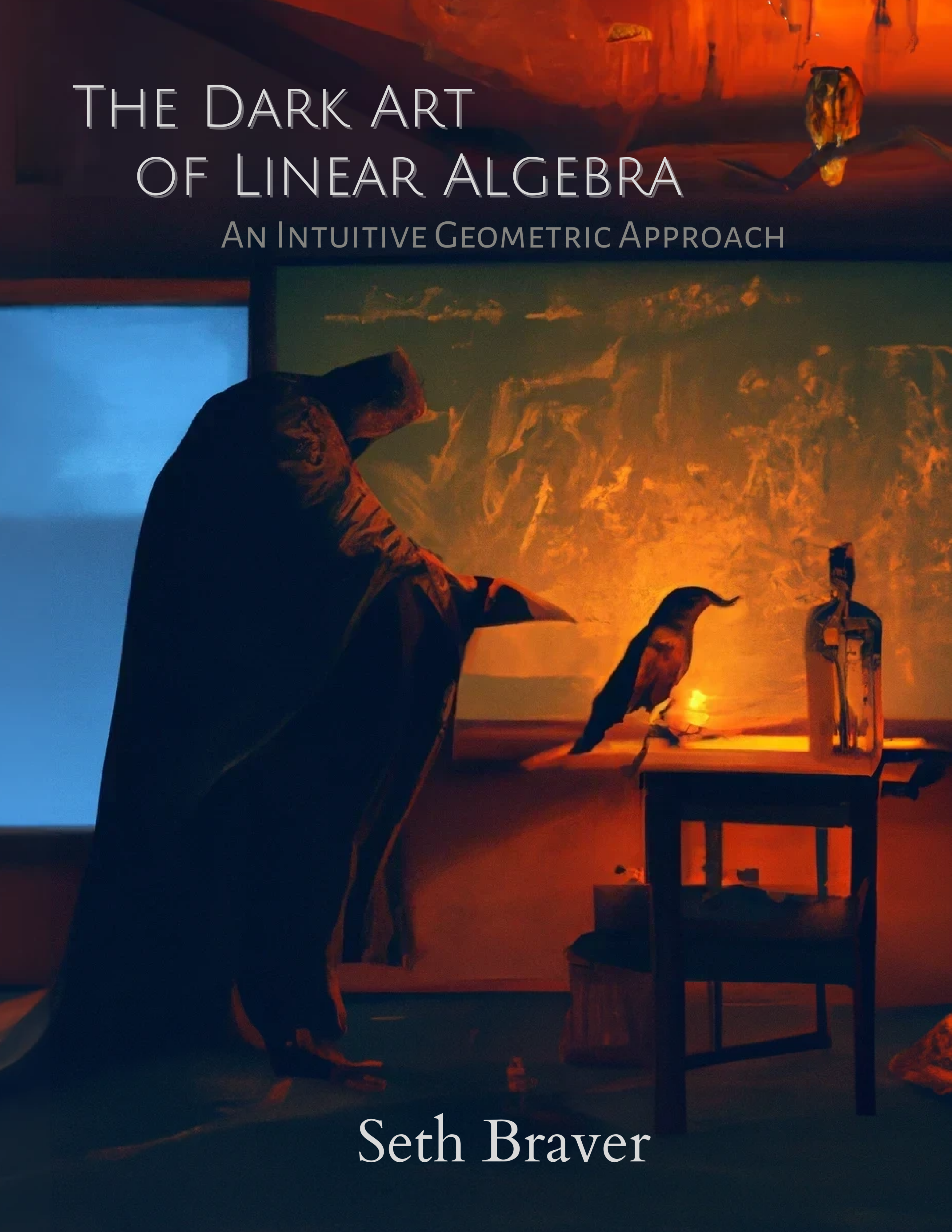
The Dark Art of Linear Algebra

An Intuitive Geometric Approach



Seth Braver

The Dark Art of Linear Algebra

An Intuitive Geometric Approach

Seth Braver

South Puget Sound Community College

VECTOR VECTORUM BOOKS

Vector Vectorum Books

Olympia, WA

Cover photo created by the author with help from Dall-E 2
(which features much linear algebra under the hood)

ISBN-13: 979-8-9881402-1-4
Library of Congress Control Number: 2023906846

This book is occasionally updated with minor corrections and changes. Most recent update: 11/5/23.

Contents

Acknowledgements

The Dark Art of Linear Algebra was my pandemic project, so I acknowledge - but do not thank - Covid for ushering it into existence. Conversely, I thank linear algebra (and the mathematicians who developed it) for providing such a diverting subject over which to brood while I hid from the virus.

My linear algebra students at South Puget Sound Community College in the Winter 2023 quarter deserve special acknowledgement for helping me test drive a draft of *The Dark Art*.* I'll single out (quadruple out?) Tobin Wheeler, Cal Holiday, Bahaa Alattar, and Alex Rice, whose questions in class motivated me to revise and substantially improve portions of the text.

I've been gratified to receive occasional emails from readers who have appreciated my previous books. Perhaps some of these friendly people are now reading this book, too. If so, thank you again. Your taste is impeccable.

Thanks to Wu Li, Empress of Cathay.
Thanks to Gulliver, the fastest basset in the West.

Thanks to my parents, who will read this page – and perhaps two more – with a certain amount of pride, but will remain none the wiser as to what linear algebra might be.

Thanks to Shannon, who has endured much, and without whom I might be hard pressed to endure little. May our candelabra continue, as the poem has it, "lighting up a night that's ours, not yet yours or mine" for many decades to come.

* In fact, some of the drafts of chapters that I gave my students to read still bore an earlier working title for the book: *Sex, Crime, and Linear Algebra (Volume 3: Linear Algebra)*.

Preface for Teachers

Of making many books there is no end;
and much study is a weariness of the flesh.
- Ecclesiastes 12:12

Linear algebra is still new. Its first devoted textbook for undergraduates, Paul Halmos's *Finite-Dimensional Vector Spaces,* was published only in 1942. As an undergraduate in the late 1990s, I was assigned a newish textbook that's now considered a classic in its own right: *Linear Algebra Done Right* by Sheldon Axler, one of Halmos's "grandstudents".* As fate would have it, I went on to become one of Axler's grandstudents. *The Dark Art of Linear Algebra* thus continues a family tradition. But the line of descent ends here. Childless and alone, I roam the manor's halls, clutching my candelabra, conscious of the ancestral portraits gazing at me from the walls.†

We know how to write calculus textbooks. We've known this for centuries, and as a result, almost all calculus texts now tend towards the same form. Naturally, there are variations in authors' expository skill, level of rigor, interest in applications, and so forth. But pick up a random calculus textbook and we know what to expect inside: not just the topics, but the order in which they will appear, and we even know which hoary old homework problems will accompany them.‡

Linear algebra is different. We don't yet know the right way to teach it to an audience of beginners, and I'm confident that neither of the two most common approaches today will stand the test of time. The traditional abstract approach, which starts from the vector space axioms, is, for all its mathematical elegance, opaque to beginners. On the other hand, the current fashion for introducing linear algebra via systems of linear equations is so mind-numbingly dull as to constitute a crime against art.

In *The Dark Art of Linear Algebra,* I take a third approach, emphasizing geometry and intuition and delaying systems of linear equations until after students have developed a strong grasp of linear maps (and the matrices that represent them) and have mastered linear algebra's core vocabulary: span, linear independence, basis, subspace, and so forth. This third way is rare in textbooks, but a series of beautiful videos develops its outlines: Grant Sanderson's "The Essence of Linear Algebra", available on his YouTube channel 3Blue1Brown. Sanderson's videos make excellent complementary material for students reading *The Dark Art of Linear Algebra*, since our approaches are so similar in spirit.

Market forces have produced generations of grotesquely bloated textbooks, but self-publication has mercifully freed me from them. I've had the luxury of writing a textbook that – miracle of miracles – students can read from cover to cover while taking a single college course. Completing a textbook is a good feeling. We should not deny our students this small pleasure.

Should you spot any typos or errors, please let me know. I can't offer you extra credit (as I do for my students), but I can offer you thanks, and I will praise your name as I ruthlessly expunge any imperfections you've identified. Should you use this book in a class, I'd especially appreciate hearing from you.

---

* Axler was Halmos's "grandstudent" in the sense that Axler's PhD advisor, Donald Sarason, was advised in *his* PhD by Halmos. Such mathematical begettings are fun to trace online at the Mathematics Genealogy Project. Karl Friedrich Gauss, it turns out, is one of my direct mathematical ancestors. Eleven intervening generations separate us.

† Forgive me, great-great-grandfather Halmos. My textbooks have illustrations, and I will sometimes sacrifice rigor for intuition. Forgive me, grandfather Axler, for I have used determinants. I honor you both. My book is meant to reach different audiences than either of yours do, and as such, it is animated by a somewhat different spirit, but I trust that it will extend our family's legacy of high quality linear algebraic exposition.

‡ My own calculus textbook, *Full Frontal Calculus*, is unusual in several respects (it is short, for instance, and uses infinitesimals), but even it adheres to the basic script – hoary old problems and all.

Preface for Students

Just remember that where there's no linear
there's no delineation. Try and stay focused.

- The Thalidomide Kid, from Cormac McCarthy's *The Passenger*.

Calculus is the summit of the high school math curriculum. Few climb it, but everyone that knows it exists. But linear algebra? How many people have even heard of it? Is it like ordinary algebra? Is it harder than calculus? Does anyone actually use it in applications? What *is* this dark art?

Geometrically speaking, linear algebra is concerned only with flatness: lines, planes, and hyperplanes. This sounds limited (and indeed it is), but those very constraints are what make the subject fundamental. Compare trigonometry. Who needs a whole subject devoted to measuring triangles? Well, everyone does. Any polygon can be chopped up into triangles, so if you understand triangles, you understand polygons. Moreover, the basic trigonometric functions take on their own life, transcend their humble origins, and become central to all periodic phenomena. Similarly, who needs linear algebra? Everyone does. There's a sense, familiar from calculus, in which non-linear phenomena can be reduced (on an infinitesimal scale) to linear phenomena, so understanding the linear world helps us grasp the nonlinear world as well. Moreover, linear algebra's basic functions, linear *transformations* (and the matrices that represent them), take on their own lives, transcend their origins, and become indispensable tools throughout mathematics, science, engineering, computer science, statistics, and even economics. In today's world – especially its technological side – linear algebra is probably the single most frequently applied part of mathematics.

You should read *The Dark Art of Linear Algebra* slowly and carefully, with your pen and paper at hand. When I omit algebraic details, you should supply them. When I use a phrase such as "as you should verify", you should do so. Only *after* reading a section should you try to solve the problems at the end. When you encounter something you don't understand, mark the relevant passage and try to clear it up – first on your own, then by discussing it with your classmates or teacher.

Conceptual understanding is especially crucial in linear algebra – much more so even than in calculus. Unlike many calculus computations, those involved in linear algebra are straightforward, even childish. They can be tedious, to be sure, but there's nothing here like the difficulty of cracking a tough integral. You need to be able to do these linear algebraic computations, of course, but that's not what learning linear algebra is about. Your main job is to understand how linear algebra's many concepts fit together, which in turn will let you understand which computations to make, and why they are appropriate.

For supplementary material, I highly recommend Grant Sanderson's series of videos "The Essence of Linear Algebra", which you can find on his YouTube channel, 3Blue1Brown. Sanderson's beautiful animations bring linear transformations to life in ways that simply aren't possible on the printed page. The general approach he takes in his videos is similar to mine in this book, although he does not discuss linear algebra's computational aspects. I also recommend trawling the internet for examples of how linear algebra is applied in your fields of interest. Examples abound, and these will give you further impetus to learn linear algebra – even if you can't at first fully understand the applications.

But enough throat clearing. Let's begin.

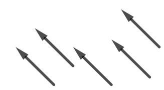
Chapter 1
Vectors

Adding Vectors

Bring me my Bow of burning gold:
Bring me my Arrows of desire

- William Blake, "And did those feet in ancient time"

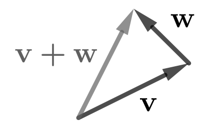
What are vectors? The answer to this innocent-sounding question will evolve as you delve ever deeper into linear algebra, but we'll begin by thinking of vectors as *arrows*: directed line segments like the one I've labelled $\mathbf{v}$ at right.* Vectors are thus geometric objects.



Vectors that point in the same direction and have the same length are said to be *equal*. Thus, the five vectors to the right of this paragraph are all equal. Or stated differently, they are all representations of the same vector, which we may *translate* anywhere we wish.

Remarkably, vectors are geometric objects that we can add. This should strike you as peculiar. Normally we add numbers (or symbols that represent numbers), not shapes.

Definition. (Vector Addition) Let $\mathbf{v}$ and $\mathbf{w}$ be vectors.
Translate them so that $\mathbf{w}$ begins where $\mathbf{v}$ ends.
We define $\mathbf{v} + \mathbf{w}$ to be the vector from $\mathbf{v}$'s tail to $\mathbf{w}$'s tip.

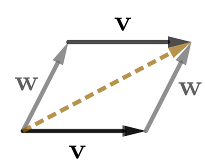


This definition is natural if we think of vectors as representing displacements. If, say, we think of a vector as corresponding to a move in a board game ("go this distance in this direction"), then vector addition corresponds to a compound move: $\mathbf{v} + \mathbf{w}$ gives the net effect of "move $\mathbf{v}$" followed by "move $\mathbf{w}$".

Vector addition obeys some familiar algebraic properties.

Claim 1. Vector addition is *commutative*: $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$.

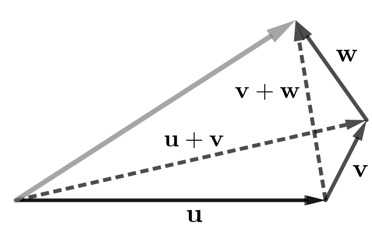
Proof. Any two vectors $\mathbf{v}$ and $\mathbf{w}$ determine a parallelogram like the one at right. The dashed vector on the diagonal splits it into two triangles. Observing each triangle in turn, we see that the dashed vector is, on the one hand, $\mathbf{v} + \mathbf{w}$, and on the other hand, $\mathbf{w} + \mathbf{v}$. Hence, $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ as claimed. ■



As a little bonus, the figure in the preceding proof reveals an alternate but equivalent way to add vectors: Translate them so that their tails coincide; their sum is a *diagonal* of the parallelogram they determine. (Specifically, it is the diagonal emanating from the point where the tails lie.)

Claim 2. Vector addition is *associative*: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.

Proof. We can arrange any three vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ tip-to-tail, as at right. Draw the (grey) vector from $\mathbf{u}$'s tail to $\mathbf{w}$'s tip. The diagonals of the resulting quadrilateral are clearly $\mathbf{u} + \mathbf{v}$ and $\mathbf{v} + \mathbf{w}$. Adding $\mathbf{w}$ to the former yields the grey vector. Adding the latter to $\mathbf{u}$ yields it, too. Equating the grey vector's two representations as vector sums reveals that $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$, as claimed. ■

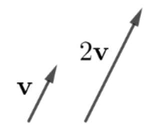


* To indicate that a symbol represents a vector, we put it in boldface: $\mathbf{v}$. When writing by hand, where boldface is impractical, we indicate that a symbol represents a vector by topping it with a little arrow like this: $\vec{v}$.

Stretching Vectors

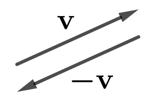
Now stretch your imagination a little…

- The Sphere, E.A. Abbott's *Flatland* (Ch. 16)



If we stretch $\mathbf{v}$ to c times its original length (while preserving its direction), the resulting vector is called $c\mathbf{v}$. This deceptively simple notation marries the geometry of stretching to the algebra of multiplication. Still, it's a funny sort of multiplication since it combines objects of two different species: it combines a vector and a number.

In linear algebra, we call numbers *scalars* because we use them to scale (i.e. to stretch or compress) vectors.* For this reason, multiplying a vector by a scalar is called **scalar multiplication**. I'll note two special cases: Scaling a vector by 1 leaves it unchanged. Scaling by 0 yields a vector without length (and hence without direction) that we call **the zero vector** and denote by $\mathbf{0}$.† Geometrically, $\mathbf{0}$ is a mere point, and it corresponds to *the absence of displacement*. Adding $\mathbf{0}$ to any other vector leaves the latter unchanged.



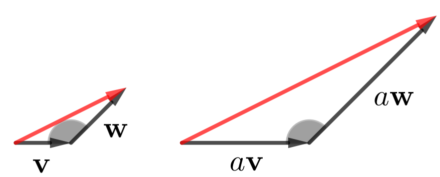
We can also multiply vectors by *negative* scalars. We define $-\mathbf{v}$ as $\mathbf{v}$ with its direction reversed. This definition ensures that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$, since the net result of adding these vectors tip-to-tail is clearly an absence of displacement. Now that we know what it means to scale a vector by -1, we can interpret scaling by any other negative. For example, multiplying by -3 can be decomposed into a sequence of two simpler operations: multiplying by -1 (which reverses direction) and multiplying by 3 (which stretches by a factor of 3).‡ Thus scaling a vector by -3 reverses its direction and triples its length.

Scalar multiplication and vector addition are linked by a pair of distributive properties.

Claim 3. For any vectors $\mathbf{v}$ and $\mathbf{w}$ and scalars a and b, the following distributive properties hold:

i) $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}.$ ii) $a(\mathbf{v} + \mathbf{w}) = a\mathbf{w} + a\mathbf{v}.$

Proof. If we think of $\mathbf{v}$ as a displacement, property (i) merely asserts the obvious. For example, it tells us that 5 "$\mathbf{v}$-steps" will get you to the same place as taking 3 "$\mathbf{v}$-steps" and then 2 "$\mathbf{v}$-steps". (Or by taking 9 "$\mathbf{v}$-steps" and then 4 "*backwards* $\mathbf{v}$-steps".)



As for property (ii), given any $\mathbf{v}, \mathbf{w}$, and a, we can form triangles such as those in the figure. The grey angles are equal (since their corresponding legs are parallel), so the triangles are similar (by SAS similarity). Hence, *all* the sides of the triangle on the right (not just the two labeled ones) must be a times the corresponding sides of the triangle on the left. By vector addition, the unlabeled sides are $\mathbf{v} + \mathbf{w}$ and $a\mathbf{w} + a\mathbf{v}$, so it follows that $a(\mathbf{v} + \mathbf{w}) = a\mathbf{w} + a\mathbf{v}$, as claimed. ■

We've now endowed vectors with a rudimentary *algebra*. Happily, vector algebra's basic rules resemble ordinary algebra's so closely that our experience with the latter lets us manipulate the symbols of vector algebra with ease. Flying vector algebra's spaceship turns out to be no harder than riding your old bicycle.

---

* Calling numbers "stretchers" would have worked just as well, but it would sound undignified, so "scalars" it is.

† Note the boldface: $\mathbf{0}$ is the zero vector; 0 is the number zero. In handwriting, the zero vector is denoted $\vec{0}$.

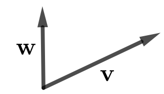
‡ We're quietly using an unexciting associative property of scalar multiplication: $(cd)\mathbf{v} = c(d\mathbf{v})$. (Here, $c = 3$ and $d = -1$.)

Exercises.

1. True or false: $0 = \mathbf{0}$.

2. (Vector subtraction)

a) In a sense, you already know how to subtract vectors, since vector subtraction is just *adding* a negated vector: $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w})$. Draw $\mathbf{v} - \mathbf{w}$ on the figure at right.



b) If it's not there already, translate the vector you drew so that it begins at $\mathbf{w}$'s tip. You've just discovered something interesting: When $\mathbf{v}$ and $\mathbf{w}$ emanate from the same point, $\mathbf{v} - \mathbf{w}$ is the vector that joins their tips and points towards $\mathbf{v}$'s tip.

c) There's another way to see the same fact that you just discovered: It's algebraically obvious that $\mathbf{v} - \mathbf{w}$ is the vector that, when added to $\mathbf{w}$, yields $\mathbf{v}$. Now look back at the figure. Which vector has that property?

3. Most people grasp the importance of commutativity, but find associativity more subtle. Let's dwell on it.

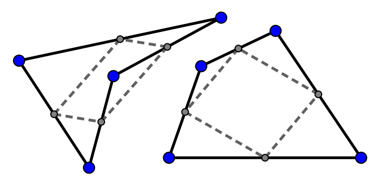
a) Ordinary subtraction is nonassociative. [e.g. $(1-5)-2 \neq 1-(5-2)$.] Think of another nonassociative operation and provide a specific counterexample.

b) Subtraction may have been a bit "too easy"; it isn't even commutative, so we already knew it was a poorly behaved operation. Can you think of an example of a *commutative* but nonassociative operation?

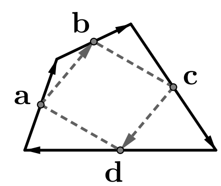
c) In Chapter 2 we'll meet an example of an associative but noncommutative operation that we'll use throughout the rest of the course: matrix multiplication. Watch for it!

4. (Order from chaos)

Pick four random points in the plane. Join them up to get a quadrilateral. We've generated this quadrilateral randomly, but hidden within it, a highly ordered object lurks: the midpoints of the random quadrilateral's sides are always the vertices of a *parallelogram*. Why? You can prove this using basic high school geometry, and you should try to do that. Interestingly, you can also prove this with vectors.



a) We can think of any quadrilateral as consisting of four vectors, as in the figure at right. We want to prove that the dashed figure within it is a parallelogram. I claim that it will be a parallelogram if and only if the sum of the figure's two dashed vectors is $\mathbf{0}$. Explain why this is so.



b) How can we prove that the sum of the two dashed vectors is $\mathbf{0}$? With some algebraic symbol shuffling! Your job is to figure out how to do this. I'll give you two hints. First, one of the dashed vectors is $(1/2)\mathbf{a} + (1/2)\mathbf{b}$. Explain why and find a similar expression for the other one. Second, what is $\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}$?

5. The properties of vector algebra that we've established justify such "obvious" statements as this:

$$(\mathbf{v} + \mathbf{w}) + (2\mathbf{v} + 3\mathbf{w}) = 3\mathbf{v} + 4\mathbf{w}.$$

This may look obvious, but there's quite a lot of algebra at work under the hood. Normally, we don't attend to such fine details, but dwelling on them at least once is good for your soul. So, let's carefully prove the preceding algebraic statement.

a) Referring to algebraic properties that we've established, justify each equals sign:

$$(\mathbf{v} + \mathbf{w}) + (2\mathbf{v} + 3\mathbf{w}) = \mathbf{v} + \big(\mathbf{w} + (2\mathbf{v} + 3\mathbf{w})\big) = \mathbf{v} + \big(\mathbf{w} + (3\mathbf{w} + 2\mathbf{v})\big) = \mathbf{v} + \big((\mathbf{w} + 3\mathbf{w}) + 2\mathbf{v}\big)$$

b) Continue the argument by justifying these equals signs:

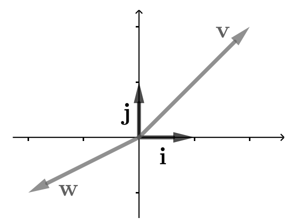
$$\mathbf{v} + \big((\mathbf{w} + 3\mathbf{w}) + 2\mathbf{v}\big) = \mathbf{v} + \big((1\mathbf{w} + 3\mathbf{w}) + 2\mathbf{v}\big) = \mathbf{v} + \big((1+3)\mathbf{w} + 2\mathbf{v}\big) = \mathbf{v} + (4\mathbf{w} + 2\mathbf{v}).$$

c) Now complete the argument. That is, carefully justify why the last expression equals $3\mathbf{v} + 4\mathbf{w}$ as claimed.

Standard Basis Vectors

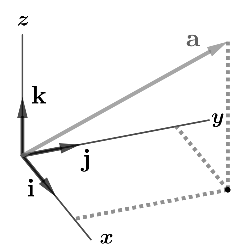
I is for IDA who drowned in a lake,
J is for JAMES who took lye by mistake.
- Edward Gorey, *The Gashleycrumb Tinies*

Readers familiar with vectors from another class may well be wondering, "Where are the coordinates?" Fear not. Your friends are safe. I've withheld them thus far only to demonstrate that vector algebra's basic properties are deep: They do not depend upon coordinates.* Coordinate systems are artificial. Geometry presents pure objects, and we impose coordinates on them. Coordinates separate us – ever so slightly – from the intrinsic nature of geometric objects, which is why mathematicians like to work "coordinate free" when possible. Still, we'd be fools to reject coordinate systems altogether. On their hardy scaffolding, we'll build our vectorial tower to heights otherwise unimaginable.



We'll begin in two-dimensional space, where the unit vectors in the positive x and y directions play special roles. We call these vectors $\mathbf{i}$ and $\mathbf{j}$, and observe that *every* vector in the plane can be represented as a weighted sum of them. (Examples: In the figure, $\mathbf{v} = 2\mathbf{i} + 2\mathbf{j}$. Similarly, $\mathbf{w} = -2\mathbf{i} - \mathbf{j}$.) For reasons that will become clear later, we call $\mathbf{i}$ and $\mathbf{j}$ the plane's **standard basis vectors**. One more term: A vector whose tail is at the origin is called a **position vector**. Clearly, the position vector pointing to (a, b) can be expressed as $a\mathbf{i} + b\mathbf{j}$. This simple observation will be our key to translating statements about *coordinates* into statements about *vectors* (and vice-versa).

We can add vectors *algebraically* by expressing them as weighted sums of the standard basis vectors. Consider vectors $\mathbf{v}$ and $\mathbf{w}$ in the figure above. What is their sum? We can certainly find it geometrically, but since we've already expressed these vectors in terms of $\mathbf{i}$ and $\mathbf{j}$, proceeding algebraically gives us a much simpler alternative: $\mathbf{v} + \mathbf{w} = (2\mathbf{i} + 2\mathbf{j}) + (-2\mathbf{i} - \mathbf{j}) = \mathbf{j}$.†



The standard basis vectors for three-dimensional space are $\mathbf{i}$, $\mathbf{j}$, and… $\mathbf{k}$, the unit vector in the positive z-direction.‡ Writing spatial vectors in terms of these three helps us add and scale them *algebraically*. For example, the figure at right shows vector $\mathbf{a} = 2\mathbf{i} + 3\mathbf{j} + 3\mathbf{k}$. If we want to scale $\mathbf{a}$ by 2 and then subtract the result from $\mathbf{b} = -8\mathbf{i} + 5\mathbf{j} + \mathbf{k}$, we can do so with a simple calculation:

$$\mathbf{b} - 2\mathbf{a} = (-8\mathbf{i} + 5\mathbf{j} + \mathbf{k}) - 2(2\mathbf{i} + 3\mathbf{j} + 3\mathbf{k}) = -12\mathbf{i} - \mathbf{j} - 5\mathbf{k}.$$

Thank God for algebra, since the vectors $\mathbf{b}$ and $(\mathbf{b} - 2\mathbf{a})$ are tricky to draw. Still, you should try to visualize them and confirm that the calculation's result makes intuitive geometric sense. (Exercise 2 should help.)

---

* Consider an analogous numerical case. Let's call a numerical property "shallow" if it isn't about the number itself, but only about its symbolic *representation*; we'll call a property "deep" if it's genuinely intrinsic to the number. Primality is deep. A mathematically sophisticated alien civilization would recognize that the number we denote 181 is prime. In contrast, the fact that this number reads the same forward as backwards is a shallow property, an accident of our base-10 positional notation. In the aliens' system of representing numbers, this number may *not* be a palindrome. Imposing coordinates on geometric objects is akin to imposing base-10 representations on numbers. You won't get far until you do it, but when you do it, you also introduce some surface-level accidents that separate us just a bit from the pure mathematical objects themselves.

† Reminder: That last equals sign's validity depends on the properties of vector algebra that we proved in this chapter's first two pages. Spelling out the details would involve the sorts of contortions we went through in exercise 5. Never again!

‡ **K** is for KATE who was struck with an axe.

