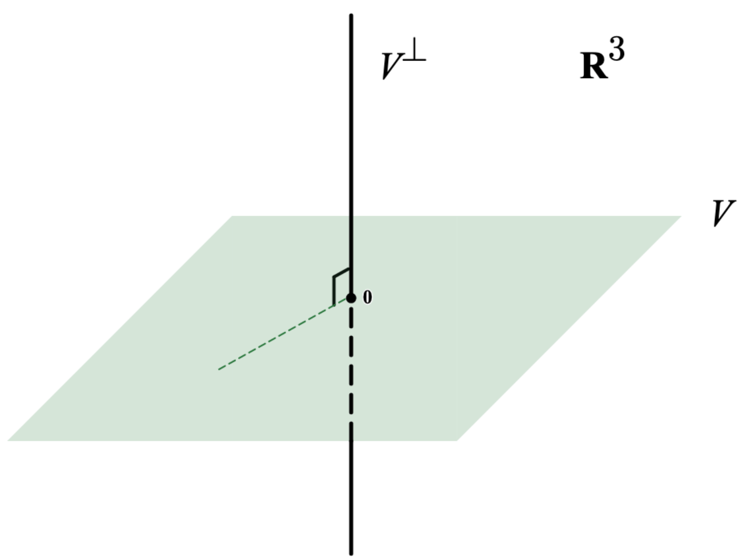


Figure 11.4: The orthogonal complement of a plane in $\mathbf{R}^3$ is a line

Reading Question 11L. If V is a line in $\mathbf{R}^3$, then what does $V^\perp$ look like? Just try to think about this one geometrically.

RQ

Exercise 11I. Prove that $V^\perp$ is indeed a subspace of $\mathbf{R}^n$.

Exercise 11J. Prove that $(V^\perp)^\perp = V$.

Exercise 11K. Show that if $\mathbf{x}$ is an element of both V and $V^\perp$, then $\mathbf{x}$ is the zero vector.

Example 11.10. Let A be an $n \times k$ matrix with rows $\mathbf{a}_1, \ldots, \mathbf{a}_n$. Take $\mathbf{x} \in \ker A$. Then,

$$\begin{aligned} A\mathbf{x} = \mathbf{0} &\iff \mathbf{a}_i \cdot \mathbf{x} = 0 \text{ for all } i \\ &\iff \mathbf{x} \perp \mathbf{a}_i \text{ for all } i \\ &\iff \mathbf{x} \perp \operatorname{span}\{\mathbf{a}_1, \ldots, \mathbf{a}_n\} = \operatorname{row} A. \end{aligned}$$

So:

$$\ker A = (\operatorname{row} A)^\perp.$$

Similarly:

$$\ker A^{\mathrm{T}} = (\operatorname{im} A)^\perp.$$

RQ

Reading Question 11M. Let

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

Find bases for $\operatorname{im} A, \ker A, \ker A^{\mathrm{T}}$, and $\operatorname{row} A$. Verify the identities $(\operatorname{row} A)^\perp = \ker A$ and $(\operatorname{im} A)^\perp = \ker A^{\mathrm{T}}$.

Exercise 11L. Prove the very last claim in Example 11.10. (You want to explain why $\mathbf{x} \in \ker A^{\mathrm{T}}$ if and only if $\mathbf{x}$ is perpendicular to the columns of A.)

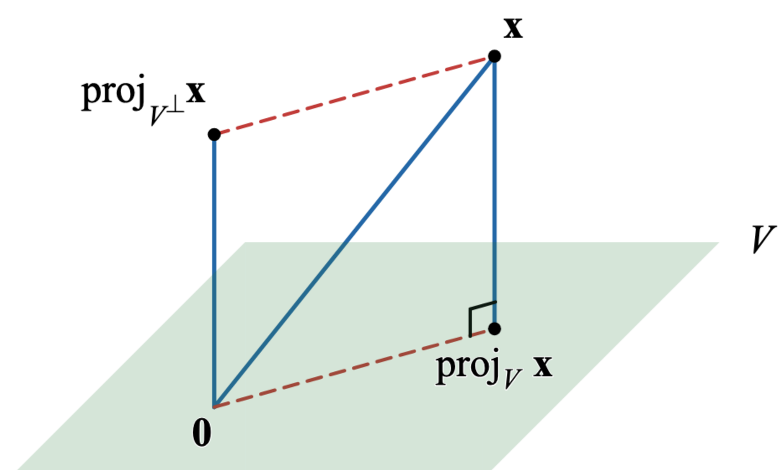
Let V be a k-dimensional subspace of $\mathbf{R}^n$. As you know, V has a basis. Given any basis for V, there is a process called the Gram-Schmidt procedure that can be used to "convert" the basis into an orthogonal basis. Further, you can use the same procedure to extend an orthogonal basis for V to an orthogonal basis for all of $\mathbf{R}^n$. We do not plan to cover this procedure in detail, but we will give you a low-dimensional example later. Take it for granted that, for any k-dimensional subspace V of $\mathbf{R}^n$, there is an orthogonal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ for $\mathbf{R}^n$ such that:

- $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ is an orthogonal basis for V, and
- the remaining vectors $\{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n\}$ form a basis for $V^\perp$.

$$\overbrace{\{\underbrace{\mathbf{u}_1, \ldots, \mathbf{u}_k}_{\text{basis for } V}, \underbrace{\mathbf{u}_{k+1}, \ldots, \mathbf{u}_n}_{\text{basis for } V^\perp}\}}^{\text{orthogonal basis for } \mathbf{R}^n}.$$

We can write any vector $\mathbf{x}$ in $\mathbf{R}^n$ as a linear combination of these basis vectors, where the weights are the $\mathcal{B}$-coordinates of $\mathbf{x}$. The projections of $\mathbf{x}$ onto V and $V^\perp$ are defined below:

$$\mathbf{x} = \underbrace{\left(\frac{\mathbf{u}_1 \cdot \mathbf{x}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\frac{\mathbf{u}_k \cdot \mathbf{x}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right)\mathbf{u}_k}_{\mathrm{proj}_V \mathbf{x}} + \underbrace{\left(\frac{\mathbf{u}_{k+1} \cdot \mathbf{x}}{\mathbf{u}_{k+1} \cdot \mathbf{u}_{k+1}}\right)\mathbf{u}_{k+1} + \cdots + \left(\frac{\mathbf{u}_n \cdot \mathbf{x}}{\mathbf{u}_n \cdot \mathbf{u}_n}\right)\mathbf{u}_n}_{\mathrm{proj}_{V^\perp} \mathbf{x}}$$



This diagram illustrates that every vector is the sum of its projections onto V and $V^\perp$.

Figure 11.5: Projection of $\mathbf{x}$ onto V and $V^\perp$

Example 11.11. Let

$$A = [\mathbf{v}\ \mathbf{w}] = \begin{bmatrix} 1 & 2 \\ -2 & 4 \\ 1 & 1 \end{bmatrix}$$

and let $V = \operatorname{im} A$. The set $\{\mathbf{v}, \mathbf{w}\}$ is a basis for V. But $\mathbf{v}$ and $\mathbf{w}$ are not perpendicular (compute their dot product to see this). We need to replace $\mathbf{w}$ with a vector perpendicular to $\mathbf{v}$ without changing the span. We can use

$$\mathbf{w}' = \mathbf{w} - \operatorname{proj}_{\mathbf{v}} \mathbf{w} = \mathbf{w} - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v}$$

because

$$\begin{aligned} \mathbf{w}' \cdot \mathbf{v} &= \left(\mathbf{w} - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v}\right) \cdot \mathbf{v} \\ &= \mathbf{w} \cdot \mathbf{v} - \mathbf{v} \cdot \mathbf{w} \\ &= 0. \end{aligned}$$