Every vector in the plane can be expressed in the form $a\mathbf{i} + b\mathbf{j}$. We can obtain a convenient shorthand for such expressions by plucking off the basis vectors' weights and arranging them in a tidy column:

$$\begin{pmatrix} a \\ b \end{pmatrix} = a\mathbf{i} + b\mathbf{j}.^{*}$$

For vectors in three-dimensional space, our shorthand "column vector" notation requires *three* entries. For example, we can use this column vector notation to rewrite the sum

$$(2\mathbf{i} + 3\mathbf{j} + 5\mathbf{k}) + (4\mathbf{i} - 3\mathbf{j} + 2\mathbf{k}) = 6\mathbf{i} + 7\mathbf{k}$$

as follows:

$$\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} + \begin{pmatrix} 4 \\ -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ 7 \end{pmatrix}.$$

This new notation makes vector arithmetic easier on the eye; we add "column vectors" simply by adding the numbers in corresponding slots.†

More importantly, this column vector notation serves as a bridge from the dusty land of number lists to the rich world of geometry – often the geometry of high-dimensional spaces. For example, an economist might compare different countries' economies by compiling, for each nation, an ordered list of 100 numbers. (The first might be the country's GDP, the second might be some sort of interest rate, etc.) He can then think of each country's list as a *vector* in a 100-dimensional space. This sounds exotic, but high-dimensional spaces possess "linear backbones" whose geometry is easy to understand once you've learned to visualize their two and three-dimensional analogues. As a result, our economist can now harness his visual intuition to the algebraic techniques that you'll learn throughout this book, thereby analyzing his number lists with a little help from 100-dimensional geometry.‡

In spaces of more than three dimensions, we still have one standard basis vector for each dimension. In n-dimensional space, the n standard basis vectors are the unit vectors lying along the space's n positive axes. We usually denote these standard basis vectors by $\mathbf{e_1}$, $\mathbf{e_2}$, $\mathbf{e_3}$, …, $\mathbf{e_n}$. (The $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ alphabetic notational scheme works well for two and three dimensions but trying to extend it further is awkward.§)

We'll explore n-dimensional space a bit in the next section, where we'll turn to the problem of finding a vector's length.

---

* In some mathematical contexts, the conventional shorthand for $a\mathbf{i} + b\mathbf{j}$ is simply $\langle a, b\rangle$. This angle-bracket notation saves paper, but we'll not use it in this book. The "column vector" notation is the appropriate shorthand in linear algebra because such columns interact seamlessly with the *matrices* that lie at this subject's heart.

† Scalar multiplication is equally easy: Since $\lambda(a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) = (\lambda a)\mathbf{i} + (\lambda b)\mathbf{j} + (\lambda c)\mathbf{k}$, we multiply a column vector by a scalar by multiplying each of entry in the column vector by the scalar.

‡ Linear algebra's ability to handle lists of numbers as objects – as vectors – makes it indispensable in computer science, statistics, and applications thereof, which in turn makes it indispensable throughout science and applied mathematics. Linear algebra is equally indispensable in many areas of pure mathematics, albeit for different reasons.

§ Consider what would happen if we tried to extend that notation in a space of 30 dimensions. What would the 14th standard basis vector be called? After spending some time counting on your fingers and reciting the alphabet, you'll find that it would need to be called $\mathbf{v}$, which is bound to lead to confusion. And what would, say, the 28th standard basis be called? Who knows? Fortunately, we don't have to worry about such things, since we'll use $\mathbf{e_{28}}$, which is immediately comprehensible.

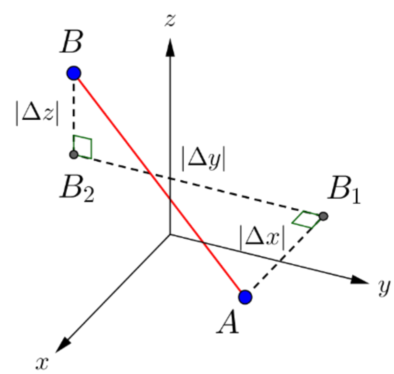
Lengths of Vectors

O Lord heal me; for my bones are sore vexed.
My soul is also sore vexed: But thou, O Lord, how long?
- Psalms 6:2-3

In this section, we'll derive a simple formula for a vector's length in n-dimensional space. But before we turn to spaces of four or more dimensions, let's recall how basic coordinate geometry works in the ordinary space of three dimensions.

In $\mathbb{R}^3$, three mutually perpendicular axes let us associate points with ordered triples of real numbers. Each point in space corresponds to an ordered triple, and vice-versa. A coordinate trio such as $(4, -2, 1)$ tells us how to reach a certain point in space from the origin: Go 4 units along the positive x-axis, then go 2 units in the direction of the negative y-axis, and finally, go 1 unit in the direction of the positive z-axis. Since three-dimensional space is equated with the set of all ordered triples of reals, we often call it $\mathbb{R}^3$.*

Among the various ways to move from point A to point B in $\mathbb{R}^3$, the scenic route – three roads, each parallel to a different axis – is the easiest to analyze, since a point travelling along it changes just one coordinate at a time. The three roads have lengths $|\Delta x|$, $|\Delta y|$, and $|\Delta z|$, and meet at right angles, so we can apply the Pythagorean Theorem twice (first to ΔAB_1B_2, then to ΔAB_2B) to find AB, thus establishing the *distance formula in* $\mathbb{R}^3$:



$$AB = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}.$$

This is, of course, the familiar distance formula in $\mathbb{R}^2$, but now with three terms instead of two under the radical.

It's tempting to guess that this pattern will hold and that the distance formula for $\mathbb{R}^n$ will be the square root of the sum of the squared changes in each of the n coordinates. This guess turns out to be correct. But rather than taking that on faith, it's worth understanding *why* this turns out to be correct. Doing so will help you learn to think about higher-dimensional spaces. Let's begin with the specific task of deriving the distance formula in $\mathbb{R}^4$.

Four-dimensional space admits *four* mutually perpendicular axes (which I'll call x, y, z, and w), which we can erect at any point. Let's imagine a two-dimensional being – a Flatlander – who lives on an ordinary two-dimensional plane (Flatland) lying within in four-dimensional space. Our Flatlander knows only his plane, and has no conception of the ambient four-dimensional space in which it lies. He might select a point in Flatland to be the origin, confidently set up x and y axes… and then have no idea where a z-axis could possibly go, declaring that the idea of a third axis perpendicular to the first two is simply inconceivable. We three-dimensional Spacelanders laugh at his parochial confusion, drawing the "inconceivable" z-axis with ease… only to find ourselves in precisely the same confusion as the Flatlander when we're asked to draw the w-axis, which should be perpendicular to each of the three existing axes.

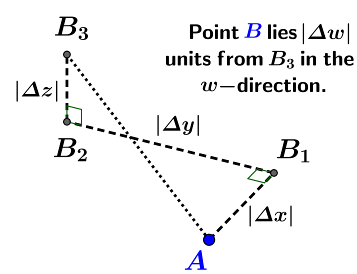
* We pronounce this "R three", not "R cubed", with the same pronunciation for analogous cases. In particular, we pronounce $\mathbb{R}^n$ like the abbreviation for a registered nurse.

We cannot draw the w-axis, but reflecting on the Flatlander's analogous difficulties will help us grope our way towards understanding. We know that the Flatlander's two-dimensional world is intersected by the "inconceivable" (to him) z-axis at just a single point, the origin. We, of course, can visualize exactly where the rest of the z-axis lies, but all the Flatlander can say is that it lies in some mysterious "elsewhere". Analogously, the w-axis intersects our three-dimensional world only at the origin, but the rest of it lies… elsewhere. Try (and fail) to picture this; you'll sympathize with the Flatlander.

Still, one thing is clear: Points in $\mathbb{R}^4$ have four coordinates. Coordinates such as $(4, -2, 1, 3)$ tell us how to reach this point from the origin: Move 4 units along the positive x-axis, then 2 in the direction of the negative y-axis, then 1 in the direction of the positive z-axis (taking us out of Flatland) and finally, we move 3 units in the direction of the positive w-axis (taking us out of Spaceland). Clearly, a point lies in Spaceland if and only if its w-coordinate is zero. Consequently, the equation of Spaceland itself, relative to the four-dimensional space we are describing, is $w = 0$.

Flatland is just one of many planes in Spaceland, which in turn is just one of many *hyperplanes* in four-dimensional space. Each plane in space extends infinitely in two mutually perpendicular dimensions, while having no thickness at all in the dimension that is perpendicular to the first two. Similarly, each hyperplane in four-dimensional space extends infinitely in three mutually perpendicular dimensions, while having no thickness at all in the dimension that is perpendicular to the first three. In three-dimensional space, the xy-plane's equation is $z = 0$, since a point in space lies on the xy-plane if and only if its z-coordinate is 0. The plane parallel to the xy-plane but lying 3 units above it has the equation $z = 3$. Similarly, in the context of four-dimensional geometry, the equations $w = 0$ and $w = 5$ correspond to parallel hyperplanes, separated by a distance of 5 units.

With these mental warmup exercises out of the way, we should now be able to derive the distance formula in $\mathbb{R}^4$ without too much trouble. Just as we did when we derived the three-dimensional distance formula, we will move from point A to point B by taking the "scenic route", but now through *four*-dimensional space along *four* roads, each parallel to a different axis. On the scenic route, just one coordinate changes per road, so their lengths are $|\Delta x|$, $|\Delta y|$, $|\Delta z|$, and $|\Delta w|$. Alas, we wretched three-dimensional creatures can't draw the journey's final leg from B_3 to B in the same picture as the first three legs (shown at right), but we can still reason about it.



Applying the Pythagorean Theorem to ΔAB_1B_2 and ΔAB_2B_3 yields $AB_3 = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$. To link AB_3 to AB (the distance we seek), we consider ΔAB_3B. We can *see* only one side of this triangle in our figure, but it is nonetheless a perfectly ordinary triangle. (It may help to compare the analogous situation for a Flatlander contemplating a triangle whose base lies in Flatland, but whose third vertex doesn't.)

Any two points determine a 1-dimensional space, a line. Any three points (provided they are not on the same line) determine a 2-dimensional space, a plane. Any four points (provided they are not all in the same plane) determine a 3-dimensional space, a hyperplane. As we proceed along our scenic route from A to B, each successive road is perpendicular to the whole space determined by the points we've already encountered on the scenic route. For instance, road B_1B_2 is perpendicular to *line* AB_1, and road B_2B_3 is perpendicular to *plane* AB_1B_2. Similarly, road B_3B is perpendicular to *hyperplane* $AB_1B_2B_3$. This means, more specifically, that B_3B is perpendicular to any line inside that hyperplane that passes through B_3. In particular, it means that B_3B is perpendicular to AB_3.

Since lines B_3B and AB_3 are perpendicular, ΔAB_3B is a *right* triangle. We know the lengths of its legs (We found AB_3 above, and $BB_3 = |\Delta w|$), so the Pythagorean Theorem will yield its hypotenuse, and with it, the distance formula for $\mathbb{R}^4$:

$$\boldsymbol{AB} = \sqrt{(\Delta \boldsymbol{x})^2 + (\Delta \boldsymbol{y})^2 + (\Delta \boldsymbol{z})^2 + (\Delta \boldsymbol{w})^2}.$$

So yes, the pattern continues – and the ideas we used to extend the distance formula from $\mathbb{R}^3$ to $\mathbb{R}^4$ work just as well to extend it from $\mathbb{R}^4$ to $\mathbb{R}^5$ and then from $\mathbb{R}^5$ to $\mathbb{R}^6$ and so on and so forth. In short, the general pattern holds in $\mathbb{R}^n$ for *all* values of n. We have therefore established the

Distance Formula in $\mathbb{R}^n$. The distance between any two points is *the root of the sum of the squared differences of the points' Cartesian coordinates.**

A vector's *length* is the distance between its tip and tail. Hence, we can use the distance formula to find a formula for a vector's length, which we'll do next. We denote the length of a vector $\mathbf{v}$ by the symbol $\|\mathbf{v}\|$.

Vector Length Formula.

If $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ is a vector expressed in Cartesian coordinates, then $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots v_n^2}$.*

Proof. Translate the vector so that its tail is at the origin, $(0,0,\dots,0)$. Its tip will be at the point $(v_1, v_2, \dots, v_n)$. As noted above, the vector's length, $\|\mathbf{v}\|$, is the distance between these points. According to the distance formula that we've just proved, this distance will be

$$\sqrt{(v_1 - 0)^2 + (v_2 - 0)^2 + \cdots + (v_n - 0)^2}\ ,$$

which simplifies to the expression claimed above. ■

And we're done. Time for some exercises.

---

* *Cartesian* coordinates are specifically based on a set of *mutually perpendicular axes*, all of which employ the same unit of length. (Equivalently, they are based on the standard basis vectors of $\mathbb{R}^n$.) This is the usual situation, but sometimes - beginning in Chapter 6 - we'll use other coordinate systems that don't meet these criteria. In such cases, the formulas above for distance and vector length do *not* hold. (See Chapter 6, Exercise 5a.) After all, we derived those formulas under the assumption that our coordinates were Cartesian. Namely, we assumed the "roads" on the "scenic routes" between points were perpendicular.

This footnote may give help you appreciate the value of working coordinate-free, as we did in this chapter's first two sections. We established most of our basic vector properties without relying on coordinates of any sort; consequently, we can be certain that those properties will still hold when we work with more exotic coordinate systems in Chapter 6 and beyond.

Exercises.

6. Given points $P = (0,0,0)$, $Q = (2,-2,2)$, $R = (2,0,1)$, and $S = (3,-1,2)$ in $\mathbb{R}^3$,

a) Are the vectors $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ equal? (The notation $\overrightarrow{AB}$ refers to the vector from point A to point B.)

b) Are the vectors $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ parallel? c) Find the lengths of vectors $\overrightarrow{PQ}$ and $\overrightarrow{RS}$.

7. True/False.

a) If vectors have the same length, they are equal. b) If vectors are equal, they have the same length.

c) $\|\mathbf{0}\| = 0$. d) The length of vector in $\mathbb{R}^{100}$ whose entries are all 1 is 1.

8. Given any nonzero vector $\mathbf{v}$, the length of $\mathbf{v}/\|\mathbf{v}\|$ is always 1. Explain why. Also explain why I included the modifier "nonzero". [*Note: Dividing a vector by its length is called "normalizing" it*.]

9. Let $\mathbf{v} = \begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$. Find the following:

a) $\mathbf{v} - \mathbf{w}$ b) $3\mathbf{v} + 2\mathbf{w}$ c) $\mathbf{v} + 2\mathbf{i} - 4\mathbf{j} - \mathbf{k}$ d) $\|\mathbf{w}\|$ e) $\left\|\frac{1}{2}(\mathbf{v} - \mathbf{w})\right\|$ f) $\mathbf{v}/\|\mathbf{v}\|$. g) $\left\|\frac{\mathbf{v}}{\|\mathbf{v}\|}\right\|$

h) $\left\|\frac{1729\mathbf{v} + 8192\mathbf{w}}{\|1729\mathbf{v} + 8192\mathbf{w}\|}\right\|$ [*Hint: No computations are necessary for this part. Just think*.]

10. Let $\mathbf{v} = 3\mathbf{e}_2 + 2\mathbf{e}_7 + \mathbf{e}_{19} + \sqrt{50}\mathbf{e}_{42}$ be a vector in $\mathbb{R}^{50}$. Find $\|\mathbf{v}\|$.

11. Flatlanders would be disturbed not only by three-dimensional space's exotic objects such as spheres that could not exist in their two-dimensional world, but also by three-dimensional space *itself*, since its roominess enables even familiar objects, such as lines, to interact in ways that are inconceivable in two dimensions. For instance, Flatlanders "know" that any pair of nonparallel lines must intersect. In three dimensions, however, nonparallel yet nonintersecting lines can exist.

a) Give an example to explain how such lines (called *skew lines*) can exist in space.

b) If we pick two lines in space at random, they will almost certainly be skew to one another. Explain why.

c) Here in three-dimensional "Spaceland", we are all quite certain that distinct planes can intersect only in a line. But in *four*-dimensional space, a pair of two-dimensional planes can intersect in a single *point*! For example, I claim that the xy-plane (the plane containing the x and y axes) and the zw-plane (defined analogously) intersect in just one point. Justify this assertion.

12. Lines in $\mathbb{R}^2$, planes in $\mathbb{R}^3$, and n-dimensional "hyperplanes" in $\mathbb{R}^{n+1}$ are all described by simple linear equations. To see why, we'll review the simplest case (lines in $\mathbb{R}^2$) and build up our understanding from there.

a) On your mother's knee, you learned that each **line** in $\mathbb{R}^2$ corresponds to a *linear* equation in two variables. (That is, an equation of the form $\boldsymbol{ax + by = c}$.) Yes, but why? To remember, let's go back to your childhood…

Imagine the line through $(0, 5)$ with slope 2, and a movable point on it that we can slide like a bead on a wire. We seek an equation that our moving point's coordinates will satisfy at *all* its possible positions on the line. To do this, let's start at $(0,5)$, taking the second coordinate, 5, as our "baseline" y-value. The line's slope, 2, indicates that if we slide the point so that its first coordinate changes by k units, its second coordinate will change by $\mathbf{2}k$ units. Thus, the point on the line whose first coordinate is x (as opposed to 0) will have a second coordinate of $y = 5 + 2x$. Because this holds for all possible values of x (i.e. for all positions of moving point), it is the line's equation. And of course, we can easily rewrite this equation in the specified form $ax + by = c$ if we wish. [Like so: $2x + (-1)y = 5$]

Your problems: (i) Understand the preceding derivation. (ii) By the same logic, convince yourself that *every* nonvertical line in $\mathbb{R}^2$ has an equation of the form $ax + by = c$. (iii) Explain why I wrote "nonvertical" in the previous sentence. Then explain why vertical lines' equations also conform to the $ax + by = c$ pattern.

b) Every nonvertical **plane** in $\mathbb{R}^3$ is determined by its z-intercept (where it crosses the z-axis) and by *two* slopes: an "x-slope" – the rate at which a moving point confined to the plane changes its third coordinate in response to *first*-coordinate changes – and its "y-slope", the rate at which the moving point's third coordinate changes in response to *second*-coordinate changes.

Given a plane's z-intercept and these two slopes, we can derive its equation. An example will explain how.

Consider a moving point on the plane passing through $(0, 0, 4)$ whose x-slope is 3 and whose y-slope is -5. We want an equation satisfied by the moving point's coordinates wherever it might happen to lie on the plane. We know it can lie at $(0, 0, 4)$, so we'll take 4 as our "baseline" third coordinate (when the others are both 0). The plane's two given slopes tell us that changing the point's first coordinate by x yields a third-coordinate change of $\mathbf{3}x$, while changing its second coordinate by y yields a third-coordinate change of $-\mathbf{5}y$.

Thus, the point on the plane whose first two coordinates are x and y will have the following third coordinate:

$$z = 4 + 3x - 5y.$$

Since this relationship among the three coordinates holds for all possible positions of point on the plane, it is in fact the plane's equation. And of course, it can easily be massaged into in the form $\boldsymbol{ax + by + cz = d}$.

Your problems: (i) Understand the preceding derivation. (ii) Convince yourself that, by the same logic, *every* nonvertical plane in $\mathbb{R}^3$ (i.e. every plane that isn't perpendicular to the xy-plane) has an equation of the form $ax + by + cz = d$. (iii) Explain why I wrote "nonvertical" in the previous sentence. Then explain why vertical planes' equations also conform to the $ax + by + cz = d$ pattern.

c) Two-dimensional space consists of a single plane, but three-dimensional space admits infinitely many planes. Similarly, $\mathbb{R}^4$ admits infinitely many **3-dimensional hyperplanes**, each one of which "looks like" a copy of $\mathbb{R}^3$. A typical hyperplane is determined by its w-intercept (the point where it cuts the w-axis) and by *three* slopes: an "x-slope", "y-slope", and "z-slope" – analogous to the two slopes that you considered above in Part B. As you'd expect, the same logic that held for lines and planes implies that 3-dimensional hyperplanes in $\mathbb{R}^4$ have equations of the form $\boldsymbol{ax + by + cz + dw = k}$.

Your problems: (i) Reflect on the fact that you know what a 3-dimensional hyperplane is like *from the inside*, since you (presumably) live in one: all of $\mathbb{R}^3$. On the other hand, it's hard to imagine what a 3-dimensional hyperplane would look like *from the outside* – from the perspective, that is, a four-dimensional being, who could see a hyperplane's situation relative to other objects in four-dimensional space. (ii) Suppose we have a 3-dimensional hyperplane that passes through the point $(0, 0, 0, 6)$ and whose x, y, and z-slopes are, respectively, $2, 1$, and 7. Find this hyperplane's equation.

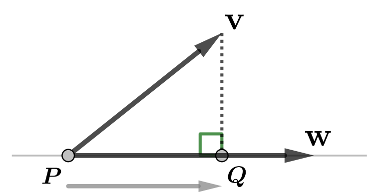
d) Flat unbounded n-dimensional objects are **n-dimensional hyperplanes**. When such beasts reside in $\mathbb{R}^{n+1}$, their equations always have the following form: $\boldsymbol{a_1x_1 + a_2x_2 + \cdots + a_nx_n + a_{n+1}x_{n+1} = c}$.

The moral of the story: Linear equations in algebra correspond to lines, planes, and hyperplanes in geometry. We'll use this insight later in the course when we interpret *systems* of linear equations geometrically.

Your problems: (i) From now on, whenever you encounter an equation such as $6x - 3y + z = 2$ or perhaps $7x + 3y + z - 4w = 0$, think to yourself, "A-ha! A plane" (in the first case) or "a hyperplane!" (in the second). (ii) Based on analogy with lower-dimensional cases, what do you think the intersection of two 3-dimensional hyperplanes in $\mathbb{R}^4$ would look like?

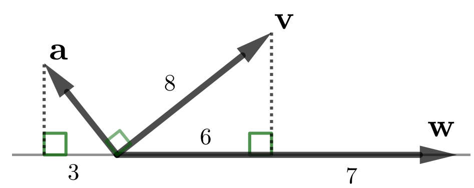
The Dot Product

To be not a man, but a projection of another man's dream –
What incomparable humiliation! What vertigo!
- Jorge Luis Borges, "The Circular Ruins".



Let two vectors $\mathbf{v}$ and $\mathbf{w}$ start at the same point P. Drop a perpendicular from $\mathbf{v}$'s tip to the line containing $\mathbf{w}$, and let Q be the point where it lands. The vector pointing from P to Q is called $\mathbf{v}$'s **orthogonal projection** on $\mathbf{w}$. In the figure, I've colored this orthogonal projection grey and moved it below the line containing $\mathbf{w}$ to make it more visible, but you should imagine it as lying on the line itself, like a shadow cast by $\mathbf{v}$.

We define $\mathbf{v}$'s **scalar projection** on $\mathbf{w}$ as the *length* of $\mathbf{v}$'s orthogonal projection on $\mathbf{w}$, with a negative sign attached to it if the orthogonal projection points in the opposite direction from $\mathbf{w}$.



Thus, in the figure at right, $\mathbf{v}$'s scalar projection on $\mathbf{w}$ is 6, while $\mathbf{a}$'s scalar projection on $\mathbf{w}$ is -3. Or consider two extreme cases: $\mathbf{v}$'s scalar projection on $\mathbf{a}$ is 0, while $\mathbf{v}$'s scalar projection on *itself* is 8. Clearly, any vector's scalar projection on itself is just its length. (So here, $\mathbf{w}$'s scalar projection on $\mathbf{w}$ is 13.)

With the idea of scalar projection in hand, we can now state the definition whose consequences we'll develop in this section.

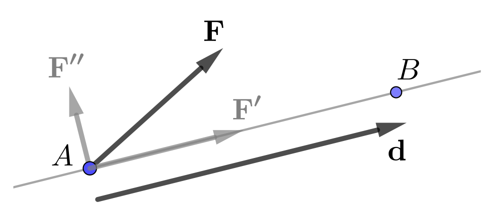
> **Definition.** The **dot product** of two vectors $\mathbf{v}$ and $\mathbf{w}$ (denoted $\mathbf{v}\cdot\mathbf{w}$) is the product of their scalar projections onto $\mathbf{w}$.

For example, in the figure above, we can see that

$$\mathbf{v}\cdot\mathbf{w} = \begin{pmatrix}\mathbf{v}'\text{s scalar}\\ \text{projection on }\mathbf{w}\end{pmatrix}\begin{pmatrix}\mathbf{w}'\text{s scalar}\\ \text{projection on }\mathbf{w}\end{pmatrix} = (6)(13) = 78.$$

The dot product arises naturally in physics, geometry, and anywhere else that vectors are applied.

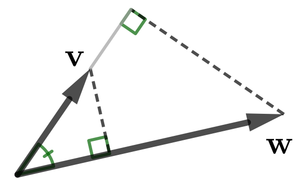
For example, you're probably familiar with the physical concept of **work**, which is defined as follows: If, as a steady force of magnitude F is applied to an object, the object moves d units *in the force's direction*, then we say that the force has done Fd units of work on the object. (So if we steadily apply 100 lbs. of force while pushing a cart 30 feet, we've done 3000 "foot-pounds" of work on the cart.) So far so simple, but what if the force is *not* in the same direction as the object's direction of motion? Suppose, for example, that the grey line in the figure represents railroad tracks, and we steadily apply force $\mathbf{F}$ to push a cart from point A to B. How much work have we done on the cart? The answer here is *not* $\|\mathbf{F}\|\|\mathbf{d}\|$, since pushing at an angle to the tracks "wastes" some of our force. If we write $\mathbf{F} = \mathbf{F}' + \mathbf{F}''$, as in the figure, we see that only the $\mathbf{F}'$ component of our force vector does work pushing the cart down the track. Accordingly, the work done on the cart is $\|\mathbf{F}'\|\|\mathbf{d}\|$, which in fact is $\mathbf{F}\cdot\mathbf{d}$. Thus we see that work, viewed in a more general setting, is simply a dot product.



We'll soon see how to use the dot product to find the angle between any two vectors. But before we can do that, we must first develop the dot product's basic algebraic and geometric properties, a process that will help fix the dot product's definition as a product of two scalar projections in your mind.

Claim 1. The dot product is *commutative*: $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$.

Proof. The figure at right depicts two typical vectors, $\mathbf{v}$ and $\mathbf{w}$. Dropping perpendiculars from each tip to the line containing the other vector produces similar right triangles. By similarity, we have



$$\frac{\mathbf{v}'\text{s scalar projection on } \mathbf{w}}{\mathbf{v}'\text{s scalar projection on } \mathbf{v}} = \frac{\mathbf{w}'\text{s scalar projection on } \mathbf{v}}{\mathbf{w}'\text{s scalar projection on } \mathbf{w}}.$$

After clearing fractions, the dot product's definition yields $\mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v}$, as claimed. ■

Commutativity lets us refer casually to two vectors' dot product without having to worry about their order.

Claim 2. (A Criterion for Perpendicularity)
Two nonzero vectors are *perpendicular* if – and only if – their dot product is *zero*.

Proof. If $\mathbf{v}$ and $\mathbf{w}$ *are* perpendicular, $\mathbf{v}$'s scalar projection on $\mathbf{w}$ is clearly zero, so $\mathbf{v} \cdot \mathbf{w}$ itself is zero. If $\mathbf{v}$ and $\mathbf{w}$ are *not* perpendicular, it's clear that the two scalar projections whose product is $\mathbf{v} \cdot \mathbf{w}$ are both *non*zero. Accordingly, $\mathbf{v} \cdot \mathbf{w}$, as the product of these two nonzero scalars, is nonzero. ■

Our next result is, if anything, even more geometrically obvious. It is often algebraically useful.

Claim 3. Dotting a vector with itself is the same as squaring its length. (That is, $\mathbf{v} \cdot \mathbf{v} = \|\mathbf{v}\|^2$.)

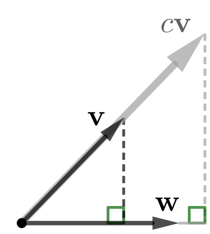
Proof. By the dot product's definition, we have

$$\mathbf{v} \cdot \mathbf{v} = \begin{pmatrix} \mathbf{v}'\text{s scalar} \\ \text{projection on } \mathbf{v} \end{pmatrix} \begin{pmatrix} \mathbf{v}'\text{s scalar} \\ \text{projection on } \mathbf{v} \end{pmatrix} = \|\mathbf{v}\|\|\mathbf{v}\| = \|\mathbf{v}\|^2. \quad ■$$

The next result has an associative flavor.

Claim 4. Scalar multiples can be "pulled to the front" of a dot product.
(That is, $(c\mathbf{v}) \cdot (d\mathbf{w}) = cd(\mathbf{v} \cdot \mathbf{w})$.)

Proof. The figure shows $\mathbf{v}$, $c\mathbf{v}$, and $\mathbf{w}$. (We'll deal with d shortly.) By similarity, the large right triangle's sides are c times as long as the smaller triangle's sides. This scaling factor justifies the *second* equals sign in what follows:



$$(c\mathbf{v}) \cdot \mathbf{w} = \begin{pmatrix} c\mathbf{v}'\text{s scalar} \\ \text{proj. on } \mathbf{w} \end{pmatrix} \|\mathbf{w}\| = \left(c \begin{pmatrix} \mathbf{v}'\text{s scalar} \\ \text{proj. on } \mathbf{w} \end{pmatrix} \right) \|\mathbf{w}\|$$
$$= c\left(\begin{pmatrix} \mathbf{v}'\text{s scalar} \\ \text{proj. on } \mathbf{w} \end{pmatrix} \|\mathbf{w}\| \right) = c(\mathbf{v} \cdot \mathbf{w}).$$

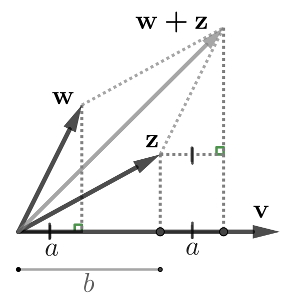
(The dot product's definition and ordinary multiplication's associativity justify the other equals signs.) Having thus deduced that scalar multiples of the *first* "factor" can be pulled to the front, we can alternately use this fact and commutativity to work our way towards the full result:

$$(c\mathbf{v}) \cdot (d\mathbf{w}) = c\big(\mathbf{v} \cdot (d\mathbf{w})\big) = c\big((d\mathbf{w}) \cdot \mathbf{v}\big) = c\big(d(\mathbf{w} \cdot \mathbf{v})\big) = c\big(d(\mathbf{v} \cdot \mathbf{w})\big) = cd(\mathbf{v} \cdot \mathbf{w}) \quad ■$$

The preceding proof, like that of Claim 1, needs a slight adjustment to account for certain cases. (Namely, the cases in which an *obtuse* angle separates the vectors, and/or if the scalar is negative.) The same is true for the next two proofs, but rather than doubling this section's size by essentially repeating each proof twice (or more), I'll leave those cases as exercises for the interested reader.*

Claim 5. The dot product distributes over vector addition.

Proof. Using the intricate figure at right, we'll show that "dotting" a vector $\mathbf{v}$ with a vector sum $(\mathbf{w}+\mathbf{z})$ yields the same thing as "dotting" $\mathbf{v}$ with $\mathbf{w}$ and $\mathbf{z}$ individually and adding the results.



Note the figure's three segments marked with a tick. (Two are equal because they are corresponding sides of congruent right triangles. Two are equal because they are opposite sides of a rectangle.) I've called their common length a, and I've introduced the symbol b for another length that we'll need: $\mathbf{z}$'s scalar projection on $\mathbf{v}$.

With this notation in place, we see that

$$(\mathbf{w}+\mathbf{z})\cdot\mathbf{v} = (a+b)\|\mathbf{v}\| = a\|\mathbf{v}\| + b\|\mathbf{v}\| = (\mathbf{w}\cdot\mathbf{v}) + (\mathbf{z}\cdot\mathbf{v}).^{\dagger}$$

Thus, the dot product distributes over vector addition as claimed.‡ ■

Next, we'll see how to compute two vectors' dot product in terms of the vectors' *Cartesian coordinates*. I'll give the argument just for vectors in $\mathbb{R}^2$, but it generalizes in an obvious way to $\mathbb{R}^n$.

Claim 6. If $\mathbf{v} = \binom{v_1}{v_2}$ and $\mathbf{w} = \binom{w_1}{w_2}$ are two vectors in $\mathbb{R}^2$ expressed in Cartesian coordinates, then

$$\mathbf{v}\cdot\mathbf{w} = v_1w_1 + v_2w_2.^{\S}$$

Proof.

$$\begin{aligned}
\mathbf{v}\cdot\mathbf{w} &= (v_1\mathbf{i}+v_2\mathbf{j})\cdot(w_1\mathbf{i}+w_2\mathbf{j}) \\
&= (v_1\mathbf{i}+v_2\mathbf{j})\cdot(w_1\mathbf{i}) + (v_1\mathbf{i}+v_2\mathbf{j})\cdot(w_2\mathbf{j}) && \text{(by Claim 5)} \\
&= (v_1\mathbf{i}\cdot w_1\mathbf{i}) + (v_2\mathbf{j}\cdot w_1\mathbf{i}) + (v_1\mathbf{i}\cdot w_2\mathbf{j}) + (v_2\mathbf{j}\cdot w_2\mathbf{j}) && \text{(Claim 5)} \\
&= (v_1w_1)(\mathbf{i}\cdot\mathbf{i}) + (v_2w_1)(\mathbf{j}\cdot\mathbf{i}) + (v_1w_2)(\mathbf{i}\cdot\mathbf{j}) + (v_2w_2)(\mathbf{j}\cdot\mathbf{j}) && \text{(Claim 4)} \\
&= (v_1w_1)\|\mathbf{i}\|^2 + (v_2w_1)(0) + (v_1w_2)(0) + (v_2w_2)\|\mathbf{j}\|^2 && \text{(Claims 2 and 3)} \\
&= (v_1w_1)1 + (v_2w_1)(0) + (v_1w_2)(0) + (v_2w_2)1 && \text{(Definitions of } \mathbf{i} \text{ and } \mathbf{j}\text{.)} \\
&= v_1w_1 + v_2w_2.
\end{aligned}$$
■

* This tradition is as old as geometry itself; in *The Elements* (c. 300 BC), Euclid often presents the proof of just one case, leaving the similar variations that cover other cases to his readers.

† The first and last equals signs are justified by the dot product's definition. The middle one is justified by distribution of ordinary multiplication over addition.

‡ Strictly speaking, the preceding proof has only established that the dot product distributes over vector addition "from the right", but the dot product's *commutativity* (Claim 1) lets us turn the "factors" in all three dot products around, thereby establishing, with no further work, that the dot can be distributed "from the left" as well: $\mathbf{v}\cdot(\mathbf{w}+\mathbf{z}) = (\mathbf{v}\cdot\mathbf{w}) + (\mathbf{v}\cdot\mathbf{z})$.

§ Note well: This result presumes that we are using *Cartesian* coordinates (i.e. based on $\mathbf{i}$ and $\mathbf{j}$). If we were to use a different coordinate system (as we often will, beginning in Chapter 6), then the coordinates of $\mathbf{v}$ and $\mathbf{w}$ would no longer represent weights for $\mathbf{i}$ and $\mathbf{j}$; as a result, our proof of Claim 6 would break down at its first step. Indeed, when we use a non-Cartesian coordinate system, the expression $v_1w_1 + v_2w_2$ typically has no geometric significance at all – and is *not* equal to the vectors' dot product.

The preceding proof clearly generalizes to $\mathbb{R}^n$, leading to the following result:

The Dot Product Formula.
We can compute the dot product of two vectors in $\mathbb{R}^n$ by summing the products of their corresponding Cartesian coordinates.*

It is this recipe that makes the dot product so easy to use. For example, consider these two vectors in $\mathbb{R}^5$:

$$\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ -3 \\ 0 \\ 2 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 3 \\ -2 \\ 3 \\ 8 \\ 5 \end{pmatrix}.$$

Is there anything we can say about them? Yes: According to our dot product recipe, we see that

$$\mathbf{v} \cdot \mathbf{w} = (1)(3) + (2)(-2) + (-3)(3) + (0)(8) + (2)(5) = 0.$$

Since their dot product is zero, Claim 2 tells us that $\mathbf{v}$ and $\mathbf{w}$ are *perpendicular*. With a simple calculation, we've just deduced something about these two vectors in $\mathbb{R}^5$ that's not obvious about them at first glance. In fact, the dot product is useful not only for detecting right angles, but for finding the angle (of any size) between any two vectors. To do that, we'll need the seventh and last result for this section.

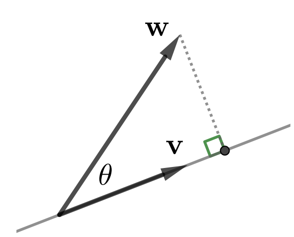
Claim 7. If θ is the angle between vectors $\mathbf{v}$ and $\mathbf{w}$, then the following relationship holds:

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta.$$

Proof. Consider the figure at right. By the definition of the dot product, $\mathbf{v} \cdot \mathbf{w}$ is the product of $\|\mathbf{v}\|$ and $\mathbf{w}$'s scalar projection onto $\mathbf{v}$. By basic right-angle trigonometry, that scalar projection is $\|\mathbf{w}\| \cos\theta$. Thus,

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta,$$

as claimed. ■



Claim 7 relates the angle between two vectors to the vectors' *dot product* and their *lengths*. Since we can compute the dot product and lengths in seconds from the vectors' components, it follows that we can find the angle between the vectors almost as quickly.

Example. Find the angle between the vectors $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j} - \mathbf{k}$ and $\mathbf{w} = \mathbf{i} + 4\mathbf{k}$.

Solution. If we call the angle θ, then Claim 7 tells us that

$$\cos\theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} = \frac{(2)(1) + (3)(0) + (-1)(4)}{\sqrt{14}\sqrt{17}} = \frac{-2}{\sqrt{238}}.$$

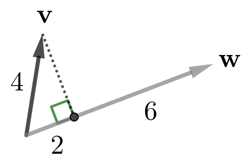
Consequently, we may conclude that the angle between the vectors is

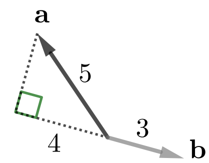
$$\cos^{-1}(-2/\sqrt{238}) \approx 97.4°.$$ ♦

* It's important that these be *Cartesian* coordinates, as discussed in the previous page's last footnote (and in Ch. 6, Exercise 5b).

Exercises.

13. In the figures below, find $\mathbf{v} \cdot \mathbf{w}$ and $\mathbf{a} \cdot \mathbf{b}$ by appealing directly to the dot product's definition.





14. We've seen that two vectors are perpendicular if their dot product is zero. What can we say about the angle between two vectors whose dot product is positive? What if the dot product is negative? Explain your answer *without using any trigonometry*.

15. To the nearest tenth of a degree, find the angle separating each pair of vectors:

a) $\mathbf{v} = \begin{pmatrix} -1 \\ 5 \\ 2 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}$. b) $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$, $\mathbf{w} = \mathbf{i} - 2\mathbf{j}$. c) $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} 6 \\ -3 \\ -1 \\ 0 \end{pmatrix}$

16. Let $\mathbf{v} = 8\mathbf{i} + 4\mathbf{j} + 3\mathbf{k}$ and $\mathbf{w} = -2\mathbf{i} + \mathbf{j} + 4\mathbf{k}$. Are these vectors perpendicular?

17. Find a vector in $\mathbb{R}^5$ perpendicular to $\mathbf{v} = \mathbf{e_1} + 2\mathbf{e_2} - 3\mathbf{e_3} - \mathbf{e_4} + 2\mathbf{e_5}$

18. Our proof of the dot product's commutativity (Claim 1) relies on a figure that appears to apply only when the vectors are separated by an acute angle. Our claim, however, is universal: commutativity holds no matter what angle separates the vectors. To establish this universality, do the following – without appealing to any of our later claims (such as the formulas we derived for the dot product in Claims 6 and 7):

a) Explain why commutativity holds if the vectors are separated by a right angle.

b) Explain why commutativity holds if the vectors are separated by an obtuse angle.

19. Claim 7 asserts that $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\| \cos\theta$, where θ is the angle between $\mathbf{v}$ and $\mathbf{w}$. Strictly speaking, we only proved the case in which the vectors are separated by an acute angle. The result does, however, hold in all cases. Prove that this is so by explaining why it holds in each of the two remaining cases: when the angle between $\mathbf{v}$ and $\mathbf{w}$ is right, and when the angle is obtuse.

20. We've seen that the dot product behaves like ordinary multiplication in certain algebraic ways: It is commutative, and it distributes over (vector) addition. Your question: Are we able to "cancel" factors from both sides of an equation in which the dot product appears? In symbols, if we know that $\mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{c}$ for three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, can we "cancel" the $\mathbf{a}$'s and conclude that $\mathbf{b} = \mathbf{c}$? If so, prove it. If not, provide a counterexample.

21. Why didn't we prove that the dot product is *associative*?
(That is, why didn't we prove that $\mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c}$ for all vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$?)

22. The **Cauchy-Schwarz Inequality** states that $|\mathbf{v} \cdot \mathbf{w}| \leq \|\mathbf{v}\|\|\mathbf{w}\|$ for any two vectors $\mathbf{v}$ and $\mathbf{w}$.

a) Verify that that the Cauchy-Schwarz inequality holds for the two vectors in each part of Exercise 15.

b) Prove that the Cauchy-Schwarz Inequality holds for *all* vectors $\mathbf{v}$ and $\mathbf{w}$.

c) The Cauchy-Schwarz Inequality contains the less-than-or-equal symbol. Can we strengthen it to strictly *less than*? Can the two sides ever be equal? If so, under what circumstances? If not, why not?

23. The **Triangle Inequality** states that $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ for any two vectors $\mathbf{v}$ and $\mathbf{w}$. Repeat all three parts of the previous problem, replacing "Cauchy-Schwarz" with "Triangle" wherever it occurs.

Chapter 2
Vocabulary

Linear Combinations, Span, Linear Independence

A combination and a form indeed
Where every god did seem to set his seal

- Hamlet, Act 3, Scene 4.

To understand linear algebra, you must first master its vocabulary. In this chapter, I'll introduce the core vocabulary, starting with the simple idea of a linear combination.

> **Definition.** Let $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$ be of any set of vectors. Then any vector of the form
>
> $$c_1\mathbf{v_1} + c_2\mathbf{v_2} + \cdots + c_n\mathbf{v_n},$$
>
> (where the c_i's are scalars) is called a **linear combination** of the original vectors.

Less formally, a linear combination of vectors is a *weighted sum* of them (where the weights are scalars).

Example 1. If $\mathbf{v} = \binom{1}{2}$ and $\mathbf{w} = \binom{2}{1}$, then $2\mathbf{v} + 3\mathbf{w} = \binom{8}{7}$ is a linear combination of $\mathbf{v}$ and $\mathbf{w}$. ♦

Example 2. If $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ are vectors, then $2\mathbf{a} + 3\mathbf{b} - 4\mathbf{c} + \mathbf{d}$ is a linear combination of them. ♦

Example 3. If $\mathbf{v}$ is a vector, then a linear combination of $\mathbf{v}$ alone is just a scalar multiple of $\mathbf{v}$. ♦

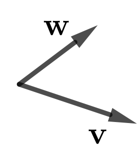
Throughout linear algebra, we often blur the distinction between *vectors* (arrows) and their *tips* (points).* This simple convention will help us make geometric sense of the following term.

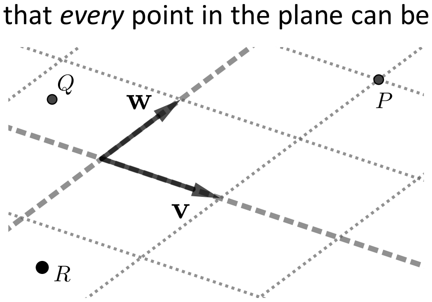
> **Definition.** The **span** of a set of vectors is the set of *all* their linear combinations.

By the convention above, we can think of the span of a vector set as a set of *points*.

Example 4. The span of any one nonzero vector $\mathbf{v}$ in $\mathbb{R}^2$ is, by definition, the set of all linear combinations of $\mathbf{v}$. By Example 3 above, this set consists of all scalar multiples of $\mathbf{v}$. Since the tips of these scalar multiples clearly lie on a line (namely, the line that $\mathbf{v}$ lies on when its tail is at the origin), we can say that $\mathbf{v}$'s span, viewed geometrically, is a line through the origin in $\mathbb{R}^2$.

Example 5. The span of $\mathbf{v}$ and $\mathbf{w}$ in the figure at right is the *plane* containing them. Vectors $\mathbf{v}$ and $\mathbf{w}$ determine a "grid" (below left) that covers this plane and reveals that *every* point in the plane can be reached with a linear combination of $\mathbf{v}$ and $\mathbf{w}$.





For example, we can reach point P with the linear combination $\mathbf{v} + 2\mathbf{w}$, and we can reach point Q with $(-3/4)\mathbf{v} + (1/2)\mathbf{w}$. I'll leave it to you to state a plausible linear combination of $\mathbf{v}$ and $\mathbf{w}$ that will reach R.

Finally, note how vectors $\mathbf{v}$ and $\mathbf{w}$ act somewhat like $\mathbf{i}$ and $\mathbf{j}$, providing the framework for a system of coordinates. We'll soon expand upon this idea. ♦

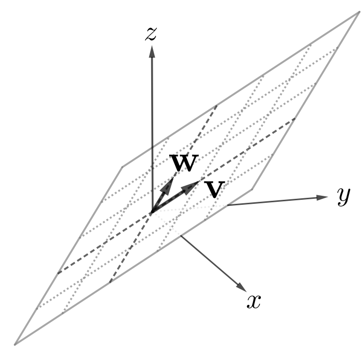
* More precisely, we associate a vector with the point where its tip lies *when its tail is placed at the origin*.

Example 6. Let $\mathbf{v}$ be any nonzero vector in $\mathbb{R}^3$. The span of $\mathbf{v}$ alone is a line through the origin. Now consider the span of $\mathbf{v}$ *and* $2\mathbf{v}$. Despite our introduction of the second vector, the set's span remains the same because its second vector, $2\mathbf{v}$, was already in the span of its first, $\mathbf{v}$.

As Example 6 suggests, if we wish to enlarge a vector set's span, then we must introduce a new vector from outside the existing span. Here is some relevant terminology that we'll use from this point forward.

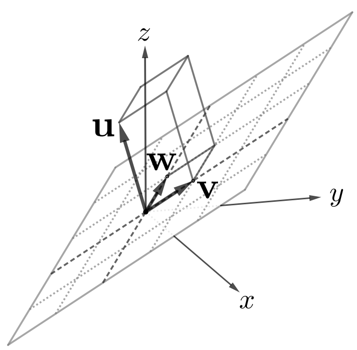
> **Definitions.** We say that a vector $\mathbf{v}$ is **linearly independent** *of a set of vectors* if $\mathbf{v}$ lies outside the span of the vectors in the set.
>
> We say that a set of vectors is itself **linearly independent** if *each* of its vectors is independent of the set of all its other vectors.

For example, the vectors $\mathbf{v}$ and $\mathbf{w}$ from Example 5 constitute a linearly independent set, while vectors $\mathbf{v}$ and $2\mathbf{v}$ from Example 6 do not. (In other words, vectors $\mathbf{v}$ and $2\mathbf{v}$ constitute a linearly *de*pendent set.)



Example 7. In the figure at right, $\mathbf{v}$ and $\mathbf{w}$ are two linearly independent vectors in $\mathbb{R}^3$. As such, their span is a *plane* through the origin. They determine a grid on this plane, as in Example 5.

If $\mathbf{u}$ is any vector that does *not* lie in the plane spanned by $\mathbf{v}$ and $\mathbf{w}$, then vectors $\mathbf{v}, \mathbf{w},$ and $\mathbf{u}$ are a linearly independent set. Exercising your visual imagination, you should be able to see that these three vectors provide the foundation for a *three-dimensional* grid: a grid encompassing all of $\mathbb{R}^3$. Whereas a two-dimensional grid cuts a plane into a neatly stacked system of parallelograms, a three-dimensional grid divides a three-dimensional space into a neatly stacked system of *parallelepipeds*, one of which is drawn at left.* You should imagine the others.



Since $\mathbf{v}, \mathbf{w}$ and $\mathbf{u}$ span all of $\mathbb{R}^3$, adding *any* fourth vector to this trio would yield a linearly *de*pendent set, since the fourth vector would already be in the span of the first three. ♦

Meditating on the preceding examples shows how limited the possibilities are for the span of a vector set. First, every vector set's span **contains the origin**, since we can always achieve $\mathbf{0}$ as a linear combination by letting all the scalar weights be zeros. Next, the span **must be a line, plane, or hyperplane.** (You met hyperplanes in Exercise 12 of Chapter 1.) By itself, a single nonzero vector spans a line. Two linearly independent vectors determine a two-dimensional grid, and thus their span will be a plane. Three linearly independent vectors determine a three-dimensional grid, and therefore their span will be a three-

* The word *parallelepiped* is properly pronounced parallel-EPI-ped, to reflect its etymological meaning: *parallels on feet* – the parallels being any two opposite faces, the feet being the edges connecting them. Many mathematicians, alas, mispronounce it parallel-uh-PIE-ped, as if trying to rhyme it with *biped*. For a charming British mispronunciation, see Canto VII of Lewis Carroll's "Phantasmagoria", in which Carroll – himself a mathematician – rhymes "parallelepiped" with "insipid".

dimensional hyperplane. And indeed, the general rule is since that n linearly independent vectors determine an n-dimensional grid, their span will be an n-dimensional hyperplane.

Fortunately, our ability to work with objects in four or more dimensions doesn't depend on our ability to visualize them directly, as the next example shows.

Example 8. Let $\mathbf{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$, $\mathbf{c} = \begin{pmatrix} 1 \\ 6 \\ 7 \\ 4 \end{pmatrix}$ be vectors in $\mathbb{R}^4$. Are they linearly independent?

Solution. Although our initial attempts to visualize this will surely fail, arithmetic will settle the question definitively: Since, as you can verify, $\mathbf{c} = \mathbf{a} + 4\mathbf{b}$, we see that one vector lies within the others' span. Accordingly, these three vectors do *not* constitute a linearly independent set. ♦

Two comments on the preceding example are in order.

First, once we've recognized that $\mathbf{c} = \mathbf{a} + 4\mathbf{b}$, we can visualize the vectors' arrangement more readily: Vectors $\mathbf{a}$ and $\mathbf{b}$ are linearly independent of each other (since they're not scalar multiples of each other), so they span an ordinary two-dimensional plane. Then, since $\mathbf{c}$ is a linear combination of $\mathbf{a}$ and $\mathbf{b}$, we know that it lies in the plane that they span. Thus, all three vectors lie in a single plane in $\mathbb{R}^4$, and more specifically, a plane *through the origin*; as discussed above, the span of any vector set contains $\mathbf{0}$.

Second, verifying that $\mathbf{c} = \mathbf{a} + 4\mathbf{b}$ was easy once I suggested it to you, but how could you have found that relationship in the first place? In Example 8, a few minutes of staring, thinking, and experimenting might have done the trick, but that's hardly a reliable "technique". (Good luck using it to spot a linear relationship among 20 vectors in $\mathbb{R}^{30}$.) In time, we'll develop an algorithmic test for linear independence, but before we can make use of it, we'll need to build up our algebraic machinery. For now, we can establish two preliminary results that will at least point us in the right direction. The first of these two results will help us demonstrate linear independence in certain special cases.

Claim 1. If each vector in a list is shown to be linearly independent of its *predecessors* on the list, then the vectors on the list constitute a linearly independent set.*

Proof. Let the list be $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$. Showing that $\mathbf{v_2}$ lies outside of $\mathbf{v_1}$'s span will establish that the span of $\mathbf{v_1}$ and $\mathbf{v_2}$ is two-dimensional. If we then show that $\mathbf{v_3}$ lies outside its predecessors' two-dimensional span, it will follow that the span of $\mathbf{v_1}, \mathbf{v_2}$ and $\mathbf{v_3}$ is *three*-dimensional… and so on and so forth until we exhaust the full set and conclude that its span is an $\boldsymbol{n}$-dimensional hyperplane. This n-dimensional hyperplane's existence implies that the vector set can't be linearly *de*pendent. After all, if one of the $\mathbf{v}_i$ lay in the span of some of the others (not necessarily its predecessors), it would be "redundant": removing it from the vector set wouldn't alter their span. (Cf. Example 6.) *If* this were so, the full set of n vectors would have the same span as the reduced set of $(n-1)$. But this *can't* be, because $(n-1)$ vectors obviously can span *no more than* $(n-1)$ dimensions, yet our full set spans $\boldsymbol{n}$ dimensions. Thus, the full set must be linearly independent, as claimed. ■

* For example, this says that to show that the vector set $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ is linearly independent, one need not explicitly demonstrate that each of the three vectors is independent of the other two. Instead, it suffices to establish just two facts: that $\mathbf{b}$ is independent of $\mathbf{a}$, and that $\mathbf{c}$ is independent of $\mathbf{a}$ and $\mathbf{b}$.

The preceding result lets us easily spot (or construct) certain "nice" linearly independent vector sets simply by thinking about their coordinates, without having to visualize the vectors in space.

Example 9. Are the following vectors in $\mathbb{R}^4$ linearly independent?

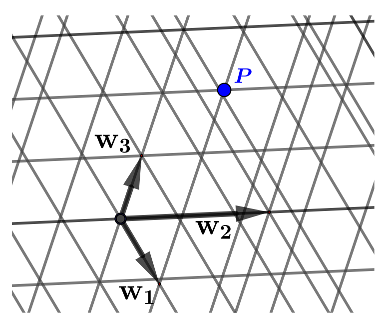
$$\mathbf{v_1} = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v_2} = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v_3} = \begin{pmatrix} 6 \\ 8 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v_4} = \begin{pmatrix} 2 \\ 3 \\ 0 \\ 5 \end{pmatrix}$$

Solution. Clearly, $\mathbf{v_2}$ is independent of $\mathbf{v_1}$, as these vectors aren't scalar multiples of one another. Next, we can easily see that $\mathbf{v_3}$ is independent of its predecessors: Any linear combination of its predecessors will have a 0 in the second slot, so *no* such linear combination could equal $\mathbf{v_3}$, which has an 8 in that slot. Finally, applying the same idea to the 4th slot shows that $\mathbf{v_4}$ is independent of its three predecessors. So, by Claim 1, these four vectors are linearly independent. ■

We'll approach our second big linear independence result by contrasting the "clean grids" induced by linearly *independent* vector sets and the "tangled grids" induced by linearly *dependent* vector sets.

Any set of linearly *independent* vectors $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$ provides the framework for a "clean grid" in $\mathbb{R}^n$: one of those familiar lattices of parallelograms, parallelepipeds, or their higher-dimensional analogues (called *parallelotopes*) that we've considered already in Examples 5 and 7 (in $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively). If we form a linear combination of our linearly independent $\mathbf{v}_i$ vectors, setting all the scalar weights to 0, the resulting linear combination will obviously be $\mathbf{0}$. It's also obvious that *no other* set of scalar weights would yield $\mathbf{0}$ as a linear combination of the $\mathbf{v}_i$ vectors, since each nonzero weight would move us into a different dimension that we haven't yet "visited", which can only take us further from the origin.