$\operatorname{span}\{\mathbf{v}, \mathbf{w}\} = \operatorname{span}\{\mathbf{v}, \mathbf{w} - \operatorname{proj}_{\mathbf{v}} \mathbf{w}\}$

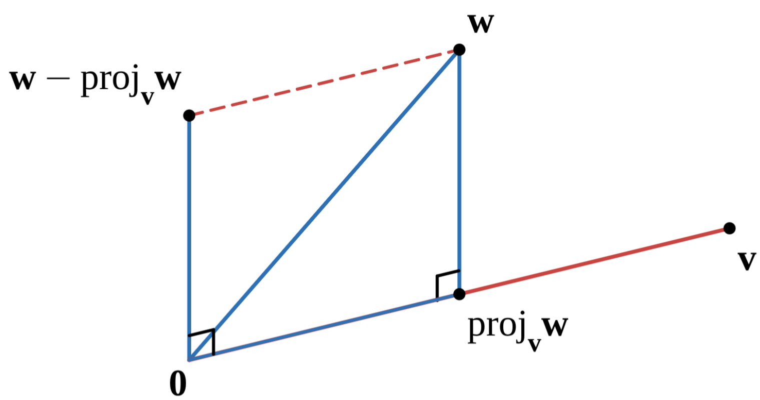


Figure 11.6: Converting $\{\mathbf{v}, \mathbf{w}\}$ to an orthogonal basis

Since $\operatorname{span}\{\mathbf{v}, \mathbf{w}\} = \operatorname{span}\{\mathbf{v}, \mathbf{w}'\}$, we have replaced our basis with an orthogonal basis. Let's compute $\mathbf{w}'$:

$$\mathbf{w}' = \mathbf{w} - \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{v} \cdot \mathbf{v}} \mathbf{v} = \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} - \frac{-5}{6} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 17/6 \\ 14/6 \\ 11/6 \end{bmatrix}.$$

We can even scale $\mathbf{w}'$ by 6 since we don't care what its length is. We now have an orthogonal basis for $V = \operatorname{im} A$:

$$\{\mathbf{u}_1, \mathbf{u}_2\} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 17 \\ 14 \\ 11 \end{bmatrix} \right\}.$$

To extend this basis for V to a basis for $\mathbf{R}^3$, we need a nonzero vector $\mathbf{u}_3$ such that $\mathbf{u}_1^{\mathrm{T}} \mathbf{u}_3 = \mathbf{u}_2^{\mathrm{T}} \mathbf{u}_3 = 0$. Thus we seek a vector in the kernel of the matrix

Think about how the row-column rule for matrix multiplication works; do you see why a vector in the kernel is perpendicular to the row vectors?

$$[\mathbf{u}_1 \ \mathbf{u}_2]^{\mathrm{T}} = \begin{bmatrix} 1 & -2 & 1 \\ 17 & 14 & 11 \end{bmatrix}.$$

You know how to find such a vector; we give one here:

$$\mathbf{u}_3 = \begin{bmatrix} -6 \\ 1 \\ 8 \end{bmatrix}.$$

Let's conclude this example by taking a vector $\mathbf{x}$ in $\mathbf{R}^3$ and computing its projections onto V and $V^\perp$. Let's use $\mathbf{x} = (2, 1, 1)$; we have

Please: check our arithmetic; compute the dot products.

$$\begin{aligned} \operatorname{proj}_V \mathbf{x} &= \frac{\mathbf{u}_1 \cdot \mathbf{x}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{u}_2 \cdot \mathbf{x}}{\mathbf{u}_2 \cdot \mathbf{u}_2}\mathbf{u}_2 \\ &= \frac{1}{6}\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \frac{59}{606}\begin{bmatrix} 17 \\ 14 \\ 11 \end{bmatrix} = \begin{bmatrix} 1104/606 \\ 624/606 \\ 750/606 \end{bmatrix}. \end{aligned}$$

You can also use the projection formula to compute the projection onto $V^\perp = \operatorname{span}\{\mathbf{u}_3\}$, but an easier way is to just use the fact that

$$\operatorname{proj}_{V^\perp} \mathbf{x} = \mathbf{x} - \operatorname{proj}_V \mathbf{x}.$$

Do this now; you should get $\operatorname{proj}_{V^\perp} \mathbf{x} = (18/101, -3/101, -24/101)$.

If you want to work with an orthonormal basis, you can just scale each vector by its length to get a unit length vector. For example, you'd replace $\mathbf{u}_1$ with $(1/\sqrt{6}, -2/\sqrt{6}, 1/\sqrt{6})$.

Let's summarize what's important about projections.

Orthogonal Projection Facts

Theorem 11.12. *Let V be a subspace of $\mathbf{R}^n$ of dimension k and let $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ be an orthogonal basis for V. Orthogonal projection onto V defines a linear transformation*

$$\operatorname{proj}_V : \mathbf{R}^n \longrightarrow V$$

which satsifies the following properties.

(1) *For every $\mathbf{x} \in \mathbf{R}^n$,*

$$\operatorname{proj}_V \mathbf{x} = \left(\frac{\mathbf{u}_1 \cdot \mathbf{x}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\frac{\mathbf{u}_k \cdot \mathbf{x}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right)\mathbf{u}_k.$$

(2) *Every $\mathbf{x} \in \mathbf{R}^n$ may be written uniquely as*

$$\mathbf{x} = \operatorname{proj}_V \mathbf{x} + \mathbf{z}$$

where $\mathbf{z} \in V^{\perp}$. (Of course, $\mathbf{z} = \operatorname{proj}_{V^\perp} \mathbf{x}$.)

(3) *If $\mathbf{x} \in V$, then $\operatorname{proj}_V \mathbf{x} = \mathbf{x}$.*

(4) *The vector $\operatorname{proj}_V \mathbf{x}$ is the closest point in V to $\mathbf{x}$ in the sense that*

$$|\mathbf{x} - \operatorname{proj}_V \mathbf{x}| < |\mathbf{x} - \mathbf{y}|$$

for all $\mathbf{y} \in V$ distinct from $\operatorname{proj}_V \mathbf{x}$.

(5) *If the basis for V is in fact orthonormal, then*

$$\operatorname{proj}_V \mathbf{x} = (\mathbf{x} \cdot \mathbf{u}_1)\mathbf{u}_1 + \cdots + (\mathbf{x} \cdot \mathbf{u}_k)\mathbf{u}_k = UU^{\mathrm{T}}\mathbf{x},$$

where $U = [\mathbf{u}_1 \cdots \mathbf{u}_k]$.

Proof. We have explained every point of the theorem except points (4) and (5). We will leave point (5) as an exercise. For point (4), take any $\mathbf{y} \in V$. Define a vector $\mathbf{d} \in V$ by

$$\mathbf{d} = \mathbf{y} - \operatorname{proj}_V \mathbf{x}.$$

Then, using item (2) above,

$$\mathbf{x} - \mathbf{y} = \mathbf{x} - \operatorname{proj}_V \mathbf{x} - \mathbf{d} = \operatorname{proj}_{V^\perp} \mathbf{x} - \mathbf{d}.$$

If we now expand the dot product $(\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y})$ and use the fact that $\mathbf{d}$ and $\operatorname{proj}_{V^\perp} \mathbf{x}$ are orthogonal (the former lies in V, the latter lies in $V^\perp$), we get

$$\begin{aligned} |\mathbf{x} - \mathbf{y}|^2 &= (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) \\ &= (\operatorname{proj}_{V^\perp} \mathbf{x} - \mathbf{d}) \cdot (\operatorname{proj}_{V^\perp} \mathbf{x} - \mathbf{d}) \\ &= |\operatorname{proj}_{V^\perp} \mathbf{x}|^2 + |\mathbf{d}|^2. \end{aligned}$$

This quantity is as small as possible when $\mathbf{d} = \mathbf{0}$ — that is, when $\mathbf{y} = \operatorname{proj}_V \mathbf{x}$. ■

Because

$$|\mathbf{x} - \operatorname{proj}_V \mathbf{x}|$$

is the smallest possible distance from $\mathbf{x}$ to any member of V, we define the **distance from x to** V to be

What's the distance from a vector to a subspace?

$$|\mathbf{x} - \operatorname{proj}_V \mathbf{x}| = |\operatorname{proj}_{V^\perp} \mathbf{x}|.$$

Exercise 11M. Prove Theorem 11.12 item (5).