On the other hand, a set of linearly *dependent* vectors $\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}$ produces a "tangled grid": a labyrinth of intersecting parallelograms, parallelepipeds, or parallelotopes. For example, the tangled grid at right is induced by three linearly dependent vectors in $\mathbb{R}^2$. Look closely and observe that it consists of *three superimposed clean grids:* one for each pair of $\mathbf{w}_i$ vectors. Every linearly dependent vector set generates a tangled grid, a messy jumble of superimposed clean grids. This superposition opens the door to something we couldn't do before: We can now "reach" $\mathbf{0}$ as a *nontrivial* linear combination of the vectors in our set. The method is simple: We follow one underlying clean grid from the origin to some other point, and then we return to the origin by a different clean grid. For example, in the picture above, we might start from the origin, follow the "$\mathbf{w_1}$, $\mathbf{w_2}$ clean grid" out to point P (it looks like $-2\mathbf{w_1} + 1.1\mathbf{w_2}$ will work), and then return home via the "$\mathbf{w_1}$, $\mathbf{w_3}$ clean grid" (something like $-\mathbf{w_1} - 3\mathbf{w_3}$ will do the trick). Summing up these results of these two displacements, we find that



$$-3\mathbf{w_1} + 1.1\mathbf{w_2} - 3\mathbf{w_3} = \mathbf{0}.$$

Lo and behold: a *nontrivial* linear combination of the linearly dependent vectors equal to $\mathbf{0}$. In fact, there are *infinitely many* ways to express $\mathbf{0}$ as a linear combination of this (or any other) linearly dependent set, since there are infinitely many points (besides P) that we might have used in our construction, each of which yields a different nontrivial linear combination equal to $\mathbf{0}$.*

Moreover, this profusion of representations in terms of linearly dependent vectors isn't limited to $\mathbf{0}$. Let $\mathbf{w}$ be *any* vector in the span of a linearly dependent vector set $\mathbf{w_1}, \mathbf{w_2}, \dots, \mathbf{w_n}$. Clearly, the set's tangled grid lets us represent $\mathbf{w}$ as a linear combination of the $\mathbf{w_i}$ in a different way for each underlying *clean* grid. If we then add any nontrivial "$\mathbf{0}$-combination" to any specific representation of $\mathbf{w}$, we'll get a different representation of $\mathbf{w}$ (thanks to the nonzero scalars in $\mathbf{0}$'s representation). Since there are infinitely many $\mathbf{0}$s that we might add to $\mathbf{w}$ this way, it follows that there are infinitely many representations of $\mathbf{w}$, too.

This stands in stark contrast to the analogous situation with linearly *independent* vectors, $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$. These generate a clean grid, and it is of course geometrically obvious that every point in a clean grid can be "reached" in just one way as a linear combination of the $\mathbf{v_i}$ vectors.

To recapitulate, if a vector set is linearly independent, then *every vector in its span* can be "reached" as a linear combination of the set's vectors in a **unique** way; on the other hand, if the vector set is linearly *de*pendent, then every vector in its span can be reached by **infinitely many** different linear combinations. From this general result, we'll now extract one little piece, which will give us a purely algebraic way to characterize linearly independent sets:†

Claim 2. (Alternate characterization of linear independence)

$$\begin{bmatrix} \mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n} \\ \text{are linearly independent} \end{bmatrix} \Leftrightarrow \begin{bmatrix} \mathbf{0} \text{ can be expressed as a linear combination of } \mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n} \\ \text{in only one way (the trivial way, where all coefficients} = 0). \end{bmatrix}$$

This alternate characterization will often serve as a bridge between geometric visualization and algebraic symbol-pushing. We'll eventually be able to use it to conduct algorithmic tests for linear independence – but only after we've developed a robust technique for solving systems of linear equations, which will be the main focus of Chapter 4.

Finally, for those who doubt their geometric intuition – or just want to supplement it with algebra – I've tucked into the footnotes an algebraic proof that a vector set admits a nontrivial linear combination equal to $\mathbf{0}$ if and only if the vector set is linearly dependent.‡

---

* Changing P isn't the only way to get new expressions for $\mathbf{0}$. We could, instead, change the pair of clean grids we used to get to and from P. Or, simpler still, we could just multiply both sides of our existing nontrivial expression for $\mathbf{0}$ by any nonzero scalar.

† Many linear algebra textbooks adopt the alternate characterization as their *definition* of linear independence. They then prove our definition of linear independence as a theorem. Although their approach is logically valid, I find it psychologically backwards; a compelling logical or mathematical argument should proceed from the intuitive to the unintuitive, not the other way around.

‡ **Proof**. Let $\mathbf{a_1}, \mathbf{a_2}, \dots, \mathbf{a_n}$ be a set of vectors. First, let's suppose they are linearly dependent. By Claim 1, at least one of them is a linear combination of its *predecessors*: $\mathbf{a_k} = c_1\mathbf{a_1} + c_2\mathbf{a_2} + \cdots + c_{k-1}\mathbf{a_{k-1}}$. Subtracting $\mathbf{a_k}$ from both sides yields a nontrivial linear combination of the vectors equal to $\mathbf{0}$. Thus, linear dependence implies the existence of a nontrivial linear combination equal to $\mathbf{0}$. Now we'll prove the converse of that statement. Suppose the given vectors admit a linear combination equal to $\mathbf{0}$. In that case, we can clearly isolate one of vectors on one side of the equation, thereby demonstrating that it lies within the span of the other vectors in the set. Hence, the vector set is linearly dependent. We've now demonstrated that linear dependence is equivalent to the existence of a nontrivial linear combination equal to $\mathbf{0}$. It follows that linear *in*dependence is equivalent to the *non*existence of a nontrivial linear combination equal to $\mathbf{0}$. ■

Exercises.

1. True/False (and explain your answer)

a) Given any set of vectors, some linear combination of them equals $\mathbf{0}$.
b) Given any set of vectors, $\mathbf{0}$ is in their span.
c) If the span of $\mathbf{v}$ and $\mathbf{w}$ is $\mathbb{R}^2$, then some linear combination of $\mathbf{v}$ and $\mathbf{w}$ equals $\pi\mathbf{i} - e\mathbf{j}$.
d) The span of *three* vectors can be all of $\mathbb{R}^2$.
e) The span of *two* vectors can be all of $\mathbb{R}^3$.
f) Three vectors in $\mathbb{R}^2$ can be linearly independent.
g) Three vectors in $\mathbb{R}^3$ can be linearly independent.
h) Three vectors in $\mathbb{R}^4$ can be linearly independent.
i) If $\mathbf{a}, \mathbf{b}, \mathbf{c}$, and $\mathbf{d}$ are linearly independent, then their span is the same as the span of $\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$.

2. Are the following vector sets linearly independent? Explain.

a) $\mathbf{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$, $\mathbf{u} = \begin{pmatrix} 6 \\ 4 \end{pmatrix}$ b) $\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 0 \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} 0 \\ 1 \\ 5 \\ 0 \end{pmatrix}$, $\mathbf{u} = \begin{pmatrix} 0 \\ 0 \\ 8 \\ 2 \end{pmatrix}$ c) $\mathbf{i}$, $\mathbf{j}$, $2\mathbf{i} + 3\mathbf{k}$, $\mathbf{k}$

3. Can the zero vector be in a linearly independent set of two or more vectors? Why or why not?

4. Describe, geometrically, the span of each of the following vector sets in $\mathbb{R}^2$:

a) $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$, $\begin{pmatrix} -6 \\ -12 \end{pmatrix}$ b) $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 2 \end{pmatrix}$ c) $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 8 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ d) $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 12 \\ 4 \end{pmatrix}$, $\begin{pmatrix} -6 \\ -2 \end{pmatrix}$ e) $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$

5. Give an example of…
a) A set of five vectors in $\mathbb{R}^2$ whose span is the line $y = 3x$.
b) A set of three vectors in $\mathbb{R}^3$ whose span is the yz-plane. (That is, the unique plane containing the y and z axes.)
c) A set of five vectors in $\mathbb{R}^4$ whose span is all of $\mathbb{R}^4$.

6. As we've seen, our alternate characterization of linear independence is really just a special case of the following much stronger statement:

> If a vector set is linearly independent, then *every vector in its span* can be "reached" as a linear combination of the set's vectors in **just one way**; on the other hand, if the vector set is linearly *de*pendent, every vector in it span can be reached by **infinitely many** different linear combinations.

In this problem, you'll explore this idea a bit.

a) The vectors $\mathbf{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\mathbf{w} = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$ span $\mathbb{R}^2$, so vector $\mathbf{u} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$ is a linear combination of them.

Because $\mathbf{v}$ and $\mathbf{w}$ are linearly independent, I claim that there's only one linear combination of them equal to $\mathbf{u}$. Find it. [*Hint: Try to turn this problem into a system of two linear equations in two unknowns, where the unknowns are the scalar coefficients in the required linear combination. Solve the system to find the coefficients*.]

b) The vectors $\mathbf{a} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, and $\mathbf{c} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ span $\mathbb{R}^2$, so vector $\mathbf{d} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ is a linear combination of them.

Because vectors $\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$ are linearly *de*pendent, I claim that infinitely many distinct linear combinations of them are equal to $\mathbf{d}$. Find four distinct linear combinations that will do the trick.

Bases and Subspaces

God All-Mighty! … why haven't I realized it…? All these years I've gone around with a – *skeleton* – inside me!

- Ray Bradbury, "Skeleton"

In the previous section, I described how any n linearly independent vectors in $\mathbb{R}^n$ determine a clean *grid*: a sort of skeleton for the whole space. Informally, one might say that the n linearly independent vectors provide a "basis" for the grid. But *grid* isn't a standard linear algebra term, so we say that the vectors constitute *a basis for* $\mathbb{R}^n$ (one of many possible bases). Even so, it is vital to maintain the idea of the grid, for it reveals how each point in $\mathbb{R}^n$ can be "reached" as a *unique* linear combination of the n basis vectors. Abstracting this last feature from our mental imagery, we obtain our formal definition of a basis for $\mathbb{R}^n$.

> **Definition.** A set of vectors is called **a basis for $\mathbb{R}^n$** if every vector in $\mathbb{R}^n$ can be *uniquely* expressed as a linear combination of vectors from the set.

You can now appreciate why we call $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ the *standard* basis vectors for $\mathbb{R}^3$. Plenty of other bases for $\mathbb{R}^3$ are available – any three linearly independent vectors will do – but the standard basis is the one we usually use unless there are compelling reasons to use a different one.

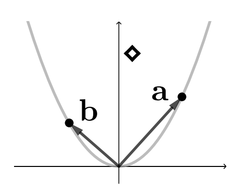
To link the idea of a basis for $\mathbb{R}^n$ to the previous section's vocabulary, you should be able to convince yourself that a set of vectors will be a basis for $\mathbb{R}^n$ if and only if two conditions hold:

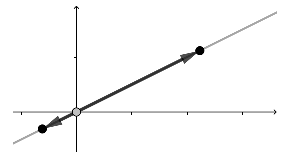
1. **The vectors must be *linearly independent*.**
   This condition ensures that the vectors determine a clean grid, so that each point in the grid has a unique *representation as a linear combination of the given vectors.*
2. **The vectors must *span* $\mathbb{R}^n$.**
   This condition ensures that the clean grid encompasses all of $\mathbb{R}^n$.

Although $\mathbb{R}^n$ is the main "vector space" in which we'll do linear algebra, we'll often find it useful to restrict our attention to lower-dimensional *subspaces* of $\mathbb{R}^n$. To define what I mean by a subspace, I'll need a term from advanced mathematics: We say that a set is **closed** *under some particular operation* if applying the operation to members of the set always yields a member of the set. For example, the set of integers is closed under multiplication (since the product of two integers is always an integer), but it is *not* closed under division (since, for example, 2/3 is not an integer). With that simple bit of terminology in hand, we can now easily define a subspace of $\mathbb{R}^n$.

> **Definition.** Any subset of $\mathbb{R}^n$ that is closed under vector addition and scalar multiplication is called a **subspace** of $\mathbb{R}^n$.

To develop a feel for this idea, let's first consider a non-example. Is the parabola at right a subspace of $\mathbb{R}^2$? Well, the vectors $\mathbf{a}$ and $\mathbf{b}$ in the figure correspond to points on the parabola, but their sum corresponds to a point – marked by a diamond – that isn't on the parabola. Since the parabola isn't closed under vector addition, it is *not* a subspace of $\mathbb{R}^2$.



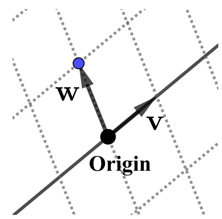


Now for an actual example, consider a line through the origin. The sum of any two vectors corresponding to points on the line (such as those in the figure) is obviously another vector on the line, so the line is closed under vector addition. It's also obvious that any such line must be closed under scalar multiplication. Consequently, any line through the origin is a subspace of $\mathbb{R}^2$.

It's not hard to catalog all of $\mathbb{R}^n$'s subspaces.

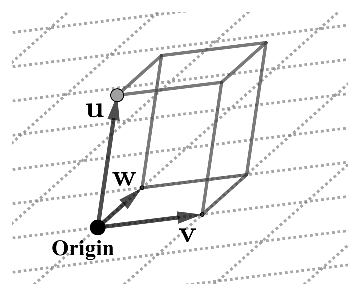
We'll start small: The origin itself is a subspace, sometimes called *the trivial subspace*. (It is closed under addition since $\mathbf{0}+\mathbf{0}=\mathbf{0}$. It is closed under scalar multiplication since $c\mathbf{0}=\mathbf{0}$ for any scalar c.)

Next, if a subspace contains a point other than the origin, then it contains the whole *line* through the origin and the point. [Proof: The subspace contains the vector corresponding to the given point, so by closure under scalar multiplication, all scalar multiples of that vector are in the subspace too. But the set of all these scalar multiples is, of course, just the line passing through the origin and the given point.]



Let's continue. If a subspace contains a line through the origin and a point not on that line, then it must contain the whole *plane* in which the line and point both lie. To see why, let $\mathbf{v}$ be any nonzero vector corresponding to a point on the line. Let $\mathbf{w}$ be the vector corresponding to the given point. Both vectors belong to our subspace. These vectors are clearly linearly independent, so they generate a grid that covers the plane in which the line and point lie. It follows that *every* point in that plane is a linear combination of $\mathbf{v}$ and $\mathbf{w}$. And thus, since subspaces are closed under vector addition and scalar multiplication (or equivalently, under linear combinations), every one of the plane's points must be in the subspace, as claimed. (Note that the plane includes the origin, since it contains a line passing through it.)

You can probably anticipate the next claim: If a subspace contains both a plane through the origin and a point not on that plane, then it must contain the whole *three-dimensional hyperplane* in which the plane and point both lie. The proof is essentially the same as that in the previous paragraph. Let $\mathbf{v}$ and $\mathbf{w}$ be any two nonzero linearly independent vectors corresponding to points in the given plane. Together with $\mathbf{u}$, the vector corresponding to the given point, they constitute a set of three linearly independent vectors in our subspace. These three vectors generate a grid covering the unique three-dimensional hyperplane in which they lie. (At right I've drawn one "cell" of this grid poking up from the given plane. You'll have to imagine the rest.) Thus, every point in the three-dimensional hyperplane is a linear combination of these three vectors, which, in turn, means – since subspaces are closed under linear combinations – that the entire three-dimensional hyperplane must be in the subspace, as claimed. (And note that this three-dimensional hyperplane – like every subspace – contains the origin.)

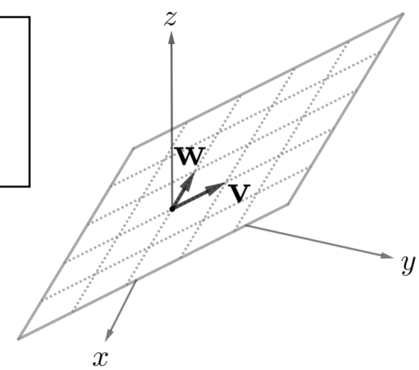


Step by step, the argument continues in the same way, one dimension at a time, and the conclusion is always the same: If a subspace of $\mathbb{R}^n$ contains a k-dimensional hyperplane through the origin *and* a point outside of that hyperplane, then the subspace contains the entire $(k+1)$-dimensional hyperplane containing them.

Our conclusion: $\mathbb{R}^n$'s nontrivial subspaces are precisely the lines, planes, and hyperplanes that pass through the origin. In Chapter 1, Exercise 12, you learned that these geometric objects are described by linear equations. Lines, planes, and hyperplanes will be our constant linear algebraic companions, so the simplicity of their equations is a blessing.

We can think of each subspace as a complete "vector space" in its own right. Accordingly, it makes perfect sense to speak of a *basis for a subspace*.

> **Definition.** A set of vectors is called **a basis for a subspace of $\mathbb{R}^n$** if each vector in the subspace has a *unique* expression as a linear combination of the set's vectors.



For example, the figure at right depicts $\mathbb{R}^3$ and one of its many two-dimensional subspaces, a plane through the origin. The two vectors $\mathbf{v}$ and $\mathbf{w}$ in the figure clearly constitute a basis for this subspace.

Thanks to our work above, we can easily describe all of $\mathbb{R}^3$'s subspaces. It has a *zero*-dimensional subspace (the trivial subspace $\{\mathbf{0}\}$), infinitely many *one*-dimensional subspaces (lines through the origin), infinitely many *two*-dimensional subspaces (planes through the origin), and a single *three*-dimensional subspace: $\mathbb{R}^3$ itself. (For technical reasons, we take the "sub" in "subspace" in the spirit of "less than *or equal to*", so every vector space is considered a subspace of itself.)

Exercises.

7. True or false (and explain your answer):

a) Any two vectors in the plane constitute a basis for $\mathbb{R}^2$.
b) Any two nonzero vectors in the plane constitute a basis for $\mathbb{R}^2$.
c) Any two linearly independent vectors in the plane constitute a basis for $\mathbb{R}^2$.
d) Any line through the origin is a subspace of $\mathbb{R}^2$.
e) The only subspaces of $\mathbb{R}^2$ are lines through the origin.
f) Any set of linearly independent vectors in $\mathbb{R}^n$ constitutes a basis for $\mathbb{R}^n$.
g) Any spanning set of linearly independent vectors in $\mathbb{R}^n$ constitutes a basis for $\mathbb{R}^n$. (A "spanning set" of vectors is one that spans the entire space.)
h) It is possible to have three linearly independent vectors in a 2-dimensional subspace of $\mathbb{R}^3$.
i) If a set of vectors *spans* a subspace, these vectors necessarily constitute a *basis* for the subspace.
j) If we have a set of vectors that *span* a subspace, but which are linearly **de**pendent, then we can obtain a basis for the subspace by throwing away some of the vectors in the spanning set.
k) If we have a set of linearly independent vectors in a subspace, these vectors necessarily constitute a basis for the subspace.
l) If in a subspace we have a set of linearly independent vectors that do *not* span the whole subspace, we can always obtain a basis for the subspace by adding some new well-chosen vectors to our set.
m) Every subspace of $\mathbb{R}^n$ contains the zero vector.

8. Describe the subspaces of the following spaces: a) $\mathbb{R}^2$ b) $\mathbb{R}^3$ c) $\mathbb{R}^4$ d) $\mathbb{R}^5$.

9. Thinking geometrically, decide whether each of the following vector pairs constitutes a basis for $\mathbb{R}^2$.

a) $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ b) $\begin{pmatrix}5\\0\end{pmatrix}, \begin{pmatrix}0\\-1\end{pmatrix}$ c) $\begin{pmatrix}3\\1\end{pmatrix}, \begin{pmatrix}-3\\1\end{pmatrix}$ d) $\begin{pmatrix}1\\4\end{pmatrix}, \begin{pmatrix}-2\\-8\end{pmatrix}$

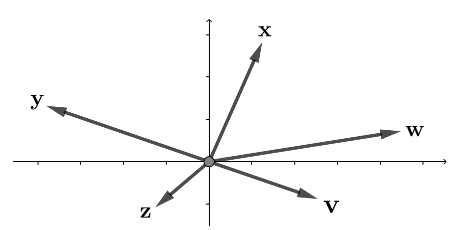
10. Let $\mathbf{b_1} = \begin{pmatrix}1\\5\\0\end{pmatrix}$, $\mathbf{b_2} = \begin{pmatrix}2\\3\\0\end{pmatrix}$, $\mathbf{b_3} = \begin{pmatrix}-2\\3\\1\end{pmatrix}$.

a) These three vectors constitute a basis for $\mathbb{R}^3$. Explain why this is geometrically obvious.
[**Hint:** *where in* $\mathbb{R}^3$ *do the first two vectors lie*?]
b) Give an example of a nonzero vector $\mathbf{c}$ such that $\mathbf{b_1}, \mathbf{b_2}$, and $\mathbf{c}$ do *not* constitute a basis for $\mathbb{R}^3$.

11. We defined a subspace as being closed under vector addition and scalar multiplication. Are both conditions necessary? Or does every set that satisfies one condition automatically satisfy the other? Prove your claim.

12. Recall that a ratio of polynomials is called a *rational function*. Is the set of all rational functions closed under addition? Is it closed under multiplication? Is it closed under differentiation? Is it closed under integration?

13. The figure at right shows five vectors in $\mathbb{R}^2$.
Which of the following sets constitute a basis for $\mathbb{R}^2$?
Justify your answers.

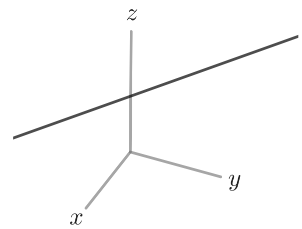


a) $\{\mathbf{v}, \mathbf{w}\}$ b) $\{\mathbf{v}, \mathbf{x}\}$ c) $\{\mathbf{v}, \mathbf{y}\}$
d) $\{\mathbf{v}, \mathbf{z}\}$ e) $\{\mathbf{z}, \mathbf{w}\}$ f) $\{\mathbf{v}, \mathbf{w}, \mathbf{x}\}$
g) $\{\mathbf{v}, \mathbf{x}, \mathbf{y}\}$ h) $\{\mathbf{v}\}$ i) $\{\mathbf{v}, \mathbf{w}, \mathbf{x}, \mathbf{y}\}$

14. For each of the following equations, state whether the graph is a subspace of $\mathbb{R}^3$. If not, explain why it isn't. If so, give a basis for the subspace, and explain how you know it is a subspace.

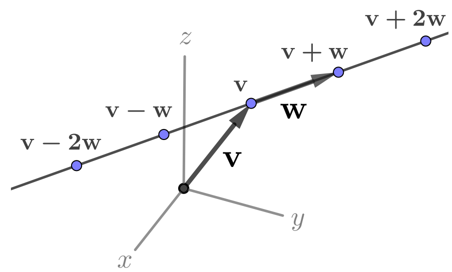
a) $z = 0$ b) $z = 1$ c) $x + y + z = 0$ d) $x + y + z = 1$
e) $z = x^2 + y^2$ f) $z = 2x - 3y$ g) $x^2 + y^2 + z^2 = 1$ h) $x^2 + y^2 + z^2 = 0$

Parametric Representations of Subspaces (and Affine Spaces)



A line, plane, or hyperplane that does *not* contain the origin (such as the line at right) is called an **affine space**. These are not proper vector spaces since they aren't closed under vector addition or scalar multiplication. They do, however, play important supporting roles in linear algebra. In this section, we'll discuss how to represent both subspaces and affine spaces in terms of parametric linear equations.

We'll begin with two examples in $\mathbb{R}^3$, starting with the line in the figure above. Pick any of the line's points and let $\mathbf{v}$ be its corresponding vector. Let $\mathbf{w}$ be any vector parallel to the line. The figure at right makes it clear that each of the line's points corresponds to a vector $\mathbf{v} + t\mathbf{w}$ for some scalar t.* Accordingly, the compact expression



$$\mathbf{v} + t\mathbf{w}$$

captures the full line as t ranges over the reals. For any specific line, we can expand this abstract representation by substituting in the specific components of $\mathbf{v}$ and $\mathbf{w}$.

For example, let's consider the line through the point $(1, 2, 3)$ and parallel to the vector $\mathbf{i} + 3\mathbf{j} + \mathbf{k}$. By the ideas above, we may represent this line as

$$\begin{pmatrix}1\\2\\3\end{pmatrix} + t\begin{pmatrix}1\\3\\1\end{pmatrix}. \text{ Or, after doing the vector addition, } \begin{pmatrix}1+t\\2+3t\\3+t\end{pmatrix}.$$

The components of this last vector can be read as individual expressions for the x, y, and z coordinates of a moving point that traverses the line as t ranges over the real numbers. That is, we can describe the line by the following set of parametric equations:

$$x = 1 + t$$
$$y = 2 + 3t$$
$$z = 3 + t$$

Using these equations, we can answer simple questions about the line.†

We'll often reverse this process. Given a set of parametric equations, we'll bundle them into vectors, which will let us recognize that they describe a line in space. For example, the three parametric equations

$$x = 2 + 3t$$
$$y = 1 - 5t$$
$$z = 9 + 2t$$

can be repackaged as a single vector equation:

* I've drawn the cases in which t is an integer between -2 and 2, but t could be any real number whatsoever. The point halfway along the $\mathbf{w}$ vector, for example, corresponds to $\mathbf{v} + (1/2)\mathbf{w}$.

† **Example**: Where does the line cross the xy-plane? Well, this occurs when $z = 0$. By our z-equation, *this* happens when $t = -3$. But when $t = -3$, the corresponding x and y values are -2 and -7. Thus, the line crosses the plane at $(-2, -7, 0)$.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2+3t \\ 1-5t \\ 9+2t \end{pmatrix}.$$

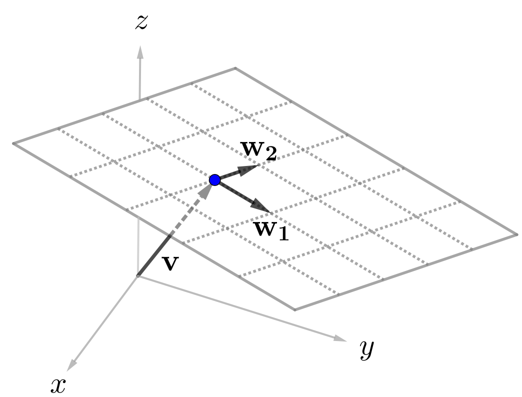
We can then distill the right-hand side into a linear combination of two vectors:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 9 \end{pmatrix} + t\begin{pmatrix} 3 \\ -5 \\ 2 \end{pmatrix},$$

which we should now be able to recognize as a vector representation of a line in $\mathbb{R}^3$. Namely, the line that passes through point $(2,1,9)$ and is parallel to the vector $3\mathbf{i} - 5\mathbf{j} + 2\mathbf{k}$.

So much for lines. Let's turn our attention to planes.

Pick any point of a plane and let $\mathbf{v}$ be its corresponding vector. Let $\mathbf{w_1}$ and $\mathbf{w_2}$ be any two linearly independent vectors parallel to the plane. Because they are linearly independent, these two vectors determine a clean grid on the plane. The figure at right makes it clear that each point on the plane corresponds to a vector of the form



$$\mathbf{v} + t\mathbf{w_1} + s\mathbf{w_2}$$

for scalars t and s. Accordingly, this expression captures the full plane as t and s range over the real numbers. This representation of a plane can be expanded, in any concrete instance, by substituting in the vectors' coordinates.

For example, consider the plane through the point $(1, 1, 3)$ containing the (suitably translated) vectors $\mathbf{i} + 2\mathbf{j} - \mathbf{k}$ and $-\mathbf{i} + 2\mathbf{j}$. By the ideas above, we may represent this line as

$$\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} + t\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + s\begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}. \quad \text{Or, after doing the vector addition,} \begin{pmatrix} 1+t-s \\ 1+2t+2s \\ 3-t \end{pmatrix}.$$

The components of this last vector can be read as individual expressions for the x, y, and z coordinates of a moving point that traverses the plane as t and s range over the real numbers. That is, we can describe the plane by the following set of parametric equations:

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

As we saw earlier, we can reverse the process, distilling from parametric linear equations a vector expression whose geometric nature we can then discern. For example, the parametric equations

$$\begin{aligned} x &= 3 + 5t - 3s \\ y &= 2 + 8t + s \\ z &= 5 - t + 4s \end{aligned}$$

can be repackaged as a single vector equation:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 + 5t - 3s \\ 2 + 8t + s \\ 5 - t + 4s \end{pmatrix},$$

whose right-hand side can then be decomposed into a linear combination of three vectors:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} + t\begin{pmatrix} 5 \\ 8 \\ -1 \end{pmatrix} + s\begin{pmatrix} -3 \\ 1 \\ 4 \end{pmatrix}.$$

Since the two rightmost vectors are linearly independent, we should now be able to recognize the preceding expression as a vector representation of a plane in $\mathbb{R}^3$. Namely, the plane that passes through point $(3, 2, 5)$ and is parallel to both $5\mathbf{i} + 8\mathbf{j} - \mathbf{k}$ and $-3\mathbf{i} + \mathbf{j} + 4\mathbf{k}$.