Exercise 11N. To define orthogonal projection, we used an orthogonal basis for V. You might worry that if we pick a different orthogonal basis, then we'll get a different projection. But this is not the case. Prove this as follows. Suppose we have two orthogonal bases $\mathcal{B}$ and $\mathcal{B}'$ for V, and take a vector $\mathbf{x} \in \mathbf{R}^n$. Suppose $\mathbf{a}$ and $\mathbf{a}^\perp$ are the projections of $\mathbf{x}$ onto V and $V^\perp$ using the first basis, and suppose $\mathbf{b}$ and $\mathbf{b}^\perp$ are the projections using the second basis. Then,

$$\mathbf{a} + \mathbf{a}^\perp = \mathbf{x} = \mathbf{b} + \mathbf{b}^\perp.$$

Use Exercise 11K to prove that $\mathbf{a} = \mathbf{b}$ and $\mathbf{a}^\perp = \mathbf{b}^\perp$.

Exercise 11O. Let

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

and let

$$\mathbf{b} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}.$$

- Show that the columns of A are orthogonal, and compute $\widehat{\mathbf{b}} = \operatorname{proj}_{\operatorname{im} A} \mathbf{b}$.
- Use the computation you just did to find a solution to $A\mathbf{x} = \widehat{\mathbf{b}}$.

Exercise 11P. Let

$$\mathbf{y} = \begin{bmatrix} 5 \\ -9 \\ 5 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} -3 \\ -5 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}.$$

Let $W = \operatorname{span}\{\mathbf{u}_1, \mathbf{u}_2\}$.

- Show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal set.
- Compute $\operatorname{proj}_W \mathbf{y}$ and $\operatorname{proj}_{W^\perp} \mathbf{y}$.
- Compute the distance from $\mathbf{y}$ to W.

Exercise 11Q. Let

$$\mathbf{y} = \begin{bmatrix} 4 \\ 8 \\ 1 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 2/3 \\ 1/3 \\ 2/3 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2/3 \\ 2/3 \\ 1/3 \end{bmatrix}.$$

Let $U = [\mathbf{u}_1 \quad \mathbf{u}_2]$.

- Compute $U^{\mathrm{T}}U$. Explain why this computation proves that the columns of U form an orthonormal basis for $\operatorname{im} U$.
- Compute UU^{T} and use this matrix to compute $\operatorname{proj}_{\operatorname{im} U} \mathbf{y}$.

Exercise 11R. Let V be a subspace of $\mathbf{R}^n$. What's the kernel of the linear transformation proj_V?

§11.5 Least squares solutions to linear systems

Let's solve the problem we set out to solve in this chapter's introduction: finding approximate solutions to inconsistent linear systems $A\mathbf{x} = \mathbf{b}$.

If we let $\widehat{\mathbf{b}} = \operatorname{proj}_{\operatorname{im} A} \mathbf{b}$, then the linear system $A\mathbf{x} = \widehat{\mathbf{b}}$ is consistent since $\widehat{\mathbf{b}} \in \operatorname{im} A$. We know how to use row reduction to solve such a system. Moreover, since $\widehat{\mathbf{b}}$ is the *closest* vector to $\mathbf{b}$ for which our system is consistent, we regard our solution of $A\mathbf{x} = \widehat{\mathbf{b}}$ as the "best approximate solution" of our original equation $A\mathbf{x} = \mathbf{b}$. This solution is referred to as a **least-squares solution** of the original equation.

What's a least-squares solution to a linear system?

There is one hitch, however. We need to be able to compute $\widehat{\mathbf{b}}$, which requires that we have an orthogonal basis for $\operatorname{im} A$. If such a basis is not readily available, we might hope for a different strategy, and the good news is that there is one.

RQ

Reading Question 11N. Do Exercise 11O. Please make sure you understand why, when computing $\widehat{\mathbf{b}}$, you actually also found a solution to $A\mathbf{x} = \widehat{\mathbf{b}}$! In that exercise it wasn't too difficult because the columns of A are orthogonal.

Look back at Example 11.10. Since $\mathbf{b} - \widehat{\mathbf{b}} \in (\operatorname{im} A)^{\perp} = \ker A^{\mathrm{T}}$, we have $A^{\mathrm{T}}\mathbf{b} = A^{\mathrm{T}}\,\widehat{\mathbf{b}}$. So if

$$\begin{aligned} A\mathbf{x} &= \widehat{\mathbf{b}}, \text{ then} \\ A^{\mathrm{T}}A\mathbf{x} &= A^{\mathrm{T}}\,\widehat{\mathbf{b}}, \text{ and so} \\ (A^{\mathrm{T}}A)\mathbf{x} &= A^{\mathrm{T}}\mathbf{b}. \end{aligned}$$

This means that least squares solutions to $A\mathbf{x} = \mathbf{b}$ coincide with with ordinary solutions to the so-called **normal equation**

What's the normal equation?

$$(A^{\mathrm{T}}A)\mathbf{x} = A^{\mathrm{T}}\mathbf{b}.$$

This observation provides a workable way to find least squares solutions when an orthogonal basis for $\operatorname{im} A$ is not evident. You can also use the normal equation to determine whether there is a *unique* least squares solution. Since the matrix $A^{\mathrm{T}}A$ is square, there is a unique least squares solution if and only if the matrix $A^{\mathrm{T}}A$ is invertible. In this case,

$$\mathbf{x} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\mathbf{b}.$$

RQ

Reading Question 11O. Let

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

and let

$$\mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}.$$

Compute $A^{\mathrm{T}}A$ and $A^{\mathrm{T}}\mathbf{b}$. The set of least-squares solutions to $A\mathbf{x} = \mathbf{b}$ is exactly the set of solutions to $A^{\mathrm{T}}A\mathbf{x} = A^{\mathrm{T}}\mathbf{b}$. Solve this system to find a least-squares solution. Is it unique?

Example 11.13. Suppose we have the following 6 points in the plane, perhaps modeling a planet's elliptical orbit:

$$[[9.94, 3.28], [5.43, 1.10], [6.74, 6.32], [2.51, 5.61], [11.0, 5.87], [1.80, 2.28]].$$

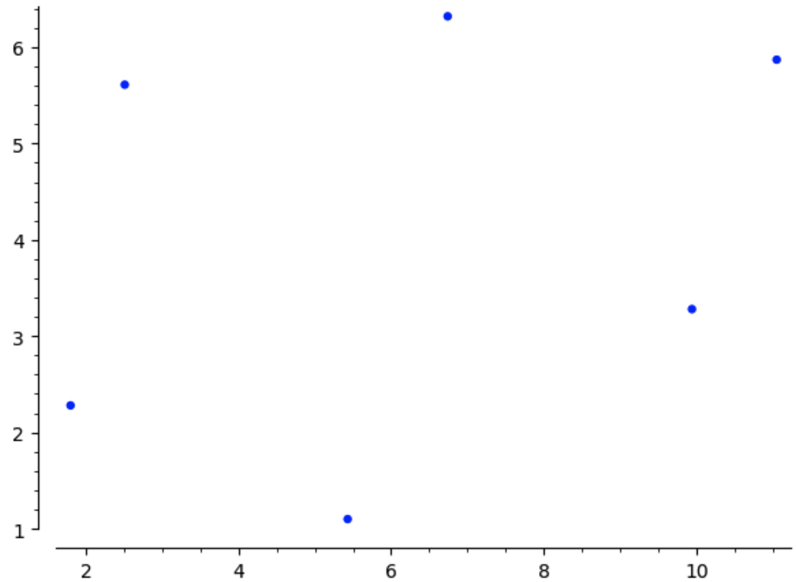


Figure 11.7: Planetary data

An ellipse is an example of a conic section (often just called a conic). Other conics are hyperbolas and parabolas. A conic has the form