All that we've just done with lines and planes in $\mathbb{R}^3$ we can do just as easily with lines, planes, and hyperplanes in $\mathbb{R}^n$. And if you've grasped the geometry of the examples in $\mathbb{R}^3$, your mind's eye should (almost!) be able to see the analogous cases in $\mathbb{R}^n$, even if you can't draw pictures of them on paper. Here's an example.

Consider the following set of five parametric equations:

$$\begin{aligned} x_1 &= 2 + 3t_1 + 2t_2 + 4t_3 \\ x_2 &= 8 \ \ - t_1 + \ \ t_2 + 3t_3 \\ x_3 &= \qquad t_1 \qquad + 4t_3 \\ x_4 &= 9 \qquad\ - 7t_2 \\ x_5 &= 2 \qquad\qquad\quad + t_3 \end{aligned}$$

Here we have five variables, x_1, x_2, x_3, x_4, and x_5, whose values depend on three parameters: t_1, t_2, t_3. As these range over the real numbers, the point $(x_1, x_2, x_3, x_4, x_5)$ sweeps out some sort of geometric object in $\mathbb{R}^5$ whose nature will be clearer once we've rewritten the parametric equations in vector form:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \underbrace{\begin{pmatrix} 2 \\ 8 \\ 0 \\ 9 \\ 2 \end{pmatrix}}_{\substack{\text{point} \\ \text{in } \mathbb{R}^5}} + \underbrace{t_1\begin{pmatrix} 3 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2\begin{pmatrix} 2 \\ 1 \\ 0 \\ -7 \\ 0 \end{pmatrix} + t_3\begin{pmatrix} 4 \\ 3 \\ 4 \\ 0 \\ 1 \end{pmatrix}}_{\substack{\text{three-dimensional} \\ \text{hyperplane in } \mathbb{R}^5}}.$$

We can't draw pictures in $\mathbb{R}^5$, but our two preceding figures give us a good sense of what's going on here. The first vector on the right-hand side above corresponds to a "base point" in $\mathbb{R}^5$. The next three vectors are linearly independent (verify this), and thus span a three-dimensional hyperplane. It follows that as the parameters t_1, t_2, t_3 range over the real numbers, our moving point $(x_1, x_2, x_3, x_4, x_5)$ sweeps out the three-dimensional hyperplane that passes through the point $(2, 8, 0, 9, 2)$ and is parallel to the span of the three rightmost vectors in the expression above.

Exercises.

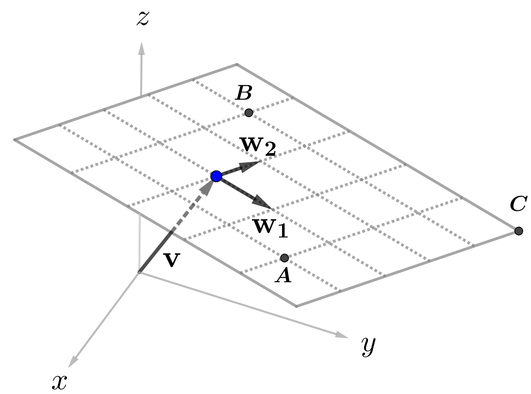
15. In the preceding section, the figure at right appears (without points A, B, and C) along with these words:

"The figure at right makes it clear that each point on the plane corresponds to a vector of the form

$$\mathbf{v} + t\mathbf{w_1} + s\mathbf{w_2}$$

for scalars t and s."

To ensure that this truly *is* clear to you, express the vectors corresponding to the three points A, B, and C in the form specified above.



16. Describe the figure represented by each set of parametric linear equations. Sketch it when possible, and decide whether each is a subspace or an affine space:

a) $x = 1 + 2t$, $y = 2 + 2t$

b) $x = 1 + 2t$, $y = 3t$, $z = 2 - t$

c) $x = 2 + t$, $y = 3 - t$, $z = t$, $w = 6 + 2t$

d) $x = t + 2s$, $y = -t + 3s$, $z = 2t - s$

e) $x = 1 + t - 2s$, $y = 2 - 2t$, $z = 3 + 2s$

f) $x = t_1 + t_2 + t_3$, $y = 2t_1 - t_2 + t_3$, $z = t_1 + t_2 + 2t_3$, $w = 3t_3$

g) $x_1 = 1 + t_1 + 3t_2$, $x_2 = 2t_2$, $x_3 = 2 + 3t_2$, $x_4 = 5 - 7t_1 + 3t_2$, $x_5 = 8 - t_1 + 5t_2$

h) $x = 2t + 4s$, $y = t + 2s$, $z = 5t + 10s$

i) $x = 1 + t_1 + t_3$, $y = 2 + t_2$, $z = 3 + t_1 + t_3$

17. Parametric representations of geometric objects are not unique. Thus, for example, we can represent the line in the plane whose equation is $2x + 3y = 6$ by many different sets of parametric equations. Explain why this is so, and give three different examples of parametric equations that represent this line.

18. Find parametric equations for…

a) the line in $\mathbb{R}^2$ through the point $(3, 1)$ in the direction of the vector $(4, -1)$.
b) the line in $\mathbb{R}^3$ through the point $(3, 1, 2)$ in the direction of the vector $(4, -1, 1)$.
c) the plane in $\mathbb{R}^3$ through points $(3, 1, 2)$, $(4, -1, 1)$, and $(-2, 3, 1)$.

19. Give parametric equations for the following subspaces and describe them geometrically.

a) the span of $\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 6 \\ 1 \\ 0 \end{pmatrix}$ b) the span of $\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 6 \\ 3 \\ 3 \end{pmatrix}$

20. A line is a one-dimensional object; two points are needed to determine a line. A plane is a two-dimensional object; three points are needed to determine a plane.

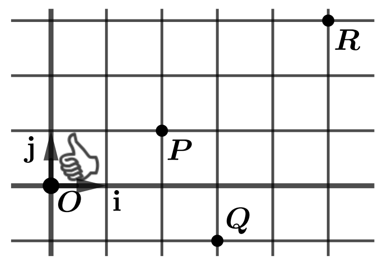
a) How many points are needed to determine a 3-dimensional hyperplane in $\mathbb{R}^n$ (for $n \geq 4$)?
b) How many points are needed to determine an m-dimensional hyperplane in $\mathbb{R}^n$?
c) Will any three points in space determine a unique plane, or are there exceptions? If there are exceptions, name three points that do *not* determine a unique plane.

Chapter 3
Linear Transformations and Matrices

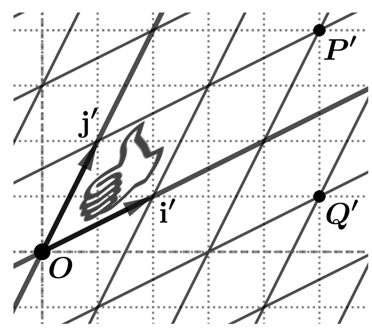
Linear Maps and Their Matrices

As Gregor Samsa awoke one morning from uneasy dreams, he found himself transformed in his bed into a gigantic insect.
- Franz Kafka, "The Metamorphosis"

Calculus focuses on real-valued functions of real numbers; we visualize these functions as static graphs. Linear algebra focuses on vector-valued functions of vectors; we visualize these functions dynamically, as distortions of space itself. More specifically, linear algebra's functions are "grid transformations": they work by dragging the tips of some or all of the standard basis vectors to new locations, thus transforming the grid they generate. These special functions are called **linear transformations** (or linear *maps*).



For example, the graph at right shows the standard graph-paper grid generated by $\mathbf{i}$ and $\mathbf{j}$. (Point O is the origin. Points P, Q, R and the enthusiastic right hand are all just stage props that will help us illustrate the effect of a linear map in a few moments.) Now let's go ahead and apply our first linear transformation: We will drag the tips of $\mathbf{i}$ and $\mathbf{j}$ to points $(2, 1)$ and $(1, 2)$ respectively. When we transform $\mathbf{i}$ and $\mathbf{j}$ this way, we transform the entire grid that they generate. The result is shown below, in the second figure.



The original grid of squares (displayed now as a ghostly background) has been replaced by a grid of parallelograms generated by $\mathbf{i}'$ and $\mathbf{j}'$, the names I'll use for the transformed images of $\mathbf{i}$ and $\mathbf{j}$. Clearly, this linear transformation will distort any figure lying in the plane, such as the hand. But one crucial feature remains constant: Even though the linear map moves most of the points in the plane, it preserves their coordinates *relative to the grid*. For example, point P (first figure) corresponds to vector $2\mathbf{i} + \mathbf{j}$; its image, point P' (second figure) corresponds to vector $2\mathbf{i}' + \mathbf{j}'$. In both cases, we get to the point by following precisely the same "marching orders": Start from the origin, take two steps forward in the direction of the first "axis", and then take one step forward in the direction of the second "axis". We usually describe this by saying that *linear transformations preserve linear combinations of vectors*.

You should verify on the pictures above that the same phenomenon happens with points Q and Q'. Finally, although point R's image, R', is "offscreen", we know where it must lie: In the first figure, R corresponds to $5\mathbf{i} + 3\mathbf{j}$, so in the second, R' must "by linearity" correspond to $5\mathbf{i}' + 3\mathbf{j}'$. It would thus lie somewhere just beyond the right edge of this page.

Every linear transformation – not just the one above – preserves linear combinations of vectors. This is their defining feature, the very thing that makes linear transformations special.

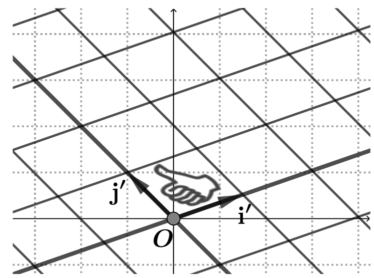
To summarize, a linear map's geometric essence is a grid transformation induced by dragging the tips of the standard basis vectors to new places. A linear map's algebraic essence, which is a direct reflection of its geometric essence, is the preservation of linear combinations: Every vector is a certain linear combination of the standard basic vectors; when subjected to a linear map, its transformed image will be precisely that same linear combination of the *transformed* basis vectors. It follows that if we know how a linear map transforms the standard basis vectors, we know how it transforms *every* vector.

This means we can describe any linear transformation quite compactly. Rather than trying to illustrate its effect with figures (tedious in $\mathbb{R}^2$, difficult in $\mathbb{R}^3$, impossible in higher dimensions), we need only specify where the transformation sends the standard basis vectors. We encode this core information in a *matrix* whose i^{th} column simply records the components of the i^{th} standard basis vector's transformed image. We call the resulting matrix **the linear map's matrix** (relative to the standard basis).

Example 1. According to our definition above, the matrix

$$\begin{pmatrix} 1.5 & -1 \\ 0.5 & 1 \end{pmatrix}$$

represents the linear map that sends $\mathbf{i}$ to $\mathbf{i}' = 1.5\mathbf{i} + 0.5\mathbf{j}$ and sends $\mathbf{j}$ to $\mathbf{j}' = -\mathbf{i} + \mathbf{j}.$ I've drawn the transformation at right. With or without this visual aid, we can easily determine where this linear transformation will send any given vector.



For instance, where will it send $2\mathbf{i} + 3\mathbf{j}$? Well, even without the visual aid, we know that linear maps preserve linear combinations, so this vector will be sent to $2\mathbf{i}' + 3\mathbf{j}'$, which is

$$2\begin{pmatrix} 1.5 \\ 0.5 \end{pmatrix} + 3\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

You can also confirm this by looking at the figure: Starting from the origin, two "$\mathbf{i}'$-steps" followed by three "$\mathbf{j}'$-steps" brings us to $(0, 4)$ on the original square grid. ♦

Mathematicians have invented a powerful algebraic shortcut: Instead of writing out the sentence

"The linear transformation whose matrix is $\begin{pmatrix} 1.5 & -1 \\ 0.5 & 1 \end{pmatrix}$ sends $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ to $\begin{pmatrix} 0 \\ 4 \end{pmatrix}$",

we express the same idea more concisely as an equation:

$$\begin{pmatrix} 1.5 & -1 \\ 0.5 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

We refer to the operation on the left side of this equation as **matrix-vector multiplication**. However, this "multiplication" should really make you think of *function evaluation*. This matrix (or better yet, the linear map that it represents) takes $2\mathbf{i} + 3\mathbf{j}$ as its input, and produces $4\mathbf{j}$ as its output. Once you have internalized the idea that matrix-vector multiplication simply indicates where a given linear map sends a given vector, it becomes very easy to reduce the process to a few mindless turns of an algebraic crank.

Consider a perfectly general expression of matrix-vector multiplication:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

What will the product be? Well, by definition, this matrix represents the linear map that sends $\mathbf{i}$ and $\mathbf{j}$ to

$$\mathbf{i}' = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{and} \quad \mathbf{j}' = \begin{pmatrix} c \\ d \end{pmatrix}.$$

Because linear transformations preserve linear combinations, it follows that our generic linear map must send the generic vector $v_1\mathbf{i} + v_2\mathbf{j}$ to the vector $v_1\mathbf{i}' + v_2\mathbf{j}'$, which is

$$v_1\begin{pmatrix} a \\ b \end{pmatrix} + v_2\begin{pmatrix} c \\ d \end{pmatrix}.$$

And thus we have our answer:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = v_1\begin{pmatrix} a \\ b \end{pmatrix} + v_2\begin{pmatrix} c \\ d \end{pmatrix}.$$

This is a crucial result. Let us rephrase it and sanctify it in a box:

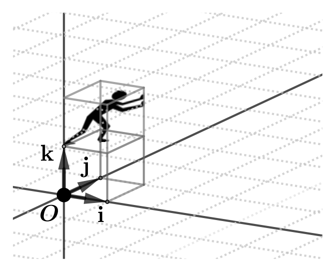
> A **matrix-vector product** is a *weighted sum of the matrix's columns*, where the i^{th} column's weight is the vector's i^{th} entry.

Example 2. $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ -1 \end{pmatrix} = 3\begin{pmatrix} 2 \\ 1 \end{pmatrix} - 1\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} \mathbf{5} \\ \mathbf{1} \end{pmatrix}$.

As you should verify, the matrix here represents the linear map whose "before and after" pictures appear on the first page of this section. Moreover, as you should also verify, this example's input and output vectors correspond to points Q and Q' on those figures. ♦

Every matrix-vector multiplication can be understood as mapping one vector to another via a linear transformation, a geometric operation whose effect is entirely determined by the columns of the matrix. This is linear algebra's central geometric idea, comparable in importance to calculus's key geometric ideas: all derivatives can be understood as slopes, and all definite integrals can be understood in terms of area. This isn't the only way to think of matrix-vector multiplication, but it is the fundamental one to which we'll return time and time again. It follows that this section may be the most important one in the entire book. To reinforce its ideas, let's consider a few more examples.

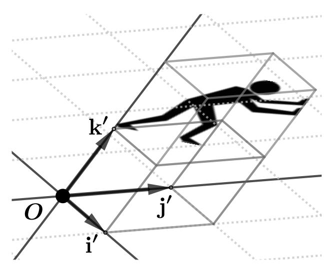
Example 3. Although grids and linear transformations are difficult to draw in $\mathbb{R}^3$, they are not hard to imagine. The standard basis vectors generate a grid of cubes (as the figure at right suggests), and we can induce a linear map by dragging the tips of some or all of the standard basis vectors to new locations, as I've done in the figure below, where the cubes have become parallelepipeds.



If I tell you that the matrix of this transformation is

$$\begin{pmatrix} 1.5 & 2 & 1 \\ 0 & 1 & -0.5 \\ -0.5 & 0 & 2 \end{pmatrix},$$

then you can easily determine where it sends any point/vector in $\mathbb{R}^3$. The first column of the matrix, for example, tells us that $\mathbf{i}$ is scaled a bit and pushed under the xy-plane to end up at $\mathbf{i}' = 1.5\mathbf{i} - 0.5\mathbf{k}$. Similarly, we can read the fates of $\mathbf{j}$ and $\mathbf{k}$ directly from the matrix.

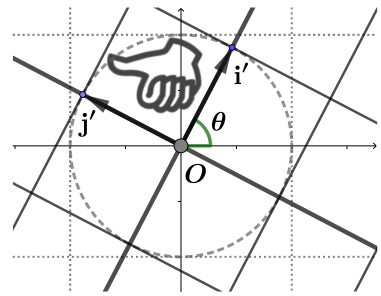


Now consider any old vector: say, $\mathbf{4i} - 3\mathbf{j} + \mathbf{6k}$. Where will our linear map send it? Following our matrix-vector multiplication recipe (in the box above), we see that the map will send the vector to

$$\begin{pmatrix} 1.5 & 2 & 1 \\ 0 & 1 & -0.5 \\ -0.5 & 0 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ -3 \\ 6 \end{pmatrix} = 4\begin{pmatrix} 1.5 \\ 0 \\ -0.5 \end{pmatrix} - 3\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + 6\begin{pmatrix} 1 \\ -0.5 \\ 2 \end{pmatrix} = \begin{pmatrix} \mathbf{6} \\ \mathbf{-6} \\ \mathbf{10} \end{pmatrix}. \quad ♦$$

In $\mathbb{R}^4$ or still higher-dimensional spaces, linear maps may be impossible to visualize, and they are certainly impossible to draw, but writing down matrices that represent them is easy. You'll have the chance to do that in the exercises, but for now, let us return to $\mathbb{R}^2$ for an eminently familiar transformation.

Example 4. Rotating the plane counterclockwise around the origin through an angle of θ is a transformation, but is it a *linear* transformation? If so, how can we represent it as a matrix?



Solution. We've defined a linear map as a grid transformation induced by dragging the tips of standard basis vectors to new places. A rotation about the origin clearly qualifies, because rotating $\mathbf{i}$ and $\mathbf{j}$ through θ induces a rotation of the whole grid.

Since a rotation is a linear map, we can represent it as a matrix. As discussed above, the columns of this matrix will be the rotated images of $\mathbf{i}$ and $\mathbf{j}$, expressed as column vectors.

The dashed circle in the figure is the *unit* circle, and by the unit-circle definitions of the sine and cosine functions, we can see that $\mathbf{i}' = (\cos\theta)\,\mathbf{i} + (\sin\theta)\,\mathbf{j}$. Those same definitions yield $\mathbf{j}' = [\cos(\theta + 90°)]\mathbf{i} + [\sin(\theta + 90°)]\mathbf{j} = (-\sin\theta)\,\mathbf{i} + (\cos\theta)\,\mathbf{j}$. It follows that the rotation matrix is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

For example, if we wanted to know where the point $(1.5, 0.2)$ ends up after rotating it $22°$ counterclockwise around the origin, we could now find the answer very easily. The point will end up at the point corresponding to the vector

$$\begin{pmatrix} \cos 22° & -\sin 22° \\ \sin 22° & \cos 22° \end{pmatrix}\begin{pmatrix} 1.5 \\ 0.2 \end{pmatrix} = 1.5\begin{pmatrix} \cos 22° \\ \sin 22° \end{pmatrix} + 0.2\begin{pmatrix} -\sin 22° \\ \cos 22° \end{pmatrix} \approx \begin{pmatrix} 1.32 \\ 0.63 \end{pmatrix}.$$

In other words, after the rotation, point $(1.5, 0.2)$ ends up at approximately $(\mathbf{1.32}, \mathbf{0.63})$.

Similarly, if we wished to rotate, say, $(0.5, -0.7)$ *clockwise* around the origin by $137°$, the same type of calculation will work if we let $\theta = -137°$ (Note that negative!):

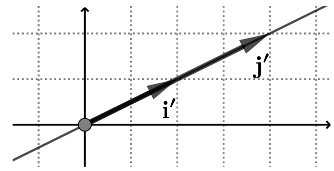
$$\begin{pmatrix} \cos(-137°) & -\sin(-137°) \\ \sin(-137°) & \cos(-137°) \end{pmatrix}\begin{pmatrix} 0.5 \\ -0.7 \end{pmatrix} = 0.5\begin{pmatrix} \cos(-137°) \\ \sin(-137°) \end{pmatrix} - 0.7\begin{pmatrix} -\sin(-137°) \\ \cos(-137°) \end{pmatrix} \approx \begin{pmatrix} -0.84 \\ 0.17 \end{pmatrix}.$$

Thus, after the rotation, point $(0.5, -0.7)$ ends up at approximately $(-\mathbf{0.84}, \mathbf{0.17})$. ♦

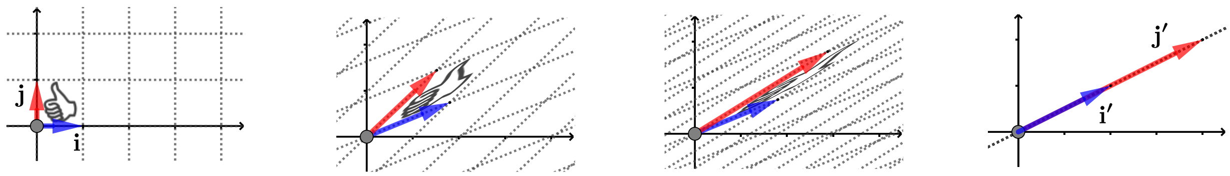
So far, we've considered *local* changes wrought by linear maps. But what do such maps achieve *globally*? The linear maps of $\mathbb{R}^2$ that we've met so far have moved points around the plane, but the global output of each such map was always… the plane. The analogous story held in our one example of a map of $\mathbb{R}^3$. This raises a question: Must *every* linear map's global output simply be the space on which it is defined? No. This has happened so far only because each map's matrix has had *linearly independent* columns, which ensured that the standard basis vectors were just mapped onto a new *basis* for the same old space. But if a map were to have a matrix with linearly *de*pendent columns, then the clean standard grid with which we started would be mapped onto a tangled grid, at least one of whose generating vectors lies in the span of the others, with the obvious result that at least one dimension of the global output would collapse.

Example 5. Discuss the linear map whose matrix is $\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$.

Solution. The columns of this map's matrix are linearly *de*pendent: The transformed images of $\mathbf{i}$ and $\mathbf{j}$ lie along the same line. It follows that the columns' span *is* that line. The linear map crushes the standard two-dimensional grid of squares down into the line, and thus, although the linear map is defined on $\mathbb{R}^2$, its global output is nothing but this lone line, a *one*-dimensional subspace of $\mathbb{R}^2$. ♦

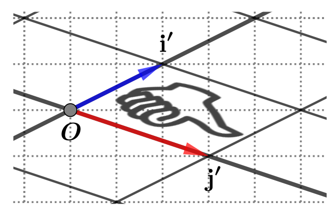


It helps to think of a linear map as a continuous transformation, with the standard basis vectors (and grid) morphing into their images. For instance, in the previous example, we might imagine $\mathbf{i}$ and $\mathbf{j}$ lengthening and approaching one another (as in the sequence of pictures below), with both coming to rest on the line $y = x/2$, at which point all of two-dimensional space has folded up, fanlike, into a line.



But suppose one of the vectors *didn't* stop at the line, but crossed over it. Then what would happen?

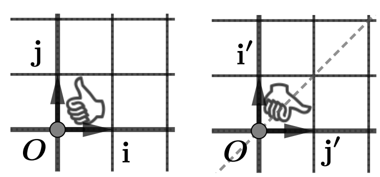
The figure at right shows such a situation. I've dragged $\mathbf{j}'$ past $\mathbf{i}'$, bringing its tip to rest at point $(3, -1)$. Vectors $\mathbf{i}'$ and $\mathbf{j}'$ are now linearly independent, so the map's global output is two-dimensional, but with a notable change: From its "folded up fan" state, the grid has now re-emerged "flipped over". We've gone through the looking glass, reversing the orientations of figures in the plane. In our original drawing, for example, the hand was a *right* hand, but it has now been transformed into a *left* hand, just as your own right hand becomes a left hand when you view it in a mirror. I'll say more about orientation-reversing maps in Chapter 5. For now, I'll note that the simplest orientation-reversing maps are ordinary reflections. Reflections are some of the most basic (and important) of all linear transformations.



Example 6. Find the matrix for reflection across the line $y = x$.

Solution. The before-and-after figures at right give us all we need: The tips of $\mathbf{i}$ and $\mathbf{j}$ are mapped, respectively, to points $(0,1)$ and $(1,0)$. Accordingly, the reflection matrix will be

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$



To confirm that this is correct, we note that our reflection should send any point (a, b) to (b, a). Does our matrix actually accomplish this? Yes it does:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 0 \\ 1 \end{pmatrix} + b\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} b \\ a \end{pmatrix}.$$ ♦

Exercises.

1. True or false. Explain your answers.

a) Every linear transformation fixes the origin (i.e. maps the origin to itself).
b) Every linear transformation moves all points other than the origin to new locations.

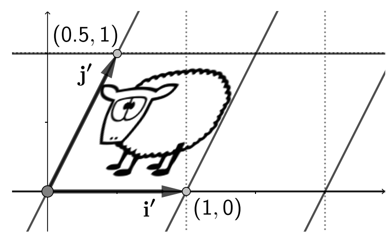
2. For each transformation of $\mathbb{R}^2$, find the associated matrix, and use it to determine where the map sends $(2, 3)$.

a) Reflection over $x = 0$ b) Reflection over $y = 0$ c) What about reflection over $y = 1$? (Careful.)
d) Rotation by θ counterclockwise about the origin.
[Don't just copy it down from the book or memorize it. Be sure you can explain where it comes from.]
e) Rotation by $30°$ counterclockwise about the origin. f) Rotation by $45°$ clockwise about the origin.
g) The "do nothing" map, which leaves the plane as it is. This map's matrix is called **the identity matrix**.

3. The linear maps represented by the following matrices distort the square grid generated by **i** and **j** into new forms. Sketch the new grids. On each, indicate (without doing any calculations!) the images of **i**, **j**, and $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j}$.

a) $\begin{pmatrix} 1.5 & 0 \\ 0 & 2 \end{pmatrix}$ b) $\begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}$ c) $\begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix}$

4. The linear map whose picture is shown at right is called a horizontal **shear**.



a) Find the shear's matrix.
b) To which point is $(10, 22)$ mapped by the shear?
c) Which point is mapped to $(10, 22)$ by the shear?
d) The shear transforms the original grid's squares into parallelograms. Does it change their areas in the process? Why or why not?
e) Does shearing the sheep change its area? Why or why not?

5. Some terminology: A matrix with n rows and n columns is called an $\boldsymbol{n} \times \boldsymbol{n}$ **matrix** (spoken as: "n by n matrix").

a) Give two examples of 2×2 matrices that map all of $\mathbb{R}^2$ onto the x-axis.
b) Give two examples of 2×2 matrices that map all of $\mathbb{R}^2$ onto the line $y = 2x$.
c) Give two examples of 3×3 matrices that map all of $\mathbb{R}^3$ onto the x-axis, thus crushing *two* dimensions.

6. The $n \times n$ **zero matrix** consists of nothing but zeros. What does it do geometrically?

7. Find the matrices that carry out the following transformations in $\mathbb{R}^3$:

a) Reflection across the xy-plane (i.e. the plane $z = 0$) b) Reflection over the *plane* $y = x$.
c) Rotation by θ about the z-axis, counterclockwise from the perspective of one looking down at the origin from a point on the positive z-axis.
d) The shear that fixes all points in the xy-plane but moves $(0,0,1)$ to $(0.5, 0.5, 1)$.
e) The "do nothing" transformation. [See Exercise 2g above.]

8. Find the matrix that…

a) reflects $\mathbb{R}^4$ across the 3-dimensional hyperplane $w = 0$ (if we call the four dimensions x, y, z, and w). You'll need to think about what this means in analogy with lower-dimensional cases.
b) reflects $\mathbb{R}^n$ across the $(n-1)$-dimensional hyperplane $x_i = 0$ (where the i^{th} dimension is x_i).
c) Describe the matrix that does nothing to $\mathbb{R}^n$. (**The $\boldsymbol{n} \times \boldsymbol{n}$ identity matrix**.)

9. Carry out the following matrix-vector multiplications:

a) $\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} 5 \\ 6 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 3 & 0 \end{pmatrix}\begin{pmatrix} 4 \\ -1 \\ 8 \end{pmatrix}$ c) $\begin{pmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \\ 0 & -1 & 2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix}$

Non-square Matrices

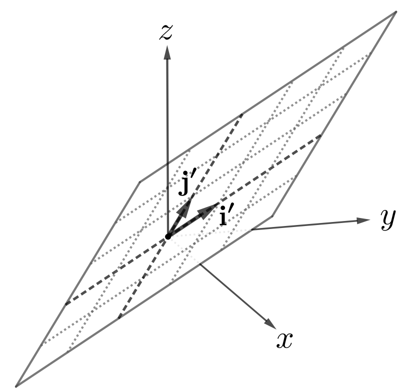
Rationalists, wearing square hats,
Think, in square rooms…
- Wallace Stevens, "Six Significant Landscapes"

So far, we've only considered linear transformations that map $\mathbb{R}^n$ into *itself*. True, we've seen that some such maps can collapse dimensions (when the matrix's columns are linearly *de*pendent), but the collapse still occurs within the ambient space of $\mathbb{R}^n$. For example, the map that orthogonally projects each vector of $\mathbb{R}^2$ onto the horizontal axis crushes all of $\mathbb{R}^2$ into a single line, but this line *isn't* $\mathbb{R}$; the projection map's output vectors are still very much vectors in $\mathbb{R}^2$. They all just happen to have a second component of zero.

Some transformations, however, map $\mathbb{R}^n$ into a different space altogether, $\mathbb{R}^m$. How does this work? By following our familiar guiding idea: If we know where the map sends $\mathbb{R}^n$'s standard basis vectors, then the fates of all the other vectors in $\mathbb{R}^n$ will be determined *by linearity* (which is a shorthand phrase for "by the preservation of linear combinations"). Maps of this sort are usually defined directly by a matrix, but unlike the matrices we've seen so far, these ones aren't square. Instead, they have n columns (one for each standard basis vector of $\mathbb{R}^n$), each of which has m entries (since each column represents a vector in the target space of $\mathbb{R}^m$). We call such a matrix an $\boldsymbol{m} \times \boldsymbol{n}$ **matrix**. Note the order: m rows, n columns.

Example 1. The 3×2 matrix

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



defines a linear transformation that sends vectors from $\mathbb{R}^2$ (this is clear because it has two columns) to $\mathbb{R}^3$ (clear because the columns have three entries). The figure at right indicates what the map/matrix does: It sends $\mathbb{R}^2$'s two standard basis vectors to $\mathbf{i}'$ and $\mathbf{j}'$, the columns of the matrix. Since these are linearly independent vectors, their span is a plane in $\mathbb{R}^3$.

Matrix-vector multiplication is still defined as the usual weighted sum of the matrix's columns. For example, where does this map send $7\mathbf{i} - 3\mathbf{j}$? An easy computation yields the answer:

$$\begin{pmatrix} 2 & -1 \\ 0 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 7 \\ -3 \end{pmatrix} = 7\begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix} - 3\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{17} \\ \mathbf{-3} \\ \mathbf{11} \end{pmatrix}. \quad ♦$$

Exercises.

10. In a matrix-vector product involving an $m \times n$ matrix, how many entries must the *vector* have?

11. An $m \times n$ matrix determines a linear transformation from where to where?

12. In the example above, a 3×2 matrix maps $\mathbb{R}^2$ onto a plane in $\mathbb{R}^3$. In similar geometric terms, explain what the following matrices do:

a) $\begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \end{pmatrix}$ c) $\begin{pmatrix} 2 & -2 & 4 \\ 3 & -3 & 6 \end{pmatrix}$ d) $\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$ e) $\begin{pmatrix} 1 & 5 & 3 \\ 0 & 2 & 9 \\ 6 & 7 & 2 \\ 0 & 0 & 4 \end{pmatrix}$ f) $\begin{pmatrix} 1 & 2 & 4 & 1 \\ 1 & 3 & 6 & 1 \end{pmatrix}$

Another Look at the Matrix-Vector Product

We've defined matrix-vector multiplication as a weighted sum of the matrix's columns and we know what it signifies: the *transformation* of the given vector by the given matrix (i.e. the linear map that it encodes). I'll now introduce a formula that lets us quickly find any particular entry in a matrix-vector product. This simple formula is useful in proofs and speeds up our hand computations of matrix-vector products.

In the proof that follows, I'll introduce some basic matrix notation that we'll often use in the future. We symbolize a matrix with an uppercase letter such as A. To indicate a particular entry in the matrix, we use the same letter, but in *lowercase*, along with two subscripts to indicate the entry's row and column. For example, $a_{1,3}$ signifies the entry in matrix A's first row and third column. (When the context allows, we sometimes minimize eye strain by omitting the comma in the subscript, writing $a_{1,3}$ simply as a_{13}.)

With this notation in hand, let's state and prove our "i^{th}-Entry Formula".

> **Matrix-Vector Multiplication ($\boldsymbol{i}^{\text{th}}$-Entry Formula).**
> If A is a matrix and $\mathbf{v}$ is a vector, we can compute $\boldsymbol{A}\mathbf{v}$'s $\boldsymbol{i}^{\text{th}}$ entry with a *dot product*:*
> $$\boldsymbol{A}\mathbf{v}\text{'s } \boldsymbol{i}^{\text{th}} \text{ entry} = (i^{\text{th}} \text{ row of } A) \cdot \mathbf{v}$$

Proof. Consider the product of a general $m \times n$ matrix A and a general column vector $\mathbf{v}$:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i,1} & a_{i,2} & \cdots & a_{i,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

By matrix-vector multiplication's definition, the product $A\mathbf{v}$ is a weighted sum of A's columns, where the weights are $\mathbf{v}$'s entries. Thinking a bit about vector addition and scalar multiplication reveals that the i^{th} entry in a weighted sum of column vectors is just… the weighted sum of the column vectors' i^{th} entries (with the same weights). It follows that the i^{th} entry of $A\mathbf{v}$ must be

$$v_1 a_{i,1} + v_2 a_{i,2} + \cdots + v_n a_{i,n} \,,$$

which is, as claimed, the dot product of A's i^{th} row with the given vector $\mathbf{v}$. ■

Example 1. In the following matrix-vector multiplication the product's 2nd entry is the dot product of the matrix's second row (in the box) and the vector:

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

Thus, it is $3(4) + 1(-1) + (-1)6 = \mathbf{5}$, as you can check by doing the full product the old way. ♦

* This is a purely formal "dot product" - mere shorthand for "sum of the products of corresponding entries in two number lists". That it works out like this is just a happy accident; it has no connection to the actual geometric definition of the dot product.

When doing matrix-vector multiplications by hand, it's quicker to repeatedly use this i^{th}-entry formula, mentally finding the product's entries one at a time, than it is to write out the weighted sums of columns. Make up a few matrix-vector multiplications on your own and do them both ways; you'll see what I mean. Because of this computational ease, I predict that you'll soon be using the dot product for all your matrix-vector multiplications, but please, please, please… never forget that matrix-vector multiplication is first and foremost a weighted sum of columns, an operation with clear geometric *meaning*, as we discussed at length in this chapter's first section. In contrast, this section's formula is just a happy algebraic accident.*

Because the i^{th}-entry formula reduces matrix-vector multiplication to dot products, whose properties *we've already proved*, it can often serve as a "bridge" in proofs, linking the unknown back to the known. For example, let's use it prove that matrix-vector multiplication can be distributed over vector addition.

Claim. For any matrix A and vectors $\mathbf{v}$ and $\mathbf{w}$ for which the following expressions are defined,

$$A(\mathbf{v}+\mathbf{w}) = A\mathbf{v} + A\mathbf{w}\,.$$

Proof. The expressions on both sides of the equals sign represent column vectors. To prove that these two column vectors are equal as claimed, we must show that all their corresponding entries are equal. To this end, we note that for all i, we have that

$A(\mathbf{v}+\mathbf{w})$'s i^{th} entry

$= (i^{\text{th}}$ row of $A) \cdot (\mathbf{v}+\mathbf{w})$ (by the i^{th}-entry formula)

$= (i^{\text{th}}$ row of $A) \cdot \mathbf{v} + (i^{\text{th}}$ row of $A) \cdot \mathbf{w}$ (distributing the dot over vector addition)

$= A\mathbf{v}$'s i^{th} entry $+$ $A\mathbf{w}$'s i^{th} entry (by the i^{th}-entry formula)

$= (A\mathbf{v} + A\mathbf{w})$'s i^{th} entry (by definition of column vector addition).

Since all entries of $A(\mathbf{v}+\mathbf{w})$ and $A\mathbf{v} + A\mathbf{w}$ are equal, the two vectors themselves are equal. ■

Exercises.

13. Redo each of the matrix-vector products from Exercise 9, but this time, do not write out the weighted sums of the columns. Instead, just use the i^{th}-entry formula to mentally compute the entries of the product.

14. The scalar multiple of a *matrix* is defined as you'd expect: Each entry in the matrix is multiplied by the scalar.

a) If A is any old matrix, describe the geometric relationship between the linear maps represented by A and $2A$.

b) Prove that for any matrix A and scalar c, the following holds: $A(c\mathbf{v}) = c(A\mathbf{v})$. That is, scalar multiples can be 'pulled through' matrix-vector multiplication. [*Hint*: Use the ideas from the proof of the claim above.]

c) Prove that matrix-vector multiplication preserves linear combinations: $A(c\mathbf{v} + d\mathbf{w}) = c(A\mathbf{v}) + d(A\mathbf{w})$. [*Hint*: All the hard work has been done already. No need to reinvent the wheel.]

15. If A is a 5×5 matrix, describe the vector $\mathbf{v}$ we'd need to use as an input so that the output $A\mathbf{v}$ would be…

a) A's 5th column. b) The sum of A's 2nd and 3rd columns. c) 3 times A's 4th column.

d) 3 times A's 4th column minus 5 times A's 1st column.

The problem's moral: We can use matrix-vector multiplication to form linear combinations of a matrix's columns.

* Regrettably, many linear algebra books *define* matrix-vector multiplication in terms of this shortcut formula – long before they even discuss linear transformations. As a result, students end up memorizing a strange rule without any hope of understanding why matrix-vector multiplication exists in the first place.

Matrix Multiplication

It's a mistake to think the practice of my art has come easily to me. I assure you, dear friend, no one has given so much care as I to the study of composition.

- Mozart, to conductor Johann Baptist Kucharz

As you learned long ago, one way to build new functions from old ones is through function *composition*. For example, if we compose the squaring function $f(x) = x^2$ with the doubling function $g(x) = 2x$, we will obtain, depending on the order of composition, either $f\big(g(x)\big) = 4x^2$, or $g\big(f(x)\big) = 2x^2$.

We can also compose linear maps. Composing two yields a third, which of course has its own matrix. We wish to understand how the matrices of the two "parent maps" relate to the matrix of their child, the composite map. We'll begin with an example of finding a composite map's matrix from first principles. Then, after we establish an algebraic theorem about matrix multiplication, we'll return to the same example and view it from another perspective.

Example 1. Let R represent a 90° counterclockwise rotation about the origin, and let F represent a flip (i.e. a reflection) across the horizontal axis. A little thought yields

$$R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ and } F = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Let us consider the composite map that first rotates, then flips. We will call this map's matrix FR. Note the "right-to-left" order: FR means reflect first, then flip. This isn't as odd as it might seem. It's analogous to function notation, where $f(g(x))$ means first apply g, then f.

What does matrix FR look like? We can determine this matrix's columns in the usual way: Rotating $\mathbf{i}$ by 90° then flipping the result over the horizontal axis turns it into $-\mathbf{j}$. Rotating $\mathbf{j}$ by 90° and flipping the result yields $-\mathbf{i}$. Recording these transformed images in a new matrix, we obtain

$$FR = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

♦

We human beings understand that the three matrices above are related through a series of geometric operations that we can visualize. But a computer – a soulless box of chips and wires – can understand none of this. It can only follow orders. Can we program a computer to obtain FR from F and R in a purely algebraic way, by following simple instructions that don't require *thinking* about rotations and reflections? We can. By definition of function composition, we know that $(FR)\mathbf{i} = F(\boldsymbol{R}\mathbf{i})$, and the latter expression can guide us as we program a computer to compute FR's first column mindlessly:

$$F(R\mathbf{i}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\left(\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

You and I can understand the *meaning* of what the computer is doing here: It first determines where $\mathbf{i}$ is mapped by the rotation matrix, and then determines where this transformed image of $\mathbf{i}$ is sent by the flip. But this means nothing at all to the computer, which just blindly follows orders, doing lots of arithmetic – adding, subtracting, and multiplying numbers – and storing the results where we tell it to store them.

Because notation such as FR looks like multiplication, we will in fact call it **matrix multiplication**. In a moment, we'll have an algorithm for doing matrix multiplication quickly. But never let the algorithm distract you from the core meaning: *Matrix multiplication corresponds to the composition of linear maps*. That is why we are interested in it in the first place. Every instance of matrix multiplication can be thought of as the composition of two linear transformations. Never forget.

Now we'll derive two formulas for the product AB of two matrices A and B. They will give us two different perspectives on a matrix product: a *column* perspective (often useful in proofs) and an *entry* perspective (useful in proofs, and also for computing matrix products by hand.)

Matrix Multiplication (Column Perspective).

If A and B are matrices, then AB looks like this:

$$AB = \begin{pmatrix} | & \cdots & | \\ A\mathbf{b_1} & \cdots & A\mathbf{b_n} \\ | & \cdots & | \end{pmatrix},$$

where $\mathbf{b}_j$ represents B's j^{th} column.

Proof.

We must show that AB's j^{th} column is $A\mathbf{b}_j$ for all j. To this end, we note that, for any j, we have

AB's j^{th} column
$= \mathbf{e}_j$'s image under the composite map* (by definition of a linear map's matrix)
$= A(B\mathbf{e}_j)$ (by definition of AB)
$= A\mathbf{b}_j$ (by definition a linear map's matrix) ■

We'll use this column perspective right away – to derive the *entry* perspective!

Matrix Multiplication (Entry Perspective).

If A and B are matrices, we can compute AB's $\boldsymbol{ij}^{\text{th}}$ **entry** (that is, its entry in row i, column j) with a *dot product*:†

$$AB\text{'s } ij^{\text{th}} \text{ entry} = (A\text{'s } i^{\text{th}} \text{ row}) \cdot (B\text{'s } j^{\text{th}} \text{ column})$$

Proof.

AB's ij^{th} entry
$=$ the i^{th} entry in AB's j^{th} column (definition of ij^{th} entry)
$=$ the i^{th} entry in $A\mathbf{b}_j$ (column perspective of matrix multiplication)
$= (A\text{'s } i^{\text{th}} \text{ row}) \cdot \mathbf{b}_j$ (i^{th}-entry formula for matrix-vector multiplication)
$= (A\text{'s } i^{\text{th}} \text{ row}) \cdot (B\text{'s } j^{\text{th}} \text{ column})$ (definition of $\mathbf{b}_j$) ■

* Recall from Chapter 1 that in $\mathbb{R}^n$ (when $n > 3$), we call the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$.

† This is another purely formal "dot product" – a sum of products of corresponding entries in two number lists – with no real connection to the dot product's geometric definition.

From the 'entry perspective', matrix multiplication is a mere algorithmic grind in which we crank out the product one entry at a time, doing the arithmetic in our heads. To see how this works out in practice, let's redo an earlier example. In Example 1, we composed a rotation and a flip, finding the composition's matrix from first principles: We thought geometrically about where the composition would map $\mathbf{i}$ and $\mathbf{j}$, and then we made their transformed images our composition matrix's columns. But now we can redo this problem, producing the composition's matrix in seconds. We begin by writing down our two matrices, along with an empty matrix – soon to be filled in – to represent their product:

$$FR = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} & \\ & \end{pmatrix}.$$

From here, our task is nothing but simple mental arithmetic: FR's top left entry (that is, row 1, column 1) will be, according to the entry perspective, the dot product of F's row 1 and R's column 1.

$$FR = \begin{pmatrix} \boxed{1 \quad 0} \\ 0 \quad -1 \end{pmatrix}\begin{pmatrix} \boxed{\begin{matrix} 0 \\ 1 \end{matrix}} & \begin{matrix} -1 \\ 0 \end{matrix} \end{pmatrix} = \begin{pmatrix} & \\ & \end{pmatrix}.$$

This dot product is $(1)(0) + (0)(1) = 0$, so we put a 0 in the top left corner of our product matrix:

$$FR = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \\ & \end{pmatrix}.$$

Doing the appropriate dot products for the three remaining entries, we end up with, as you should verify,

$$FR = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}.$$

Naturally, this is the same matrix we produced in Example 1, but now we've obtained it by mindlessly turning an algebraic crank, heedless of the *meanings* of any of the matrices involved. Reducing matrix multiplication to an algorithm frees us to concentrate on larger things, and in that sense, it is a blessing… provided that we never forget matrix multiplication's underlying meaning, which is now hidden neatly under the hood: Matrix multiplication corresponds to the *composition* of linear maps.

While Example 1's matrices are here on the page, let's go ahead and compute RF, too:

$$RF = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Note well: $RF \neq FR$. We've discovered something important: **Matrix multiplication is *not* commutative**. This may sound exotic, but on further reflection, it should feel completely natural to you – and to anyone who knows not just how but *why* we multiply matrices. Matrix multiplication corresponds to the composition of linear maps, and everyone knows that when we compose two functions, *the order matters* (as this section's first paragraph reminded us). Hence, order must matter in matrix multiplication, too.

On the other hand, **matrix multiplication is associative**. An algebraic proof of this is unilluminating, but we don't need one. We can see why it is true simply by thinking about matrix multiplication's *meaning*. If we think of each matrix as "doing something" – namely, carrying out the corresponding map – then $A(BC)$ means "do C then B… then do A." On the other hand, $(AB)C$ means "do C… then do B then A." Obviously, the result is the same either way, so it follows that $A(BC) = (AB)C$, as claimed.

Exercises.

16. Carry out the following multiplications using the 'entry perspective' on matrix multiplication:

a) $\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ -1 & 1 \end{pmatrix}$ b) $\begin{pmatrix} 0 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$ c) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 3 \\ 0 & 4 & 0 \end{pmatrix}$

17. Although matrix multiplication is noncommutative in general, certain specific pairs of matrices A and B have the property that $AB = BA$. Think of some specific examples of pairs of 2×2 matrices like this.
[***Hint***: *Think geometrically. Think of linear maps that yield the same result can be done in either order*.]

18. The $n \times n$ *identity matrix*, which you met in Exercise 8c, has 1s on its "main diagonal" (upper left to lower right) and 0s elsewhere. It is denoted by the letter I. Explain why $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$ holds for all $n \times n$ matrices A.

19. In Exercise 11, you saw that an $m \times n$ matrix represents a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. (Note the order!)

a) Suppose A is a 5×2 matrix, and B is a 2×3 matrix. Explain why (in terms of the underlying linear maps) the matrix AB is defined, but BA is undefined.
b) What are the dimensions of matrix AB in the previous part? It represents a linear map from where to where?
c) Now suppose C is a 3×5 matrix. Given the dimensions of A and B above, is the matrix product $BCAB$ defined? If so, find its dimensions.

20. Let $M = \begin{pmatrix} 1 & 2 \\ 3 & 0 \\ -2 & 1 \end{pmatrix}$ and $N = \begin{pmatrix} 2 & -3 & 1 \\ 1 & 1 & 0 \end{pmatrix}$. Find the following (if they are defined): $MN,\ NM,\ M^2,\ N^2,\ NMN$.

21. Consider the linear map of the composition that rotates points in the plane by $120°$ counterclockwise about the origin and then reflects the results over the horizontal axis.

a) Find its matrix. b) Where does this operation send the point $(3, 8)$?
c) If we call this composition's matrix M, find M^2. Why does the matrix that you found make intuitive sense?

22. Prove that scalar multiples can be 'pulled through' matrix multiplication: $(cA)(dB) = (cd)AB$.

23. The matrix that undoes the action of a square matrix A is called the **inverse matrix** of A, and is denoted A^{-1}. Thus, by definition, A^{-1} is a matrix such that $A^{-1}A = I = AA^{-1}$ (where I is the identity matrix).* In this exercise, you'll play with the idea of inverse matrices and learn two important computational facts about them.

a) Let A be the 2×2 matrix of a $90°$ rotation counterclockwise about the origin. Describe what A^{-1} does. Then find A^{-1}. Check your work by grinding out $A^{-1}A$ and $A^{-1}A$. Is the product what you expected?
b) Only *square* matrices can have inverses. Explain why.
c) Prove that matrix inverses are *unique*. (i.e. if B and C are inverses of A, then $B = C$.)
d) Not every square matrix has an inverse. Explain why, for example, the 2×2 zero matrix (see Exercise 6), which crushes the whole plane down into the origin, is not invertible.
e) **Important fact 1.** If A and B are two invertible matrices, then $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$. *Note that reversal of order!* (You can grasp this result intuitively by thinking of A as "put on your shoes" and B as "put on your socks". With those in mind, state the meanings of $A^{-1},\ B^{-1},\ AB,\ BA, (AB)^{-1},$ and $B^{-1}A^{-1}$. The equality will make sense.) Prove the result formally by showing that $B^{-1}A^{-1}$ satisfies the definition of AB's inverse.
f) Let M be the matrix in Exercise 21. Find M^{-1} by using Part D. Check your work by finding M^{-1} another way.
g) **Important fact 2**: If A is an invertible matrix and c is a scalar, then $(\boldsymbol{cA})^{-1} = \boldsymbol{c}^{-1}\boldsymbol{A}^{-1}$. Explain why.

---

* If we wish to show that a matrix B is in fact A^{-1}, we must – by this definition – verify two separate things: $BA = I$ and $AB = I$. But we'll prove later (Ch. 5, Exercise 12) that each of these things *implies the other*, so to verify both, we just need to verify *one*. This is tricky to prove rigorously, but it's intuitively plausible for those (like you!) who understand that matrix multiplication is a kind of *composition*; if B is the map that "undoes" whatever A does, then it makes sense that A is the map that "undoes" B.

Matrix Addition and Transposition

Algebra which cannot be translated into good English
and sound common sense is bad algebra.

- William Kingdon Clifford, *The Common Sense of the Exact Sciences* (Chapter 1, Section 7)

If matrices A and B have the same dimensions, we define their **matrix sum** $A + B$ as the matrix we obtain by adding the corresponding entries in A and B. Thus, for example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} = \begin{pmatrix} 2 & 5 & 8 \\ 6 & 9 & 12 \end{pmatrix}.$$

Matrix subtraction is defined analogously. Compared to matrix multiplication, matrix addition is boring; it lacks a universal geometric meaning, and thus cannot easily be translated into "sound common sense". That said, it is not "bad algebra". Adding and subtracting matrices enriches matrix algebra in subtle ways, usually by simplifying linear algebraic algorithms, which is no small thing. We'll see, for example (in Chapter 7) that our algorithm for finding a matrix's *eigenvalues* will involve matrix subtraction. It will also involve scaling a matrix, an operation you met in Exercise 14. Moreover, there are some instances in which matrix addition does have a tangible interpretation, as you'll see in Exercise 24.

The **transpose** of any matrix M is the matrix whose *columns* are M's *rows* (taken in the usual order). The symbol for M's transpose is M^T. Thus, for example, if

$$M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad \text{then we have} \quad M^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Transposes assume greater prominence as one moves deeper into linear algebra. In time, we'll see how transposes relate to both inverse matrices and *symmetric matrices*, which you'll first meet in exercise 27. As with matrix addition, it's difficult to appreciate transposition's value at the outset, since it lacks a clear geometric interpretation. Still, it's good to make its acquaintance early on. In the exercises that follow, you'll develop a few simple algebraic properties of the transpose, which we'll use later.

Exercises.

24. A grayscale image, such as the one at right, is ultimately an array of pixels. This particular one, for example, is composed of over 100,000 pixels: It is 328 pixels high and 314 wide. When you or I look at it, we behold the King of Rock and Roll and the King of the Beasts. In contrast, a sufficiently tiny bug crawling on the image would see a mere patchwork of shaded squares. In still further contrast, a *computer* would "see" such an image as a matrix. Each pixel in the grayscale image is one of 256 possible shades of gray in a black-to-white spectrum. The different shades in the spectrum are assigned numerical codes, the extreme values being Black = 0 and White = 255. Throughout the range, lower codes correspond to darker shades, higher codes to lighter ones. A computer would "know" this image as a 328 × 314 matrix, whose 102,992 entries are the codes for the shades of gray of the pixels in the corresponding position. For example, the first several rows of this giant matrix would consist of 255s (or numbers close to 255) since the top of the image is black. Call this matrix M. As we'll now see, we can alter the image by doing matrix arithmetic.

a) Roughly speaking, what would the image corresponding to the matrix $2M$ look like? (Assume throughout this exercise that any code greater than 255 is interpreted as if it were 255, and any negative code as if it were 0.)
b) What about $.5M$?
c) What about M^T?
d) What about $255U - M$, where U is the 328×314 matrix consisting entirely of 1s.
e) What about MF, where F is the 314×314 matrix with 1's on its *off* diagonal (bottom left to top right) and 0's elsewhere. [**Hint**: *The 'column perspective' on matrix multiplication should help here*.]
f) What about FM?
g) What about M', the matrix that remains after removing the bottom 14 rows of M?
h) What about FM'?

25. Matrix addition obeys the expected distributive properties, as you should now prove:

a) $(A + B)\mathbf{v} = A\mathbf{v} + B\mathbf{v}$ [**Hint**: *Explain why the sides' i^{th} entries are equal*.]
b) $(A + B)C = AC + BC$ [**Hint**: *Explain why the sides' ij^{th} entries are equal*.]
c) $A(B + C) = AB + AC$

26. In this problem, you'll learn a few algebraic facts about transposes. For parts b – d, your best strategy for the proof is to explain why the sides' ij^{th} entries are equal.

a) A curious and occasionally useful fact: We can compute a dot product with matrix multiplication, one factor of which involves a transpose. Namely, if $\mathbf{v}$ and $\mathbf{w}$ are column vectors, then

$$\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T\mathbf{w}.$$

Explain why. [**Note**: *To make this work, we blur the distinction between a real number and a 1×1 matrix*.]
b) The transpose of a sum is the sum of the transposes: $(M + N)^T = M^T + N^T$. Prove it.
c) Scalar multiples can be pulled through a transpose: $(cM)^T = c(M^T)$. Prove it.
d) The transpose of a product has an important order-reversing property: $(MN)^T = N^T M^T$. Prove it. (The order reversal should remind you of what you learned about matrix *inverses* in Exercise 23c.)
e) Explain why the previous property extends to 3 or more matrices, so that, for example, $(ABC)^T = C^T B^T A^T$.

27. A matrix A is said to be *symmetric* if $A = A^T$.

a) Can a non-square matrix be symmetric? If so, give an example. If not, why not?
b) Write down some symmetric matrices of various sizes.
c) Make up any old 2×3 matrix, M. Compute MM^T. What do you notice? Now compute $M^T M$. Magic, no?
d) In fact, given *any* matrix whatsoever, its product with its transpose (in either order) will always be symmetric. Explain why. (The cleanest - but not the only - way to prove this uses the property you proved in Exercise 26d.)

Abstract Linear Algebra: A Trailer

There is no branch of mathematics, however abstract, which may not someday be applied to phenomena of the real world.