$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

There is no way to choose the coefficients above so that all of our points lie on the same conic, but we can set up a least squares system to solve in order to find a conic that's the best fit. To ensure that we don't get the stupid solution where all the coefficients above are zero, we will assume $f = -1$. (This is equivalent to picking a conic that doesn't go through the origin.) We get an inconsistent linear system using our data points (x_i, y_i) as follows. Turn the list of equations

$$ax_i^2 + bx_iy_i + cy_i^2 + dx_i + ey_i = 1$$

into a matrix-vector equation:

$$\begin{bmatrix} ax_1^2 + bx_1y_1 + cy_1^2 + dx_1 + ey_1 \\ \vdots \\ ax_6^2 + bx_6y_6 + cy_6^2 + dx_6 + ey_6 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\underbrace{\begin{bmatrix} x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_6^2 & x_6y_6 & y_6^2 & x_6 & y_6 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}}_{\mathbf{b}}.$$

Plugging in our data,

$$A = \begin{bmatrix} 98.804 & 32.603 & 10.758 & 9.9400 & 3.2800 \\ 29.485 & 5.9730 & 1.2100 & 5.4300 & 1.1000 \\ 45.428 & 42.597 & 39.942 & 6.7400 & 6.3200 \\ 6.3001 & 14.081 & 31.472 & 2.5100 & 5.6100 \\ 122.10 & 64.864 & 34.457 & 11.050 & 5.8700 \\ 3.2400 & 4.1040 & 5.1984 & 1.8000 & 2.2800 \end{bmatrix}.$$

Using a computer, we found that $A^{\mathrm{T}}A$ is invertible, so the linear system $A\mathbf{x} = \mathbf{b}$ has a unique least squares solution

$$\mathbf{x} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\mathbf{b} = (-0.018257, 0.012583, -0.065470, 0.17555, 0.44727)$$

which corresponds to the conic

$$-0.018257x^2 + 0.012583xy - 0.065470y^2 + 0.17555x + 0.44727y - 1 = 0.$$

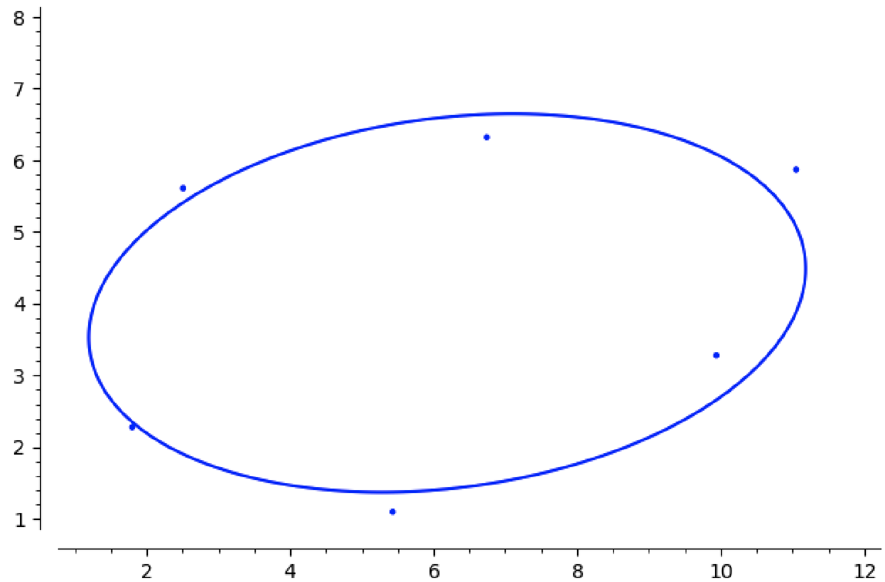


Figure 11.8: Best fit conic

As you can see, the conic we found is an ellipse and it appears to fit the data quite well!

Exercise 11S. Look back at the setup for Exercise 7B. The table gives 5 points in $\mathbf{R}^3$ of the form (x, y, r). Use least squares to find the "best fit" plane for these points and use that plane as a model to predict how a user would rate a movie with an x-value of 2 and a y-value of 3.

Why are the solutions to $A\mathbf{x} = \widehat{\mathbf{b}}$ called "least-squares" solutions to $A\mathbf{x} = \mathbf{b}$? Recall that in this situation the vector $\widehat{\mathbf{b}} = A\mathbf{x}$ is closest to $\mathbf{b}$ in the sense that $\mathbf{x}$ minimizes

$$|\widehat{\mathbf{b}} - \mathbf{b}| = |A\mathbf{x} - \mathbf{b}|.$$

This is equivalent to $\mathbf{x}$ minimizing $|A\mathbf{x} - \mathbf{b}|^2$, which is a sum of squares. Look back at the example at the very beginning of this chapter where we discussed how to find a "best fit" line to a set of points (p_i, q_i) in the plane. The solution vector $\mathbf{x} = (c, m)$ to $A\mathbf{x} = \widehat{\mathbf{b}}$ in that case gave the best fit line $y = c + mx$. If we let $r_i = c + mp_i - q_i$ (so, r_i is the difference between the y-value predicted by the model $y = c + mx$ and the actual y-value q_i), then the pair (c, m) minimizes the sum

$$r_1^2 + \cdots + r_n^2.$$

Orthonormal bases and data compression §11.6

Suppose that $\mathbf{x}$ is a vector in $\mathbf{R}^n$, that

$$\mathcal{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$$

is an orthonormal basis for $\mathbf{R}^n$, and that $U = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_n]$. We know in this case that $U^{-1} = U^T$, and accordingly that $U^T\mathbf{x}$ is the coordinate vector for $\mathbf{x}$ with respect to the basis $\mathcal{U}$:

$$[\mathbf{x}]_{\mathcal{U}} = U^T\mathbf{x} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{x} \\ \mathbf{u}_2 \cdot \mathbf{x} \\ \vdots \\ \mathbf{u}_{n-1} \cdot \mathbf{x} \\ \mathbf{u}_n \cdot \mathbf{x} \end{bmatrix}.$$

Now, whether we use standard coordinates or $\mathcal{U}$-coordinates to represent $\mathbf{x}$, we need n numbers to describe $\mathbf{x}$ completely: either the numbers x_i (in standard coordinates) or the numbers $\mathbf{u}_i \cdot \mathbf{x}$ (in $\mathcal{U}$-coordinates). So if we wish to store the vector $\mathbf{x}$ electronically or transmit it to someone else, we must store or transmit n numbers.

Imagine, though, that n is very large (so $\mathbf{x}$ is a very "long" vector). Perhaps $\mathbf{x}$ contains the price of oil every day for the past 100 years, or the color values of every pixel in a family photo, or the amplitude of sound every 20,000th of a second on your favorite dance track. In this situation we might want to *compress* $\mathbf{x}$ — that is, we might want to approximate $\mathbf{x}$ with some vector $\widehat{\mathbf{x}}$ that can be stored or transmitted more parsimoniously.