- Nicolai Ivanovich Lobachevski

We can extend the scope and power of linear algebra by considering it from a higher level of abstraction. Abstraction, however, comes at a cost: diminished intuition. Accordingly, too much abstraction is usually a poor choice for one's initial foray into a subject. Part of the goal of this book is to help you develop such a strong intuition for linear algebra in a concrete setting – where vectors are arrows – that you'll be able to appreciate abstract linear algebra, where "vectors" can be all sorts of unexpected things. And even though we've just begun our journey through this book, we've made enough progress already to warrant a brief ascent into the abstract skies – just enough to give you a glimpse of what linear algebra looks like from a higher perspective. Consider this section a trailer for this course's sequel. Later in the book, I may refer in passing to this section, but it is not essential for what follows, and may be skipped.

What are **vectors**? This is the question with which we began Chapter 1. In the abstract perspective, we simply say that a vector is any element of a **vector space**, which itself can be a set of any kind so long as it is closed under addition and multiplication by scalars.* Naturally, the spaces in which we've worked ($\mathbb{R}^n$ and its subspaces) are all examples of vector spaces, but many other sets qualify as well.

For example, $\mathcal{P}_3$, the set of all polynomials of degree 3 (or less) with real coefficients, is a vector space. After all, this set is clearly closed under addition (the sum of any two such polynomials is another one), and multiplication by scalars (multiplying a polynomial by a scalar never increases its degree). Thus $\mathcal{P}_3$ is indeed a vector space, in which the "vectors" are *polynomials*. Another example: $M_{3\times 3}$, the set of all 3×3 matrices, which is obviously closed under matrix addition and under multiplication by scalars. Another notable example comes from the study of differential equations: The set of all solutions to any specific linear differential equation (such as $3y'' + 2y' + 5 = 0$) is a vector space.

For the rest of this section, I'll stick with $\mathcal{P}_3$ for some concrete examples.

Defined abstractly, a **linear transformation** is a function defined on a vector space that preserves linear combinations. In symbols, T is a linear transformation of a vector space if $T(\mathbf{v} + \mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$ and $T(c\mathbf{v}) = cT(\mathbf{v})$ for all scalars c and all vectors $\mathbf{v}, \mathbf{w}$ in the space.

One example of a linear transformation of $\mathcal{P}_3$ is *differentiation*. This might be easier to see if we signify the operation of taking a derivative with a D, rather than the usual prime. For in that case, we have $D\big(p(x) + q(x)\big) = D\big(p(x)\big) + D(q(x))$ and $D\big(cp(x)\big) = cD(p(x))$ for any polynomials $p(x)$ and $q(x)$, as you learned in your first calculus class. Indeed, you may have even learned in that class that these two properties (derivative of a sum is the sum of the derivatives, constant multiples can be pulled through the derivative) are called the derivative's "linearity properties". Now you know the source of that name.

All of Chapter 2's concepts (linear independence, span, basis, etc.) can be defined abstractly so that they can apply to any vector space. For example, consider the four simple "vectors" $x^3,\ x^2,\ x,\ 1$ in $\mathcal{P}_3$. Linear combinations of these four vectors have the form $ax^3 + bx^2 + cx + d$. Ah, but *every* vector in $\mathcal{P}_3$ has that form, so that means that these four vectors *span* $\mathcal{P}_3$. It's also clear that these four vectors are *linearly independent* of each other (each lies outside the span of the other three). Consequently, they constitute a *basis* for $\mathcal{P}_3$. In fact, we call them $\mathcal{P}_3$'s *standard* basis vectors. Moreover, by using a shorthand

* There's a bit more to the abstract definition of a vector space but the details are immaterial in this overview.

column vector notation, we can indicate vectors in $\mathcal{P}_3$ without explicitly writing out $\mathcal{P}_3$'s standard basis vectors. For example, consider $\mathbf{p} = 2x^3 - 4x^2 + 8x + 3$ in $\mathcal{P}_3$. We can rewrite this as the column vector

$$\begin{pmatrix} 2 \\ -4 \\ 8 \\ 3 \end{pmatrix}.$$

Every linear map of $\mathcal{P}_3$ (or any other vector space) can be represented by a *matrix*, whose columns indicate the transformed images of the vector space's standard basis vectors. For example, differentiation sends our first standard basis vector, x^3, to $3x^2$, so the first column of the differentiation matrix will be

$$\begin{pmatrix} 0 \\ 3 \\ 0 \\ 0 \end{pmatrix}.$$

In fact, the full differentiation matrix will be, as you should verify,

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

And just as we used rotation matrices to carry out rotations in $\mathbb{R}^2$, we can use this differentiation matrix to carry out differentiation in $\mathcal{P}_3$. For example, we could differentiate $\mathbf{p} = 2x^3 - 4x^2 + 8x + 3$ as follows:

$$D\mathbf{p} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -4 \\ 8 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \\ -8 \\ 8 \end{pmatrix}.$$

That last column vector corresponds to $6x^2 - 8x + 8$, which is of course the correct expression. We've taken a derivative by means of matrix-vector multiplication!

Since every vector in $\mathcal{P}_3$ can be indicated by a column vector with four entries, $\mathcal{P}_3$ is intimately related to $\mathbb{R}^4$, which of course has the same property. We say that those two vector spaces are **isomorphic**, meaning that although their elements look very different on the surface (arrows in four-dimensional space in one case, polynomials of degree three or lower in the other), the two vector spaces have the same basic structure. Accordingly, we can carry insights about one space over to the other. *Every* finite-dimensional real vector space, in fact, turns out to be isomorphic to $\mathbb{R}^n$ for some value of n. It is this fact that justifies concentrating exclusively on $\mathbb{R}^n$ in a first linear algebra course. Everything you learn about linear algebra in $\mathbb{R}^n$ can be used later to understand what seem to be radically different vector spaces.

Chapter 4

Gaussian Elimination

Linear Systems

Mathematics is the art of reducing any problem to linear algebra.
- William Stein

A **linear equation in n unknowns** is any equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = a,$$

where the a_i's and a are constants. The n "unknowns" are the variables $x_1, x_2, \ldots, x_n$. If there are only a few variables, we might represent them by non-subscripted symbols instead, such as x, y, and z.

A collection of m such equations is called a **linear system** of m equations in n unknows. The system's **solutions** are the points in $\mathbb{R}^n$ whose coordinates satisfy *all* the system's equations. For example, consider the following linear system:

$$\begin{aligned} 2x + y - 3z &= 5 \\ 2x - 4y + 2z &= 0 \\ 4x - 7y + z &= -1. \end{aligned}$$

The point $(3, 2, 1)$ is a solution of the system, as you should verify. On the other hand, $(2, 1, 0)$ is *not* a solution since it satisfies the system's first two equations but fails to satisfy the third.

Systems of linear equations are linked to linear algebra because solving a linear system – regardless of how many equations or variables it contains – is equivalent to solving a *single* matrix-vector equation of the form $A\mathbf{x} = \mathbf{b}$, where $\mathbf{b}$ is a constant vector, and $\mathbf{x}$ is a variable vector consisting of the unknowns we seek. For example, we can repackage the system of equations above in the equivalent form:

$$\begin{pmatrix} 2 & 1 & -3 \\ 2 & -4 & 2 \\ 4 & -7 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ -1 \end{pmatrix}.^*$$

This offers a surprising geometric way to understand the solutions (x, y, z) of the original linear system: The system's solutions are the points of $\mathbb{R}^3$ that the matrix above maps *to* $(5, 0, -1)$.

In theory, we can solve a linear system by rewriting it as $A\mathbf{x} = \mathbf{b}$ and then left-multiplying both sides by A^{-1} (recall the discussion of inverse matrices in Chapter 3, exercise 23) to get the solution $\mathbf{x} = A^{-1}\mathbf{b}$. This idea is conceptually simple, but often practically difficult. To begin with, it frequently won't work, since not all matrices are invertible. And even when A is invertible, the process of first inverting A (we'll learn how later in this chapter) and then multiplying both sides of our equation by it, is computationally "expensive" for large matrices. We'd like a more efficient method for solving a linear system that puts less strain on a computer's processing power. We'll develop one in this chapter.

Computational efficiency is a theme pervading applied linear algebra, since some applications involve gigantic matrices with thousands or even millions of entries. Such giants can be slain only by computer programs, and even then, only by programs whose underlying algorithms are tight, with no extra "fat". An algorithm requiring, say, a billion arithmetic operations might crash a computer when one requiring only a million might be feasible. An entire subfield of linear algebra, *numerical* linear algebra, is devoted to finding computationally efficient algorithms in linear algebra.

---

* A matrix-vector product's first entry is the dot product of the matrix's first row and the vector, so here, the product's first entry is $2x + y - 3z$. Equating this with the right-hand side's first entry, 5, yields $2x + y - 3z = 5$, the first equation in the system. The same idea holds for the remaining entries.

Understanding precisely how such giant matrices - or systems of equations - arise in applications usually requires broad knowledge of the fields (physics, economics, etc.) to which linear algebra is applied. Nonetheless, one can get a general sense of how it happens by considering one characteristic example: the ubiquitous engineering technique of *finite element analysis*. Roughly speaking, an engineer designing an object (anything from a screwdriver to a spacecraft) uses CAD software to break an image of the object into a gigantic but finite mesh of points. Physical factors will then dictate that at each point in the mesh, a linear equation of some sort must be satisfied if the object is to remain functional under outside stresses (the turning of the screwdriver, say, or the passage of the spacecraft through Earth's atmosphere). Collectively, these equations constitute an enormous system of linear equations whose solution can then lead the engineer to perfect his or her design.

Linear algebra's applications abound in nearly all fields of science, computer science, statistics, and higher mathematics itself. Many of these applications ultimately reduce to solving an equation of the form $A\mathbf{x} = \mathbf{b}$, which itself is equivalent to solving a system of linear equations.* Whatever your field of interest, you can find examples of applications online, even if you may not be able to fully understand them yet. Look for some. They will serve as additional inspiration for learning linear algebra.

By Exercise 12 of Chapter 1, each equation in a linear system has a simple graph: a line, plane, or hyperplane, depending on how many variables there are. Thus, solving a linear *system* amounts to finding the points – if any – where certain lines, planes, or hyperplanes all intersect. Thanks to these shapes' 'straightness', it's geometrically clear that their intersection must either be nothing at all, a single point, or infinitely many points. No set of lines, for example, can intersect in precisely *two* points. (If lines, planes, or hyperplanes have points A and B in common, they necessarily have the entire line AB in common.)

To develop our intuition, let's imagine Goldilocks randomly generating – one equation at a time – a system of linear equations in *three* unknowns.† She wants a system with neither too many solutions, nor too few. To Goldilocks's way of thinking, a system that is "just right" is a system with a unique solution.

She begins with a "system" of one equation in three unknowns. Its graph, of course, is a *plane* in $\mathbb{R}^3$. In this rudimentary system, all of the single plane's infinitely many points are solutions. Goldilocks is dissatisfied with this two-dimensional solution space. ("This system has too many solutions!")

Generating a second random equation introduces a second plane. The two planes intersect in a *line*. The solution space has now been shaved down to a one-dimensional space, but its infinitely many points still leave poor Goldilocks nonplussed. ("This system *still* has too many solutions!")

Generating a third random equation introduces a third plane. Naturally, it intersects the line determined by the first two planes in a single point. This point, this sole survivor, is the linear system's *unique* solution – a zero-dimensional solution space that delights Goldilocks. ("This system is just right!")

Suppose Goldilocks foolishly pushes her luck and generates a fourth random equation for the system. Being a randomly generated plane in space, it will almost certainly miss the one point common to the first three planes, thus wiping that system's delicate zero-dimensional solution space out of existence. Alas, poor Goldilocks! ("Ah, now this system has too few solutions!")

Our Goldilocks story illustrates a more general phenomenon: Each equation that we add to a linear system typically reduces the *dimension* of its solution space by 1, until that dimension slips below zero

* Other applications of linear algebra boil down to solving yet another simple-looking equation, $A\mathbf{x} = \lambda\mathbf{x}$, where we seek vectors that are mapped to scalar multiples of themselves. More on this theme in Chapter 7.

† Note well: I've specified that the equations are *randomly* generated to effectively preclude "accidents" such as equations whose graphs are *parallel*. For that to happen, something special would have to be going on; parallels don't occur randomly.

and the solution space vanishes entirely.* The usual recipe for a Goldilocks-approved linear system with a *unique* solution is to have *as many equations as there are unknowns*. A linear system with fewer equations than unknowns will typically have "too many" (that is, infinitely many) solutions; a linear system with more equations than unknowns typically has "too few" solutions (that is, no solutions at all).†

Let's summarize this general tendency in a box.

> If a linear system consists of m equations in n unknowns, then typically (but not necessarily)…
>
> - If $m < n$, the system has **infinitely many solutions**.
> - If $m = n$, the system has **a unique solution**.
> - If $m > n$, the system has **no solution**.

Exercises.

1. Rewrite each of the following linear systems as a matrix equation of the form $A\mathbf{x} = \mathbf{b}$.

a) $\begin{aligned} x + 3y &= 5 \\ 2x - y &= 2 \end{aligned}$ b) $\begin{aligned} 2y + 4z &= -5 \\ x + 3y + 5z &= -2 \\ 3x + 7y + 7z &= 6 \end{aligned}$ c) $\begin{aligned} x - 2y - z &= 3 \\ 3x - 6y - 2z &= 2 \end{aligned}$ d) $\begin{aligned} x + y &= 2 \\ -x + y &= 2 \\ 2x + y &= 1 \end{aligned}$

2. Based exclusively on the number of equations and variables in Exercise 1's linear systems, how many solutions might you reasonably guess each of these systems to have?

3. Rewrite each of the following matrix-vector equations as a system of linear equations.

a) $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$ b) $\begin{pmatrix} 3 & 1 & 4 \\ 1 & 5 & 9 \end{pmatrix}\begin{pmatrix} r \\ s \\ t \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \end{pmatrix}$ c) $\begin{pmatrix} 2 & 7 \\ 1 & 8 \\ 2 & 8 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 9 \end{pmatrix}$

4. Give an example of a system of two linear equations in two unknowns with…
a) a unique solution b) no solution c) infinitely many solutions

5. Repeat the previous exercise, but with systems of *three* linear equations in two unknowns.

6. Describe an arrangement of three planes in $\mathbb{R}^3$ in which each pair intersect, yet there's no intersection of all three. These planes correspond to a linear system. How many equations, unknowns, and solutions does it have?

7. Linear systems can only have 0, 1, or infinitely many solutions. We've accepted this as geometrically intuitive, but we can also prove it algebraically. Here's a sketch of the argument, whose details you should supply:

If $\mathbf{s_1}$ and $\mathbf{s_2}$ are two solutions of a system $A\mathbf{x} = \mathbf{b}$, then every point on the line joining the tips of $\mathbf{s_1}$ and $\mathbf{s_2}$ is a solution. Hence, if a linear system has more than 1 solution, it has *infinitely many* solutions.

* Again, this is the *typical* situation. It's possible that adding an equation to a system won't reduce the dimension of its solution space (ex: if two planes intersect in a line, a third plane could *contain* the whole line rather than intersect it in a point), or that adding an equation could eliminate a solution space of 1+ dimensions altogether in one fell swoop (ex: if two planes intersect in a line, a third plane could be *parallel* to the line). But for such things to happen, the geometric stars must be aligned just so.

† Consider the still simpler scenario where Goldilocks generates a system of equations in *two* unknowns. One equation is a line, a one-dimensional solution space. ("Too many solutions!") Two equations are two lines; these typically intersect in a point, a zero-dimensional solution space. ("Just right!") Three or more equations are three or more lines; these typically have no points in common. ("Too few solutions!")

Augmented Matrices

> The vast white headless phantom floats further and further from the ship, and every rod that it so floats, what seem square roods of sharks and cubic roods of fowls, augment the murderous din.
>
> - Herman Melville, *Moby Dick* (Chapter 69, "The Funeral").

We've now seen how to compress an entire linear system into a single matrix-vector equation, $A\mathbf{x} = \mathbf{b}$. We can compress this further by ditching the equals sign and the "vector of variables" and packing all the remaining vital information into a so-called **augmented matrix**, $(A|\mathbf{b})$. Here, for example, is the same linear system expressed in three equivalent ways, the last of the which is an augmented matrix:

$$\begin{aligned} 2x + 4y - 2z &= 2 \\ 4x + 9y - 3z &= 8 \\ -2x - 3y + 7z &= 2 \end{aligned} \quad \Leftrightarrow \quad \begin{pmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ 2 \end{pmatrix} \quad \Leftrightarrow \quad \left(\begin{array}{ccc|c} 2 & 4 & -2 & 2 \\ 4 & 9 & -3 & 8 \\ -2 & -3 & 7 & 2 \end{array}\right)$$

It's easy to jump back and forth between the system on the left and the augmented matrix on the right. Viewed from afar, the method for solving a linear system that we're about to develop will look like this:

1. Translate the linear system into an augmented matrix.
2. Manipulate the augmented matrix.
3. Translate the results back into a linear system, from which we can then extract the solutions.

All the magic takes place at stage 2. The manipulations we'll use at that stage all share a special property: They all transform an augmented matrix $(A|\mathbf{b})$ into a different but **equivalent** augmented matrix $(A'|\mathbf{b}')$. By an 'equivalent' augmented matrix, I mean one whose corresponding system has the *same solutions* as the original system. Thus, the game is to manipulate the original augmented matrix in this manner until it yields a system whose solutions are obvious, for then these will be our original system's solutions, too. The technique for doing this that we'll develop in the next two sections is called *Gaussian elimination*.

For example, you'll soon be able to verify that when we apply Gaussian elimination to the augmented matrix above, we can transform it into the following equivalent augmented matrix:

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

This corresponds to the utterly trivial linear system whose equations are $x = -7$, $y = 4$, and $z = 0$. Obviously, the sole solution of this system, and hence of the *original* system, is $(-7, 4, 0)$.

Exercises.

8. Rewrite the linear systems in Exercise 1 (on the previous page) as augmented matrices.

9. Rewrite the matrix-vector equations in Exercise 3 as augmented matrices.

10. Rewrite the following augmented matrices as systems of linear equations:

a) $\left(\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}\right)$ b) $\left(\begin{array}{cc|c} 7 & 8 & -3 \\ 9 & 0 & 5 \\ 1 & 2 & 4 \end{array}\right)$ c) $\left(\begin{array}{ccc|c} 0 & 1 & 4 & 1 \\ -2 & 2 & 3 & 0 \\ 5 & -1 & 6 & 8 \end{array}\right)$

11. Rewrite the augmented matrices in Exercise 10 as equations of the form $A\mathbf{x} = \mathbf{b}$.

Row Operations

Merrily, merrily, merrily, merrily,
Life is but a dream.
- Mr. Traditional

Gaussian elimination involves three childishly simple "row operations". By deploying them strategically, we can, with patience, solve any linear system whatsoever. Rather than describing the process abstractly, let's dive in by applying it to the system that you encountered on this chapter's first page:

$$\begin{aligned} 2x + y - 3z &= 5 \\ 2x - 4y + 2z &= 0 \\ 4x - 7y + z &= -1 \end{aligned}$$

As I walk you through this first example, I'll introduce you to each row operation as they appear on stage.

The first row operation: **scale a row** by a nonzero constant. For example, starting with our system's augmented matrix (which I've written below on the left) we'll divide all the entries in its second row by 2, yielding the augmented matrix on the right:

$$\left(\begin{array}{ccc|c} 2 & 1 & -3 & 5 \\ 2 & -4 & 2 & 0 \\ 4 & -7 & 1 & -1 \end{array}\right) \; \div 2 \; \left(\begin{array}{ccc|c} 2 & 1 & -3 & 5 \\ 1 & -2 & 1 & 0 \\ 4 & -7 & 1 & -1 \end{array}\right).^{*}$$

Scaling a row amounts to multiplying both sides of *one* of the system's equations by a nonzero constant. As all algebra students know, doing so preserves *that equation's* solutions. Moreover, the system's other equations are unchanged by the row scale, so the row scale operation preserves *the system's* solutions. Hence, the two augmented matrices above – before and after the row scale – are *equivalent*.

The second row operation is even simpler: **swap two rows**. For example, swapping the first two rows of our second augmented matrix yields a third:

$$\left(\begin{array}{ccc|c} 2 & 1 & -3 & 5 \\ 1 & -2 & 1 & 0 \\ 4 & -7 & 1 & -1 \end{array}\right) \; R_1 \leftrightarrow R_2 \; \left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 2 & 1 & -3 & 5 \\ 4 & -7 & 1 & -1 \end{array}\right).$$

Because swapping rows doesn't change any of the system's equations, it obviously preserves its solutions. Thus, this third augmented matrix is equivalent to the original one.

We mostly use the first two row operations to prepare the ground for the third, which does all our heavy elimination work. Typically, we use the first two operations to insert a 1 in some strategic place. Here, for instance, we used the first two operations to put a 1 in the augmented matrix's upper left corner. As we'll see in the next section, getting a 1 in that spot is typically the first order of business in Gaussian elimination.[†] Once we have a well-positioned 1, we bring in the third operation, which we'll describe next.

---

[*] The $\div 2$ isn't official notation; it's just bookkeeping scratchwork to indicate how the matrices are related. When writing by hand, you should of course omit the shaded box. Different textbook authors use different bookkeeping scratchwork for this sort thing, which hardly matters, since once you've passed your first linear algebra class, writing down the step-by-step details of Gaussian elimination (like writing down the details of long division) is a purely private affair. In public, one just presents the final result.

[†] There are often multiple ways to get a 1 where you want it. Here, for example, we might have put a 1 in the upper left corner with just *one* row operation: divide the initial augmented matrix's *first* row by 2. This would certainly have achieved our goal, but at the cost of introducing unsightly fractions into our matrix.

The third row operation: **to any given row, add (or subtract) a scaled version of another.** To see this operation in action, let's apply it (twice) to our ever-changing but always-equivalent augmented matrix:

$$\left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 2 & 1 & -3 & 5 \\ 4 & -7 & 1 & -1 \end{array}\right) \begin{array}{c} \\ -2R_1 \\ \\ \end{array}{}^{*} \left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 5 & -5 & 5 \\ 4 & -7 & 1 & -1 \end{array}\right) \begin{array}{c} \\ \\ -4R_1 \end{array} \left(\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 5 & -5 & 5 \\ 0 & 1 & -3 & -1 \end{array}\right).$$

Note the result here. We've cleared out the first column. This is the *elimination* in "Gaussian elimination". The 1 now reigns supreme in column 1, presiding over nothing but zeros. Thus, in our ever-changing but always-equivalent linear system, we've arranged it so that only one equation, the top one, involves the first unknown, x.

But why does this third row operation preserve a system's solutions? The explanation isn't difficult, but it is somewhat involved, so I'll delay it until the end of this section, lest it interrupt the flow of our example.

Our next goal is to put a 1 in row 2, column 2, which we can then use to clear out the second column. There are two simple paths to getting a 1 there. One would be to swap the second and third rows. Another path, the one I'll take, is to divide the second row by 5. This latter path has the added advantage of making that row's entries smaller (without introducing fractions), which will simplify our subsequent calculations. With that plan in mind, let's make our next two moves:

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

This last augmented matrix is equivalent to the original one, and its corresponding linear system,

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

is easily solved: We just begin at the bottom equation and work our way to the top via "back substitution". The bottom equation tells us that $z = 1$. Substituting this into the middle equation, we find that $y = 2$. Substituting these y and z values into the top equation, we deduce that $x = 3$. Thus, the original system from the previous page has a unique solution: $(3, 2, 1)$, which you should verify by plugging it back in.

The example that we've just worked through shows the basic strategy behind Gaussian elimination. There are various possible outcomes we must still account for, and some terminology to introduce along the way, but you've already encountered the core ideas. The augmented matrix from which we extracted that simple linear system is an example of what we call **row echelon form**.† The English word *echelon* refers to an arrangement (often military in nature) resembling a flight of stairs, and here, with a little imagination you can imagine the zeros piled up in the lower left portion of the matrix as forming a kind of "staircase". When a matrix is in echelon form, we call its leftmost nonzero entries its **pivots**.

* If the bookkeeping scratchwork isn't clear, $-2R_1$ after a row means "subtract 2 times row 1 from this row".

† **Definition**: A matrix is in *row echelon form* if… (1) In each row, the leftmost *nonzero* entry (if there is one) lies to the right of the leftmost nonzero entries of all rows above it, and (2) If any rows consist entirely of zeros, they are at the bottom of the matrix.

Gaussian elimination lets us transform any given matrix into an equivalent matrix in *row echelon form*. In fact, we often push it further – all the way to ***reduced* row echelon form**, a row echelon form in which two extra conditions hold: (a) all pivots are 1s, and (b) the entries below *and above* the pivots are all zeros. For example, we'll push our recent row echelon form matrix all the way to *reduced* row echelon form:

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

Observe that the pivots of this last matrix are all 1s, and the entries below and above the pivots are all 0s. Hence, it is in *reduced* row echelon form, as claimed. In this form, the corresponding linear system becomes utterly trivial. Solving it doesn't even require back substitution. The system *is* the solution:

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

A nice feature of *reduced* row echelon form (as opposed to mere row echelon form) is its *uniqueness*: each matrix is equivalent to one and only one matrix in this form. This makes it useful for theoretical considerations, allowing us to refer unambiguously to *the* reduced row echelon form of a given augmented matrix $(A|\mathbf{b})$, which we denote, unsurprisingly, as $\mathrm{rref}(A|\mathbf{b})$. Thus, over the past few pages, we've shown that

$$\mathrm{rref}\left(\begin{array}{ccc|c} 2 & 1 & -3 & 5 \\ 2 & -4 & 2 & 0 \\ 4 & -7 & 1 & -1 \end{array}\right) = \left(\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right).$$

Before turning to some exercises, we must attend to an unfinished piece of business: We must make good on our earlier promise to prove that the third row operation preserves a linear system's solutions.

Claim. The third row operation described above (*to one given row, add a multiple of another*) preserves the underlying linear system's solutions.

Proof. Any linear equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = a$ can be expressed in the form $\mathbf{a} \cdot \mathbf{x} = a$, where vectors $\mathbf{a}$ and $\mathbf{x}$ consist of the coefficients and variables of the equation's left-hand side. This compact notation will help us in the argument that follows.

In an augmented matrix, take a row corresponding to $\mathbf{a} \cdot \mathbf{x} = a$, and to it add c times a row corresponding to $\mathbf{b} \cdot \mathbf{x} = b$. The result is a new system, identical to the old one, except that in place of a row corresponding to $\mathbf{a} \cdot \mathbf{x} = a$ it has one corresponding to $(\mathbf{a} \cdot \mathbf{x}) + c(\mathbf{b} \cdot \mathbf{x}) = a + cb$.

If $\mathbf{s}$ is a solution to the original system, then $\mathbf{a} \cdot \mathbf{s} = a$ and $\mathbf{b} \cdot \mathbf{s} = b$, so substituting $\mathbf{s}$ into the new linear equation, highlighted in the previous paragraph, will clearly yield a true statement. Thus, all the solutions of the original system are solutions of the new one, too.

Conversely, if $\mathbf{s}'$ is a solution to the *new* system, it satisfies the *old* row's equation $\mathbf{a} \cdot \mathbf{x} = a$. To see why, we simply note that $\mathbf{a} \cdot \mathbf{s}' = [(\mathbf{a} \cdot \mathbf{s}') + c(\mathbf{b} \cdot \mathbf{s}')] - [c(\mathbf{b} \cdot \mathbf{s}')] = (a + cb) - cb = a$, where the second equals sign holds because $\mathbf{s}'$ satisfies the *new* system's equations (which include both $\mathbf{b} \cdot \mathbf{x} = b$ and $(\mathbf{a} \cdot \mathbf{x}) + c(\mathbf{b} \cdot \mathbf{x}) = a + cb$). Thus, all the solutions of the new system are also solutions of the old one.

Having shown that the new and old systems have precisely the same solutions, we conclude that the third row operation preserves solutions, as claimed. ■

Exercises.