Let's stick with the same idea that we have been studying for the last couple of sections. In particular, let us write V for the span of the first k vectors in $\mathcal{U}$, and let us take as our approximation

$$\widehat{\mathbf{x}} = \operatorname{proj}_V \mathbf{x} = (\mathbf{u}_1 \cdot \mathbf{x})\mathbf{u}_1 + \cdots + (\mathbf{u}_k \cdot \mathbf{x})\mathbf{u}_k.$$

Recalling Theorem 11.12, we know that $\widehat{\mathbf{x}}$ is the closest point in V to $\mathbf{x}$. Moreover, the fact that that $\mathcal{U}$ is an *orthonormal* basis makes the error of approximation very easy to compute: since

$$\mathbf{x} - \widehat{\mathbf{x}} = (\mathbf{u}_{k+1} \cdot \mathbf{x})\mathbf{u}_{k+1} + \cdots + (\mathbf{u}_n \cdot \mathbf{x})\mathbf{u}_n,$$

we have

$$(\star) \quad |\mathbf{x} - \widehat{\mathbf{x}}| = \sqrt{(\mathbf{u}_{k+1} \cdot \mathbf{x})^2 + \cdots + (\mathbf{u}_n \cdot \mathbf{x})^2}.$$

In other words, *the entries of the coordinate vector* $[\mathbf{x}]_{\mathcal{U}}$ *let us immediately calculate how close* $\mathbf{x}$ *is to* $\widehat{\mathbf{x}}$. Geometrically, this is why we prefer to have a coordinate mapping that's a linear isometry.

Exercise 11T. Nothing like ($\star$) is true if the our basis is not orthonormal. Consider a basis of $\mathbf{R}^2$ of the form $\mathcal{B} = \{\mathbf{e}_1, \mathbf{w}\}$, write $V = \operatorname{span}\{\mathbf{e}_1\}$, and consider $\mathbf{x} = (1, 1)$. Then the closest member of V to $\mathbf{x}$ is $\mathbf{e}_1$, and if we take $\widehat{\mathbf{x}} = \mathbf{e}_1$ we have $|\mathbf{x} - \widehat{\mathbf{x}}| = 1$.

If $\mathcal{B}$ is an orthonormal basis (like the standard basis), the absolute value of the second coordinate of $[\mathbf{x}]_{\mathcal{B}}$ will be $|\mathbf{x} - \widehat{\mathbf{x}}| = 1$ — this is just what ($\star$) says in this context. However, if $\mathcal{B}$ is not orthonormal, there is no limit to how far from 1 the second coordinate of $[\mathbf{x}]_{\mathcal{B}}$ can be. To see this, work through the following examples.

(i) Take $\mathbf{w} = (\cos(\theta), \sin(\theta))$, for $0 < \theta < 1$. Graph the second coordinate of $[\mathbf{x}]_{\mathcal{B}}$ as a function of θ.

(ii) Take $\mathbf{w} = (0, s)$, for $0 < s < \infty$. Graph the second coordinate of $[\mathbf{x}]_{\mathcal{B}}$ as a function of s.

(iii) Parts (i) and (ii), taken together, illustrate that both parts of the definition of orthonormality are needed for ($\star$) to hold; explain.

The point is that we can easily see, from the coordinates $\mathbf{u}_i \cdot \mathbf{x}$, **which vectors $\mathbf{u}_i$ we can use to get good approximations of $\mathbf{x}$.** So here is a compression strategy for storing or transmitting an approximation of $\mathbf{x}$ electronically:

- Look at all the coordinates $\mathbf{u}_i \cdot \mathbf{x}$.
- Re-order the members of U so that the k vectors with the largest coordinates (in absolute value) are listed first. (We can either choose k ahead of time, or we can choose k so that $|\mathbf{x} - \widehat{\mathbf{x}}|$ is smaller than some specified value.)
- Store or transmit the k numbers $\mathbf{u}_1 \cdot \mathbf{x}, \ldots, \mathbf{u}_k \cdot \mathbf{x}$, along with a list or description of which k of the original vectors in $\mathcal{U}$ are being used to build $\widehat{\mathbf{x}}$.

To recontruct $\widehat{\mathbf{x}}$, we (or the recipients of our transmission) take the matrix U (with the basis elements appropriately re-ordered) and multiply it by the vector

$$[\widehat{\mathbf{x}}]_{\mathcal{U}} = \begin{bmatrix} \mathbf{u}_1 \cdot \mathbf{x} \\ \vdots \\ \mathbf{u}_k \cdot \mathbf{x} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Of course, all this is only worth doing if we can get a good approximation while taking k small relative to n — small enough that storing and transmitting k numbers (along with a description of which k vectors in the original basis $\mathcal{U}$ have been used) is much better than just storing the n numbers (either in standard or $\mathcal{U}$ coordinates) needed to fully describe the vector $\mathbf{x}$. Here, then, is where things get interesting. As we have already remarked in the context of diagonalization, similarity, and dynamical systems, **choice of basis is the foundation of applications of linear algebra**. It turns out that, in various application areas, orthonormal bases of $\mathbf{R}^n$ have been discovered that tend to allow us to approximate vectors $\mathbf{x}$ of interest very well, with relatively few basis vectors.

Time series compression §11.6.1

This example is artificially "small", but illustrates ideas that apply to much larger vectors.

Let us number the 8 months from November 2024 to June 2025 from 1 to 8. The figure below shows the U.S. seasonally-adjusted unemployment rate during this time (source: U.S. Bureau of Labor Statistics).

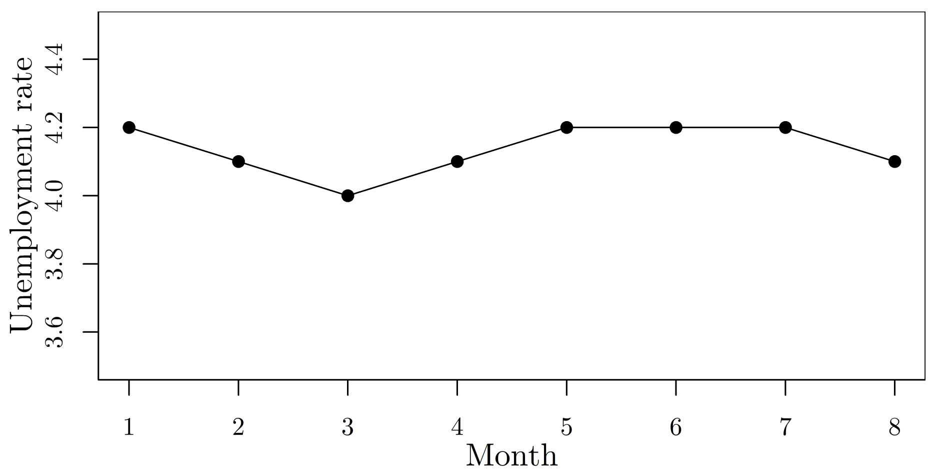


Figure 11.9: Eight months of the U.S. unemployment rate

These eight months of unemployment data can be viewed as a vector in $\mathbf{R}^8$:

$$\mathbf{x} = (4.2, 4.1, 4.0, 4.1, 4.2, 4.2, 4.2, 4.1).$$

Now, suppose we wanted to obtain an approximation $\widehat{\mathbf{x}}$ of $\mathbf{x}$ by projecting $\mathbf{x}$ onto the span of some subset of an orthonormal basis. The standard basis is very poorly suited to such a task. Even if we took the projection of $\mathbf{x}$ onto

$$\text{span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6, \mathbf{e}_7, \mathbf{e}_8\}$$

(that is, if we compressed $\mathbf{x}$ by ignoring only the *single* "least important" vector $\mathbf{e}_3$) the approximation would be lousy. The "shape" of the resulting approximation

$$\widehat{\mathbf{x}} = (4.2, 4.1, 0, 4.1, 4.2, 4.2, 4.2, 4.1)$$

is very different from $\mathbf{x}$ and $|\mathbf{x} - \widehat{\mathbf{x}}| = 4$, which (given that $|\mathbf{x}| \approx 11.70$) seems pretty big.

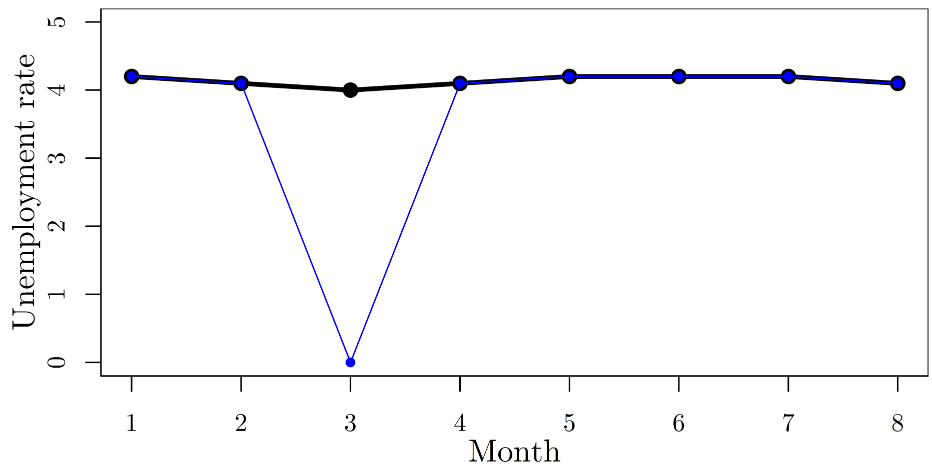


Figure 11.10: Unemployment rate (black) with a “bad” approximating vector (blue).

The reason for the deficiency of the standard basis in this situation is easy to grasp. If the vector $\mathbf{x}$ in $\mathbf{R}^n$ represents n time periods' worth of data, we expect there to be some relationships among entries of $\mathbf{x}$ that are near each other — that is, some trends in the time series. But the standard basis vectors capture no such trends.

When n is a positive power of 2 there is a particularly nice orthonormal basis called the *Hadamard basis* (a.k.a the *Walsh-Hadamard basis*, the *Hadamard-Rademacher-Walsh basis*, or the *Walsh basis*). The Hadamard basis is analogous to the Fourier basis in classical analysis; it is designed, in some sense, to detect and describe “oscillations” at various scales.

For $n = 2$, the Hadamard basis is

$$\mathcal{H}_2 = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}.$$

For $n = 4$, the Hadamard basis is

$$\mathcal{H}_4 = \left\{ \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \right\}.$$

The following figure illustrates the vectors in $\mathcal{H}_8$. Eight little horizontal axes are shown; the entry numbers of the vectors are on the horizontal axis. The entries of each vector are graphed around its horizontal axis, with the entries of each vector are joined by lines to make the "shape" of the vector easier to see. For example, the third vector in $\mathcal{H}_8$ is $(1, 1, -1, -1, 1, 1, -1, -1)/\sqrt{8}$.

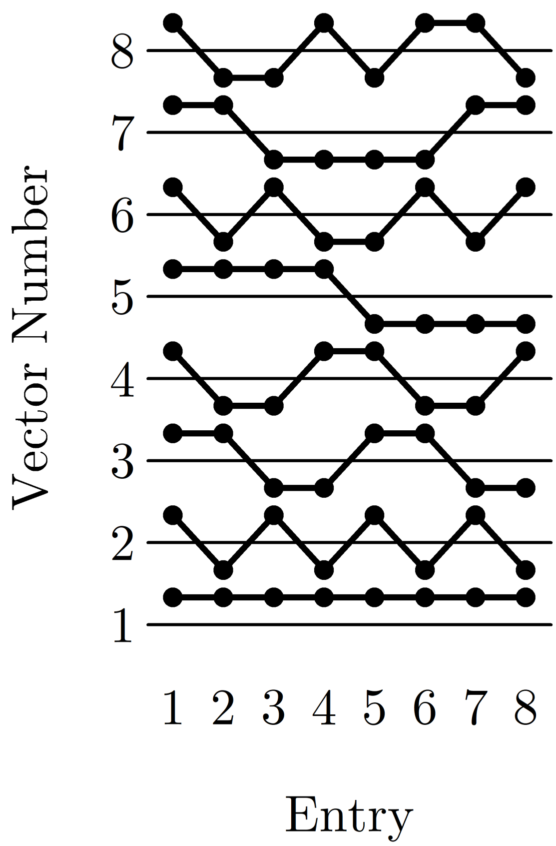


Figure 11.11: The vectors in $\mathcal{H}_8$.

Now, here is the coordinate vector of our unemployment vector **x** with respect to $\mathcal{H}_8$ (rounded to three decimal places)

$$[\mathbf{x}]_{\mathcal{H}_8} = \begin{bmatrix} 11.703 \\ 0.035 \\ 0.106 \\ 0.035 \\ -0.106 \\ -0.035 \\ 0.035 \\ 0.106 \end{bmatrix}$$

We see that coordinates 1, 3, 5, and 8 seem to be much more "important" than the others. Let us then take $\widehat{\mathbf{x}}$ to be the projection of $\mathbf{x}$ onto $V = \operatorname{span}\{\mathbf{u}_1, \mathbf{u}_3, \mathbf{u}_5, \mathbf{u}_8\}$. The resulting approximation is shown below. We have

$$\widehat{\mathbf{x}} = (4.175, 4.100, 4.025, 4.100, 4.175, 4.250, 4.175, 4.100)$$

and $|\mathbf{x} - \widehat{\mathbf{x}}| \approx 0.071$.

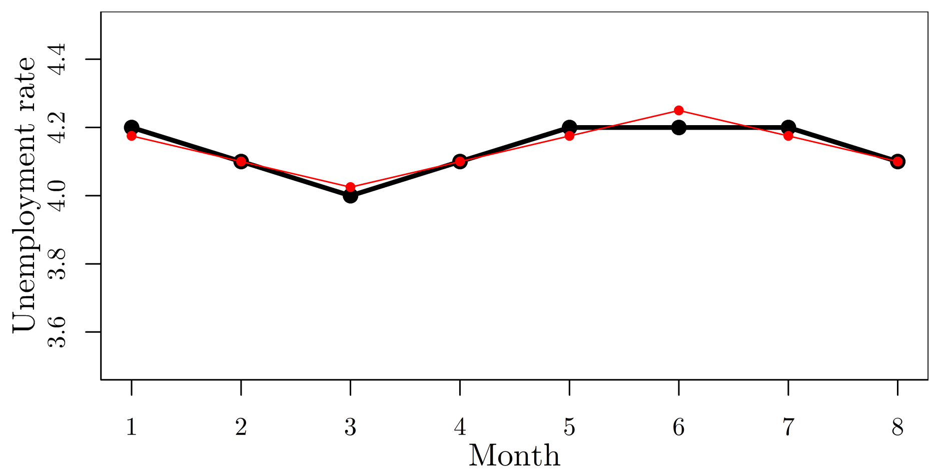


Figure 11.12: Unemployment rate (black) with a "good" approximating vector (red).

So, again, the point is that *choosing a "good" basis* let us find an accurate and parsimonious approximation to $\mathbf{x}$.

RQ **Reading Question 11P.** Verify that $\mathcal{H}_4$ really is an orthonormal basis.

Exercise 11U. Suppose that $\mathcal{U}$ is an orthonomal basis for $\mathbf{R}^n$, and that $\mathbf{u}_1$, the first vector in

$\mathcal{U}$, is a vector of the form $\mathbf{u}_1 = (a, a, \ldots, a)$. Find a formula for $\mathbf{u}_1 \cdot \mathbf{x}$, and interpret it. Relate what you find to the large first coeffient of the vector $[\mathbf{x}]_{\mathcal{H}_8}$ above.

Exercise 11V. Let's approximate the employment vector above in a slightly different way, where *you* will find the orthonormal basis. Given a vector $\mathbf{x} = (x_1, x_2, \ldots, x_8)$, let

$$\overleftarrow{\mathbf{x}} = (x_2, \ldots, x_8, x_1)$$

and let

$$\overrightarrow{\mathbf{x}} = (x_8, x_1, \ldots, x_7).$$

Define $T \colon \mathbf{R}^8 \to \mathbf{R}^8$ by

$$T(\mathbf{x}) = \frac{\overleftarrow{\mathbf{x}} + \overrightarrow{\mathbf{x}}}{2}.$$

The way to think about this function is: the input is a periodic time series, and we are averaging what happens when you shift it one unit in each direction.

1. Prove that T is a linear transformation and find its matrix representation A.
2. For each item below, you can directly use the defintion of T or you can use its matrix represenation A.
   (i) Show directly that $(1, 1, \ldots, 1)$ is an eigenvector for A with eigenvalue 1.
   (ii) Show directly that $(1, -1, 1, -1, 1, -1, 1, -1)$ is an eigenvector for A with eigenvalue -1.
   (iii) Find a pair of linearly independent (orthogonal, in fact) vectors that are eigenvectors for A with eigenvalue 0.
3. Use a computer to find an orthonormal basis of eigenvectors $\mathcal{V}$ for A. You should find that E_1 and E_{-1} each have dimesion one and E_0 has dimsension 2. There are two more real eigenvalues, and each associated eigenspace has dimension 2. (Note: if the basis $\{\mathbf{v}, \mathbf{w}\}$ your software provides for an eigenspace is not orthonormal, then normalize them first (scale them so they're unit length) and then replace them with the pair $\mathbf{v} + \mathbf{w}, \mathbf{v} - \mathbf{w}$. Do you see why this pair will now be orthogonal? You can then renormalize these new vectors to obtain an orthonormal set.)
4. Now let $\mathbf{x}$ be the employment rate vector given in this section. Compute $[\mathbf{x}]_{\mathcal{V}}$. Which four coordinates have the largest magnitudes?
5. Let V be the subspace of $\mathbf{R}^8$ spanned by the four basis vectors in $\mathcal{V}$ corresponding to the four coordinates you found in the pervious part. Compute $\widehat{\mathbf{x}}$, the projection of $\mathbf{x}$ onto V. How does the error $|\mathbf{x} - \widehat{\mathbf{x}}|$ compare to the error found using the Hadamard basis above?

The matrix A is symmetric and hence orthogonally diagonalizable by the Spectral Theorem.

Image compression §11.6.2

In this example, we give a rough idea of how .jpeg compression works. Here is a picture of Edgar sitting on the porch.

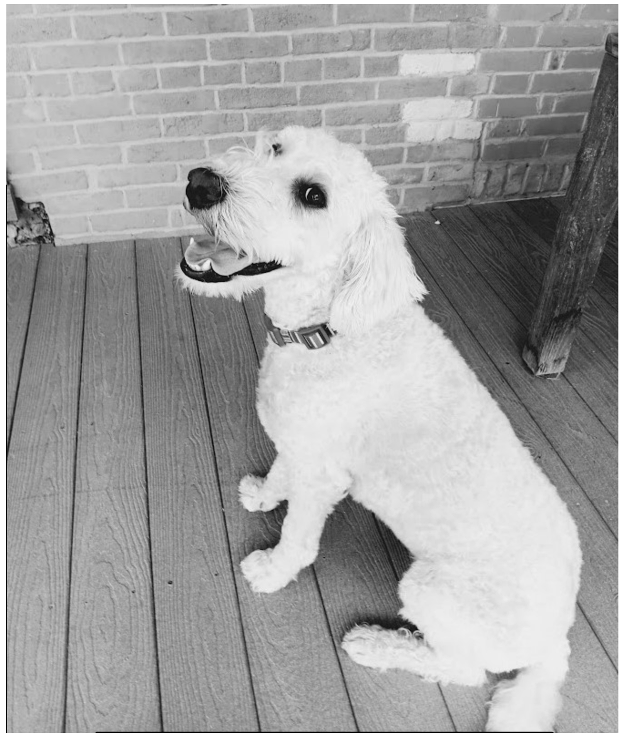


Figure 11.13: Edgar sitting on the porch.

This data for this picture are stored in a 672×800 matrix X, where $X[i, j] \in [0, 1]$ is the greyscale intensity (0 = black and 1 = white) of the pixel i places to the right and j places down from the top left corner. In other words, the picture is stored as

$$672 \times 800 = 537600$$

numbers between 0 and 1, in a specified order. Can we store a satisfactory version of the image more efficiently?

Here is the basic approach to .jpeg compression: we break the image into blocks of $b \times b$ pixels, try to come up with a more efficient version of each block, and then re-assemble these efficient versions into a putatively more efficient version of the whole image.

Figure 11.14 shows a blown-up version of the 16×16 pixel block on the top-left of Figure 11.13.

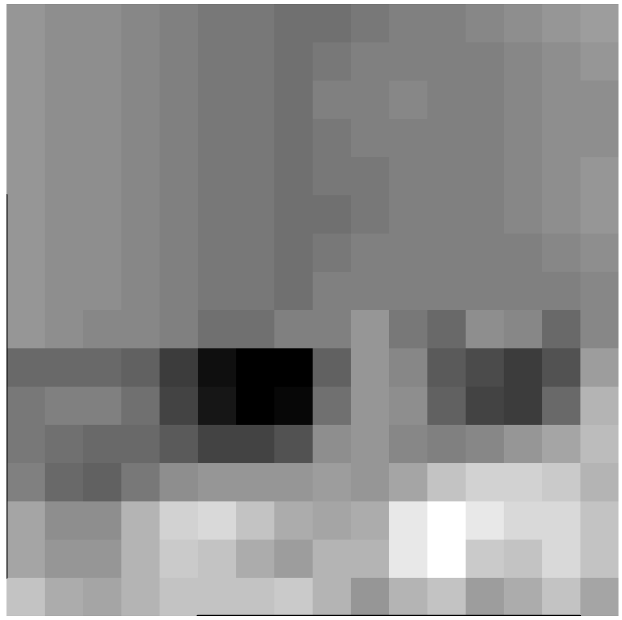


Figure 11.14: A single 16×16 pixel block.

Now, a 16×16 pixel block is pretty small. Maybe we could just take the average greyscale intensity of the whole pixel block. This would essentially amount to making each pixel 256 times larger, but would also mean that we could cut the number of different numbers we needed to store by a factor of 1/256. Let's try it!

Figure 11.15: Edgar, with constant color on each 16×16 pixel block.

This does not look very good. We will need to allow ourselves to extract more information from each pixel block. How should we do it?

By reading across successive rows from top to bottom, we can view each $b \times b$ pixel block as a vector $\mathbf{v}$ in $\mathbf{R}^{b\times b}$. It turns out that expressing $\mathbf{v}$ in terms of the standard basis for $\mathbf{R}^{b\times b}$ is very "inefficient" because, in a pixel block, we will *usually* have strong similarly between adjacent pixels and will *occasionally* have sharp contrast between adjacent pixels. This means that the vector $\mathbf{v} \in \mathbf{R}^{b\times b}$ will almost never be well-approximated by a linear combination of just a few standard basis vectors. If we choose a *different* orthonormal basis, however, then we might frequently be able to approximate $\mathbf{v}$ pretty well with a much more parsimonious linear combination of basis vectors.

The jpeg compression standard uses various versions of the so-called *discrete cosine basis* for this purpose. We will not describe this basis in detail here, though it is conceptually similar to the Hadamard basis we introduced in Example 11.6.1. For now we will content ourselves with a pictorial representation of the general idea. Let us suppose that we are dealing with 2×2 pixel blocks. Drawn as 2×2 pixel blocks, the four standard basis vectors in $\mathbf{R}^{2\times 2} = \mathbf{R}^4$ look like this.

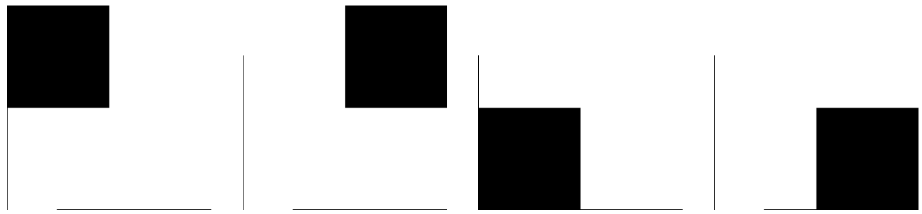


Figure 11.16: Standard basis vectors as 2×2 pixel blocks.

By contrast, the four cosine basis vectors look like this.

Figure 11.17: Discrete cosine basis vectors as 2×2 pixel blocks.

(Note: in Figure 11.16 we have reversed the usual color scheme for grayscale values. In Figure 11.17, we have normalized the vectors to have values between 0 and 1, though most of the vectors have both positive and negative values.) It is hopefully plausible that, for 2×2 pixel blocks from a "typical" image, we have a much better chance of approximating the pixel block with a linear combination of only a few basis vectors if we use the second basis. (To take a specific example: if the block is nearly monochromatic, then we can approximate it with only the first member of the second basis, but would need all four standard basis vectors.)

Let us return to Edgar. For each 16×16 block, let us do the following:

- Express the block as a vector $\mathbf{v}$ in $\mathbf{R}^{16\times16} = \mathbf{R}^{256}$;
- Calculate the coordinate vector $\mathbf{c}$ of $\mathbf{v}$ with respect to the discrete cosine basis;
- "Compress" $\mathbf{c}$ by setting to zero all entries of $\mathbf{c}$ except the 20 with the largest absolute value;
- Create a "compressed" version of $\mathbf{v}$ by assembling cosine basis vectors according to this "compressed" version of $\mathbf{c}$.

Figure 11.18 shows the result of this process, once we have re-assembled the 16×16 blocks into a single image.

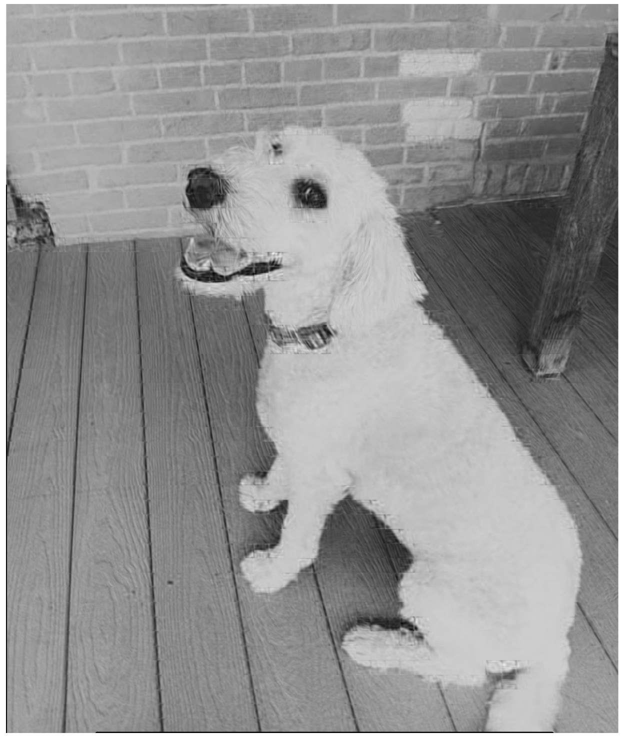


Figure 11.18: Edgar, compressed.

Not bad! To fully appreciate what we've accomplished here, we need to think about two things. The first is how parsimoniously we could store the data needed to create Figure 11.18. Well, we would need to record, for each 16×16 pixel block, (i) which discrete cosine basis vectors have non-zero coordinates and (ii) what those numbers are. That's $20 + 20 = 40$ numbers, which is less than 16 percent of the 256 numbers needed to describe the pixel block in its original form. The second is how vital a role is played by the choice of a "good" basis (in this case, the discrete cosine basis) to describe the data. Starting with Figure 11.13, we took the 20 "most important" components in each pixel block — *with respect to the discrete cosine basis* – to obtain the almost-as-good Figure 11.18. Suppose we had instead tried to take the 20 "most important" components in each pixel block with respect to the *standard basis* — this would mean just taking the 20 pixels with the darkest (or maybe the lightest) colors. Choosing "darkest," we obtain the entirely inadequate Figure 11.19.

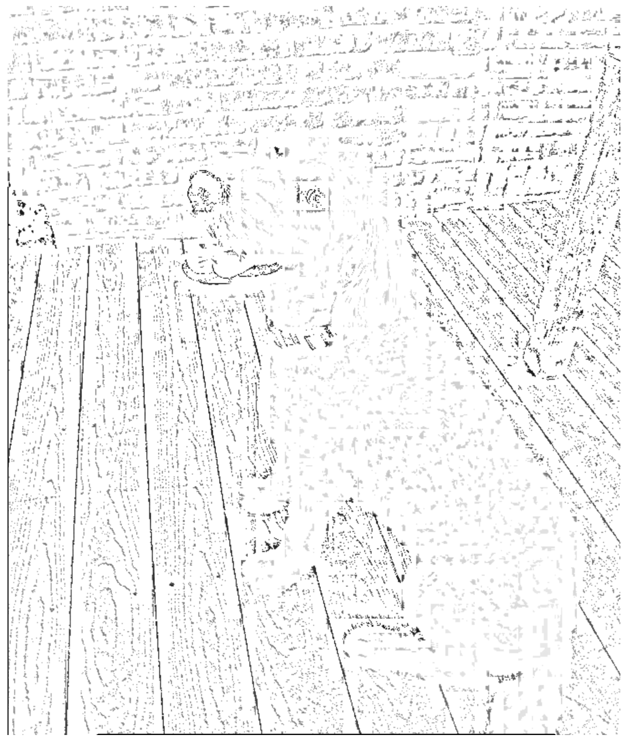


Figure 11.19: Edgar, compressed with the standard basis.

These pictures emphasize that the art of using orthonormal bases for data compression hinges on finding an appropriate orthonormal basis for the application area.

The actual jpeg compression standard works on 8×8 pixel blocks. The coordinates of each block with respect to an appropriate discrete cosine basis are calculated. Then the coordinates are discretized (with "small" coordinates being sent to zero). Finally, the information about which coordinates are used, and their discretized values, is stored in a clever and efficient way. Images are reassembled by software that takes this information and combines discrete cosine vectors appropriately.[1]

So what now? §11.7

We hope that you have enjoyed learning some of the fundamentals of linear algebra. If you have been intrigued by what you have learned, you can learn more

[1] See, for example, Gregory Wallace, "The JPEG Still Picture Compression Standard," *IEEE Transactions on Consumer Electronics* 38:1, February 1992.

about applications of linear algebra in econometrics and data science courses. In addition, the following math classes offer particularly appropriate follow-ups to this course.

- Math 215 (Abstract Mathematics): learn more about formal mathematical proof.
- Math 302 (Applied Research in Mathematics): learn more about how to use various mathematical ideas, including those from linear algebra, to study problems from business, industry and government.
- Math 331 (Abstract Algebra): learn more about algebraic structures like vector spaces (e.g. groups and rings), including their applications to geometry and combinatorics (the art of cleverly counting things).
- Math 342 (Applied Linear Algebra): learn more about what to do when matrices are *not* diagonalizable, the special structure of matrices with nonnegative entries, and more applications of projection to data analysis.
- Math 343 (Geometry): learn about linear transformations and their relatives, and how they lead us to a profound understanding of geometric ideas that have been around for millennia.
- Math 353 (Probability and Statistics): learn more about least squares and regression in probability and statistics.
- Math 361 (Chaos and Dynamical Systems): learn more about the wild and variable behavior of dynamical systems, and their applications.

We'll see you there!