12. Carry out the row operations indicated below by filling in the entries of the second matrix:

a) $\left(\begin{array}{cc|c} 4 & 8 & 12 \\ 2 & 3 & 5 \end{array}\right) \div 4 \left(\begin{array}{c|c} \quad & \quad \end{array}\right)$ b) $\left(\begin{array}{ccc|c} 3 & 6 & 9 & 12 \\ 2 & 5 & 8 & 11 \\ 1 & 4 & 7 & 10 \end{array}\right) R_1 \leftrightarrow R_3 \left(\begin{array}{c|c} \quad & \quad \\ & \\ & \end{array}\right)$

c) $\left(\begin{array}{cccc|c} 2 & 3 & 25 & 1 & 2 \\ 0 & 1 & 8 & -3 & 7 \\ 0 & -2 & -10 & 6 & 1 \end{array}\right) \begin{array}{c} \\ \\ +2R_2 \end{array} \left(\begin{array}{c|c} \quad & \quad \\ & \\ & \end{array}\right)$

d) $\left(\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 5 \\ 6 & 2 & 11 & -10 & -2 \end{array}\right) \begin{array}{c} \\ -6R_1 \end{array} \left(\begin{array}{c|c} \quad & \quad \\ & \end{array}\right)$

13. When in the course of Gaussian elimination, it becomes necessary to dissolve all the nonzero entries lying above or below a pivot, we usually alter all rows at once, rather than rewriting the augmented matrix over and over, changing one row each time.

a) Try the "all rows at once" approach here:

$$\left(\begin{array}{cc|c} 1 & -3 & 1 \\ -2 & -1 & 4 \\ 3 & 2 & 5 \\ -1 & 2 & 0 \end{array}\right) \begin{array}{c} \\ +2R_1 \\ -3R_1 \\ +R_1 \end{array} \left(\begin{array}{c|c} \quad & \quad \\ & \\ & \\ & \end{array}\right)$$

b) Now do it again, but now *you* determine the appropriate multiple of row 2 to add or subtract from each row to "clear out" the second column, putting zeros above and below the 1 in the matrix's center:

$$\left(\begin{array}{ccc|c} 1 & -4 & 5 & 0 \\ 0 & \mathbf{1} & -2 & 6 \\ 0 & 1/2 & -3 & 4 \end{array}\right) \qquad \left(\begin{array}{c|c} \quad & \quad \\ & \\ & \end{array}\right)$$

14. If A is a *square* matrix, then $A\mathbf{x} = \mathbf{b}$ usually has a unique solution since there are as many equations as unknowns. Gaussian elimination supplies us with a straightforward (albeit tedious) strategy for finding that unique solution: Begin by getting a 1 in the top left entry and then using it to clear out the rest of first column. Then get a 1 in the main diagonal's next entry (i.e. row 2, column 2) and use it to clear out the rest of column two. Continue going down the main diagonal like this until you've transformed A into the identity matrix I. At that point you'll have transformed $(A|\mathbf{b})$ to an equivalent augmented matrix $(I|\mathbf{b}')$, from which you can simply read off the solution. Use this strategy to solve the following systems. (This will involve some arithmetic with fractions. Such is life.) Check your work by substituting your solutions back into the original system.

a) $\begin{aligned} 2x + 3y &= 0 \\ 4x + 5y &= 0 \end{aligned}$ b) $\begin{aligned} 4x + 3y &= 2 \\ 7x + 5y &= 3 \end{aligned}$ c) $\begin{aligned} x + 2y + 3z &= 1 \\ 2x + 4y + 7z &= 2 \\ 3x + 7y + 11z &= 8 \end{aligned}$ d) $\begin{aligned} x + 2y + 3z &= 8 \\ x + 3y + 3z &= 10 \\ x + 2y + 4z &= 9 \end{aligned}$

15. Suppose that while carrying out Gaussian elimination on an augmented matrix $(A|\mathbf{b})$, we end up at some point with a row of the form $(0 \quad 0 \quad \cdots \quad 0 \mid c)$, where c is a nonzero constant. What can we conclude about the linear system underlying the original augmented matrix? Explain why this is so.

16. Classify each of the following as being in row echelon form, *reduced* row echelon form, or neither.

a) $\left(\begin{array}{cccc|c} 1 & 4 & 0 & 5 & 7 \\ 0 & 0 & 1 & 6 & 0 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$ b) $\left(\begin{array}{cccc|c} 0 & 1 & 8 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right)$ c) $\left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$ d) $\left(\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 2 & 0 & 5 \\ 0 & 0 & 3 & 6 \end{array}\right)$

Gaussian Elimination (Endgames and Algorithm)

> When you have eliminated all which is impossible, then
> whatever remains, however improbable, must be the truth.
> - Sherlock Holmes, in Arthur Conan Doyle's *The Sign of Four*

You met Gaussian elimination in the preceding section. Now we'll consider the various "endgames" that can arise from it. So far, we've seen just one: the case where the underlying system has a unique solution. After considering other possible endgames, I will conclude this section with a universal algorithm for Gaussian elimination, which frankly, you shouldn't take too seriously – unless you are programming a computer. No human being slavishly follows the algorithm as stated (we tend to deviate from it when we spot shortcuts or ways of avoiding fractions), but it's still interesting to see the process reduced to a short clinical sequence of steps that a machine can follow.

Recall the goal: We want to put a given augmented matrix into *reduced row echelon form*. Our strategy is to scale and swap rows until we get a 1 into a key "pivot" position, and then we clear out the rest of the 1's column by using our third row operation (*to one row, add or subtract a multiple of another*). We then move on, always working from the top row down, developing an echelon structure in the matrix as we go.

Example 1. Use Gaussian elimination to find all solutions to the following linear system:

$$\begin{aligned} x + 4y - 5z &= 0 \\ 2x - y + 8z &= 0. \end{aligned}$$

Solution. After rewriting the system as an augmented matrix, Gaussian elimination gives us this:

$$\left(\begin{array}{ccc|c} 1 & 4 & -5 & 0 \\ 2 & -1 & 8 & 0 \end{array}\right) \underset{-2R_1}{} \left(\begin{array}{ccc|c} 1 & 4 & -5 & 0 \\ 0 & -9 & 18 & 0 \end{array}\right) \underset{\div(-9)}{} \left(\begin{array}{ccc|c} 1 & 4 & -5 & 0 \\ 0 & 1 & -2 & 0 \end{array}\right) \overset{-4R_2}{} \left(\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 0 \end{array}\right).$$

This final matrix, which is in reduced row echelon form, corresponds to this linear system:

$$\begin{aligned} x \quad + 3z &= 0 \\ y - 2z &= 0. \end{aligned}$$

We can solve this system – and others resembling it – by keeping each equation's first variable on the left, while moving any others to the right. Doing so here yields the following system:

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

This tells us that in any solution, the values of x and y are determined entirely by z's value. Meanwhile, z has no constraints on it whatsoever. It is a "free variable".* If we let $z = 1$, we find that $(-3, 2, 1)$ is a solution. If we let $z = 2$, we find that $(-6, 4, 2)$ is a solution. Indeed, if we let $z = t$, where t can be any real number, we find the system's complete infinite set of solutions: **All points of the form** $(-3\boldsymbol{t}, 2\boldsymbol{t}, \boldsymbol{t})$, where t can be any real number. ♦

When a linear system has infinitely many solutions, it often helps to think about them geometrically. As an aid to such thinking, we can take our expression for the general solution, rewrite it in vector form, and then distill it into simpler parts that elucidate the solution space's geometry. Thus, in the preceding example, we could take the general solution $(-3t, 2t, t)$, and rewrite it as follows:

* Note that our free variable corresponds to a *pivotless* column in the matrix – a theme we'll develop in this section's footnotes.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3t \\ 2t \\ t \end{pmatrix} = \boldsymbol{t}\begin{pmatrix} \mathbf{-3} \\ \mathbf{2} \\ \mathbf{1} \end{pmatrix}.$$

The expression on the right shows that the system's solutions are all the scalar multiples of $-3\mathbf{i} + 2\mathbf{j} + \mathbf{k}$. That is, the solutions constitute a line in $\mathbb{R}^3$ passing through the origin. This should come as no surprise; the system's equations correspond to two planes in $\mathbb{R}^3$, which, of course, typically intersect in a line.

Example 2. Use Gaussian elimination to solve the following linear system:

$$4x - 4y = 4, \qquad 2x + y = 2, \qquad 3x - y = 0.$$

Solution. After rewriting the system as an augmented matrix, Gaussian elimination yields

$$\left(\begin{array}{cc|c} 4 & -4 & 4 \\ 2 & 1 & 2 \\ 3 & -1 & 0 \end{array}\right) \div 4 \left(\begin{array}{cc|c} 1 & -1 & 1 \\ 2 & 1 & 2 \\ 3 & -1 & 0 \end{array}\right) \begin{matrix} \\ -2R_1 \\ -3R_1 \end{matrix} \left(\begin{array}{cc|c} 1 & -1 & 1 \\ 0 & 3 & 0 \\ 0 & 2 & -3 \end{array}\right) \div 3 \left(\begin{array}{cc|c} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 2 & -3 \end{array}\right) \begin{matrix} +R_2 \\ \\ -2R_2 \end{matrix} \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{array}\right).$$

We've now found that our original system is equivalent to one containing the equation $0 = -3$. A system containing an *equation* with no solution is clearly a *system* with no solution. Hence, our original system has **no solution**. (Not surprising, since it has more equations than variables.) ♦

We know that a row of the form $(0\ 0\ \cdots 0 \mid c)$ where $c \neq 0$ indicates that the system has no solution. But what if $c = 0$? A **zero row** $(0\ 0\ \cdots 0 \mid \mathbf{0})$ corresponds to $0 = 0$. In the context of Gaussian elimination, this triviality actually tells us something: Our original system is equivalent to one *with one fewer equation*. A zero row is thus a signal that our original system harbors a "redundant" equation that could be removed from the system without changing its solution space. Redundant equations are best understood as redundant *constraints* on the solution space. As in our earlier Goldilocks example, each successive equation typically constrains the solution space to a smaller region because it is effectively demanding, "to survive in the solution space, points must lie on *my graph too*!" A redundant equation is therefore one whose demands *have already been met* by all the points that have survived all the earlier constraints. For example, consider the linear system consisting of the three equations $x = y$, $2x = 2y$, $x = -y$. Usually, 3×2 systems (i.e. 3 equations, 2 unknowns) lack solutions, but this one *has* a solution: (0,0). This is because the second equation is redundant in this system. We could eject it from the system without altering the solution space. Hence, this 3×2 system "might as well be" a $\mathbf{2} \times 2$ system… which is the sort of system we *expect* to have one solution. If we had not spotted its redundant equation at the outset and we had tried to solve the system via Gaussian elimination, we'd have ended up with a telltale zero row:

$$\left(\begin{array}{cc|c} 1 & -1 & 0 \\ 2 & -2 & 0 \\ 1 & 1 & 0 \end{array}\right) \begin{matrix} \\ -2R_1 \\ -R_1 \end{matrix} \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 2 & 0 \end{array}\right) R_2 \leftrightarrow R_3 \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right) \div 2 \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right) \begin{matrix} +R_2 \\ \\ \\ \end{matrix} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right).$$

I've put the matrix into rref for completeness's sake, but observe that the zero row turned up very early: in the second augmented matrix. Once a zero row appears, we know that we have a redundant equation. In this example, the redundancy was obvious (since two of the equations have the same graph), but redundancies are usually subtler than that. In our next example, we'll consider three linear equations in three unknowns, which we might *expect* to have a unique solution, but… one equation will turn out to be redundant, so the system might as well consist of only *two* equations in three unknowns, and will, accordingly, have infinitely many solutions. The telltale zero row will emerge during Gaussian elimination.

Example 3. Use Gaussian elimination to solve this linear system:

$$\begin{aligned} x + 2y \ + 3z &= 0 \\ 4x + 5y + 6z &= 3 \\ 7x + 8y + 9z &= 6 \end{aligned}$$

Find the system's solutions and describe them geometrically.

Solution. After rewriting the system as an augmented matrix, Gaussian elimination gives us

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 3 \\ 7 & 8 & 9 & 6 \end{array}\right) \begin{array}{c} \\ -4R_1 \\ -7R_1 \end{array} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 3 \\ 0 & -6 & -12 & 6 \end{array}\right) \begin{array}{c} \\ \div(-3) \\ \div(-6) \end{array} \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & 1 & 2 & -1 \end{array}\right) \begin{array}{c} -2R_2 \\ \\ -R_2 \end{array} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

This last matrix (which is in rref, as you should verify) has a zero row, which signals a redundant equation in the original system. Accordingly, this "might as well be" a system of *two* linear equations in three unknowns, so we expect it to have infinitely many solutions. We can write them all down as follows: First, we extract the linear system, omitting the useless $0 = 0$ equation.

$$\begin{aligned} x \quad\ - z &= 2 \\ y + 2z &= -1. \end{aligned}$$

Now, as in Example 1, we isolate each equation's first variable:

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

This shows that x and y are determined by z, which is a free variable. Letting $z = t$, where t is free to be any real number whatsoever, we find the system's set of infinitely many solutions: **All points of the form** $\boldsymbol{\big((2+t),\ (-1-2t),\ t\big)}$, where t can be any real number.

To describe these solutions geometrically, we take the preceding expression for the general solution, rewrite it in vector form, and distill it into simpler parts:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2+t \\ -1-2t \\ t \end{pmatrix} = \begin{pmatrix} \mathbf{2} \\ \mathbf{-1} \\ \mathbf{0} \end{pmatrix} + \boldsymbol{t} \begin{pmatrix} \mathbf{1} \\ \mathbf{-2} \\ \mathbf{1} \end{pmatrix}.$$

From our work in Chapter 2's final section, we recognize this as a parametric representation of a line in $\mathbb{R}^3$: the line that passes through point $(2, -1, 0)$ and is parallel to the vector $\mathbf{i} - 2\mathbf{j} + \mathbf{k}$. The infinitely many points on this line are the linear system's infinitely many solutions. ♦

For our last examples, I'll skip the elimination details and just concentrate on interpreting the rref.

Example 4. Suppose that after Gaussian elimination, we end up with this augmented matrix:

$$\left(\begin{array}{cccc|c} 1 & 0 & -3 & 2 & 5 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).^*$$

What are the solutions to the corresponding linear system? Describe them geometrically.

---

* This contains *two* pivotless columns. Their corresponding variables (z and w) will be *free* in our solution set, and it's clear why: In any pivotless column, each nonzero entry gets pushed to the right of its row's equation, where it is free to take any value. **Question**: If a pivotless column's entries are all 0, is its corresponding variable free? Try to answer this. (I'll do so in Example 5).

Solution. The corresponding linear system is

$$\begin{aligned} x \quad - 3z + 2w &= 5 \\ y + z - w &= 0. \end{aligned}$$

Pushing the non-leading variables to the right, this becomes

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

We can select the values of the two free variables, z and w, at will. They, in turn, determine the values of x and y. (If, say, $z = 0$ and $w = 1$, then $x = 3$ and $y = 1$, so $(3,1,0,1)$ is a solution.) To specify all solutions, we must introduce *two* parameters: s for z's value and t for w's value. Our solutions are thus $((\mathbf{5+3s-2t}), (-\boldsymbol{s}+\boldsymbol{t}),\ \boldsymbol{s},\ \boldsymbol{t})$**,** where s and t can be any real numbers.

To describe the solutions geometrically, we write them in vector form, and decompose them:

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 5+3s-2t \\ -s+t \\ s \\ t \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 3 \\ -1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -2 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

From our work in Chapter 2's final section, we recognize this as a parametric representation of a two-dimensional plane in $\mathbb{R}^4$. Namely, the plane that passes through point $(5,0,0,0)$ and is parallel to the vectors $3\mathbf{e_1} - \mathbf{e_2} + \mathbf{e_3}$ and $-2\mathbf{e_1} + \mathbf{e_2} + \mathbf{e_4}$. ♦

Our final example will show us that not all free variables turn up on our equations' right-hand sides.

Example 5. Suppose that after Gaussian elimination, we end up with this augmented matrix:

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

What are the solutions to the corresponding linear system? Describe them geometrically.

Solution. Extracting the linear system and pushing non-leading variables to the right, we obtain

$$\begin{aligned} x_1 &= 3 - 2x_5 \\ x_3 &= -x_4 - 3x_5 \end{aligned}$$

Clearly, x_4 and x_5 are free variables. But so is x_2: It appears in no equations since its corresponding column is $\mathbf{0}$. Thus, *nothing constrains its value*. It is free.* Accordingly, we'll need *three* parameters for our solution set. I'll use s, t, u for x_2, x_4, x_5 respectively. Our solutions will thus be all points of the form $\left((\mathbf{3-2}\boldsymbol{u}),\ \boldsymbol{s},\ (-\boldsymbol{t}-\mathbf{3}\boldsymbol{u}),\ \boldsymbol{t},\ \boldsymbol{u}\right)$, for any reals s, t, u**.** The three parameters suggest that our solution will be a three-dimensional hyperplane in $\mathbb{R}^5$, which is indeed the case:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 3-2u \\ s \\ -t-3u \\ t \\ u \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} + u\begin{pmatrix} -2 \\ 0 \\ -3 \\ 0 \\ 1 \end{pmatrix}. \quad ♦$$

* We've now explained why pivotless columns correspond to free variables – whether or not they contain any nonzero entries. On the next page, we'll parlay this insight, gleaned so stealthily down here in the footnotes, to produce a nice time-saving trick.

And that's that. You can now use Gaussian elimination to solve any linear system that confronts you. To honor this, I'll show you a time-saving trick. Once an augmented matrix is in reduced row echelon form, we can quickly extract its solutions if we know (as all careful footnote readers do!) that pivotless columns correspond to free variables. For example, in Example 1, we ended up with this augmented matrix in rref:

$$\begin{array}{c} \begin{matrix} x & y & z \end{matrix} \\ \left(\begin{array}{ccc|c} \textcircled{1} & 0 & 3 & 0 \\ 0 & \textcircled{1} & -2 & 0 \end{array}\right). \end{array}$$

For emphasis, I've circled the pivots and written the corresponding variable at the top of each column. Since the z-column is a pivotless column, z will be a free variable. (This is because if we wrote the second row out as an equation, we'd end up pushing z over to the right-hand side, where the free variables live.) On the other hand, neither x nor y will be free, since their columns are pivot columns. With that in mind, we can write down a vector expression for the solutions very quickly, as follows. We begin by assigning parameters to any free variables (I'll use t for our free variable z) while leaving the other entries blank:

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

Next, we look back at our augmented matrix, mentally turning each row in turn into an equation, whose non-leading terms we shift to the other side, replacing free variables with their associated parameter(s). This lets us fill in our solution vector's blanks. In the first row here, for example, we mentally push the $3z$ over to the other side, obtaining $x = -3z$. But we've assigned t to z, so we have $x = -3t$. Thus $-3t$ goes in our solution vector's top slot. After doing the same thing for the middle row, we end up with:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3t \\ 2t \\ t \end{pmatrix}.$$

From there, we can, of course, pull out the t if we wish, as we did in Example 1.

Similarly, in Example 4, our rref augmented matrix was – apart from my new ornamentations – this:

$$\begin{array}{c} \begin{matrix} x & y & z & w \end{matrix} \\ \left(\begin{array}{cccc|c} \textcircled{1} & 0 & -3 & 2 & 5 \\ 0 & \textcircled{1} & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right). \end{array}$$

To write down the solutions, we first observe that the pivotless columns correspond to z and w, so these are free variables. We'll assign them the parameters s and t respectively. If we put those in our solution vector and then fill in the remaining slots (for x and y) with a little mental algebra, we quickly obtain

$$\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 5 + 3s - 2t \\ -s + t \\ s \\ t \end{pmatrix},$$

which, of course, we can then decompose into a linear combination of three vectors if we wish.

This little trick will speed up your work at the end of the row reduction game, so feel free to use it.

Finally, as promised, I'll conclude this section with a formal Gaussian elimination algorithm guaranteed to reduce any matrix to its reduced row echelon form. In practice, when we do Gaussian elimination by hand, we follow the spirit of the algorithm without following it to the letter. Computers, however, cannot understand that distinction, and must be fed explicit instructions.

Gaussian Elimination Algorithm

1. a) Go to the first nonzero column (i.e. the first column whose entries aren't *all* zeros).
b) Swap rows, if necessary, to put a nonzero entry on top of this column.
2. a) Scale the top row, if necessary, to make its leftmost entry equal to 1.
b) Eliminate all nonzero entries in that 1's column (using the third row operation).
3. Consider the "submatrix" of all entries that lie both *to the right of* and *below* that 1.
a) If there are no such entries, stop the algorithm.
b) If there are, return to Step 1 in the algorithm, but now apply it to the *sub*matrix. (When you get back to the "elimination" stage, though, you should still eliminate all nonzero entries in the column, including any lying above the submatrix.)

Following the algorithm literally and following its general spirit often leads us to do the exact same things. But following its spirit sometimes gets us to our destination (rref) more comfortably and quickly. Compare the directions given to us by a GPS system: They'll always get you to your destination, but sometimes via needlessly complicated routes that would exasperate any driver who is already familiar with the roads.

Exercises.

17. Use Gaussian elimination to solve each of the following linear systems. Describe the solutions geometrically.

a) $\begin{aligned} 2x+5y-8z&=4\\ x+2y-3z&=1\\ 3x+8y-13z&=7\end{aligned}$ b) $\begin{aligned} x+3y-2z+w&=3\\ 2x+6y-3z-3w&=7\end{aligned}$ c) $\begin{aligned} x-2y+4z&=2\\ 2x-3y+5z&=3\\ 3x-4y+6z&=7\end{aligned}$

d) $\begin{aligned} 4y+z&=2\\ 2x+6y-2z&=3\\ 4x+8y-5z&=4\end{aligned}$ e) $\begin{aligned} x+y+z&=5\\ 2x+3y+5z&=8\\ 4x+5z&=2\end{aligned}$ f) $\begin{aligned} x_4+2x_5-x_6&=2\\ x_1+2x_2+x_5-x_6&=0\\ x_1+2x_2+2x_3-x_5+x_6&=2\end{aligned}$

18. Each of the following augmented matrices is already in reduced row echelon form. For each, describe the number of solutions that its associated system of equations has, and explain how the solutions are arranged geometrically (e.g. a point in space, a plane in space, a 4-dimensional hyperplane in $\mathbb{R}^8$, etc.)

a) $\left(\begin{array}{cccc|c} 1&0&0&0&0\\ 0&1&1&2&2\\ 0&0&0&0&0\end{array}\right)$ b) $\left(\begin{array}{ccc|c} 1&0&0&0\\ 0&1&0&-1\\ 0&0&1&2\\ 0&0&0&1\end{array}\right)$ c) $\left(\begin{array}{ccc|c} 1&-2&3&1\\ 0&0&0&0\\ 0&0&0&0\end{array}\right)$ d) $\left(\begin{array}{cc|c} 1&0&0\\ 0&1&0\end{array}\right)$

e) $\left(\begin{array}{ccccc|c} 0&1&3&0&2&7\\ 0&0&0&1&5&8\\ 0&0&0&0&0&9\end{array}\right)$ f) $\left(\begin{array}{ccccccc|c} 1&-2&0&0&-4&0&6&8\\ 0&0&1&0&3&5&7&9\end{array}\right)$

19. Is the vector $\mathbf{v}=\begin{pmatrix}3\\-7\\-3\end{pmatrix}$ in the span of the columns of matrix $A=\begin{pmatrix}1&-4&2\\0&3&5\\-2&8&-4\end{pmatrix}$?

20. If Example 5's augmented matrix had a third row – a *zero row* at the bottom – would the solution set change?

21. (Test for linear independence)

At the end of Chapter 2's first section, we met an "alternate characterization" of linear independence:

$$\begin{array}{c}\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n} \\ \text{are linearly independent}\end{array} \Leftrightarrow \begin{array}{c}\mathbf{0} \text{ can be expressed as a linear combination of } \mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n} \\ \text{in only one way (the trivial way, where all coefficients } = 0).\end{array}$$

In the same section, I noted that "we'll eventually be able to use [this characterization] to conduct algorithmic tests for linear independence – but only after we've developed a robust technique for solving systems of linear equations." With Gaussian elimination in hand, we can now proceed to our linear independence test.

a) Explain why the following works: To see whether $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$ are linearly independent, we make them the columns of a matrix A. We then check to see if $A\mathbf{x} = \mathbf{0}$ has any nontrivial solutions. If the only solution is $\mathbf{0}$, the vectors are linearly independent. Otherwise, the vectors are linearly dependent.

(Don't memorize this test. Instead, it's the alternate characterization of linear independence you should bear in memory. If you understand it, you should be able to reconstruct the test in seconds whenever it's needed.)

b) Check to see if the following sets of vectors are linearly independent:

i) $\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}3\\5\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}$ *ii*) $\begin{pmatrix}1\\2\\3\\4\end{pmatrix}, \begin{pmatrix}0\\1\\-2\\1\end{pmatrix}, \begin{pmatrix}2\\7\\0\\11\end{pmatrix}$ *iii*) $\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}2\\5\\7\end{pmatrix}, \begin{pmatrix}9\\1\\4\end{pmatrix}, \begin{pmatrix}6\\8\\3\end{pmatrix}$ [***Think*** *before computing*!]

22. As you've discovered by now, doing Gaussian elimination by hand can be a slightly nerve-wracking task, since one arithmetic mistake will ruin everything. (And you've had it easy. I've mostly stuck to "nice" examples: fairly small matrices whose entries have been cooked up in such a way that didn't require you to do arithmetic with fractions.) Though tedious for humans, this is precisely the sort of task at which computers excel. For this problem, figure out how to use the technology of your choice (calculator, software, computer program you've written yourself, or whatnot) to put a matrix into reduced row echelon form.*

a) Use technology to redo a few parts of Exercises 17 and 21b.

b) Determine whether either of the vectors below lies in the span of the three vectors in Exercise 21b, part *ii*. Do this by writing down an appropriate matrix equation, and then solving it with technology. If either vector does lie in the span of the three vectors above, express it as a linear combination of those three vectors.

$$\mathbf{v} = \begin{pmatrix}3\\1\\4\\1\end{pmatrix}, \qquad \mathbf{w} = \begin{pmatrix}22\\69\\16\\113\end{pmatrix}.$$

23. I mentioned that slavishly following the algorithm isn't always advisable. To see what I mean, solve the following system two ways: first as you normally would (by beginning with a row swap to get a 1 in the top left corner), and then by following the boxed algorithm instructions in a lockstep manner:

$$\left(\begin{array}{cc|c}2 & 1 & 9\\ 1 & 5 & 0\end{array}\right).$$

24. a) Is there a quadratic polynomial whose graph passes through points $(1, 0)$, $(-1, 1)$, and $(2, 3)$? If so, find it.

b) How about a cubic polynomial whose graph passes through points $(0, 1)$, $(1, 0)$, $(-1, 0)$, and $(2, 15)$?

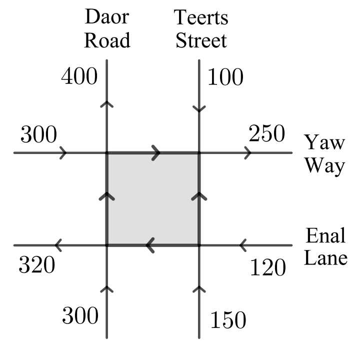
c) Is there a quadratic that passes through the four points in Part B? How about a quartic?

* Wolfram Alpha's website is easy to use. For example, to put the augmented matrix in Problem 23 in reduced row echelon form there, one would simply type **rref{{2,1,9},{1,5,0}}** into the "natural language" bar. Observe that the matrix must be entered *row by row*, not column by column.

25. The schematic drawing at right represents the four one-way streets that comprise the boundary of Azalp Plaza in downtown Ytic City. The arrows indicate the traffic's direction, and the numbers indicate the number of cars that pass along each block during rush hour.

There is no parking in the area shown in our map, so each car that enters the area during rush hour must also depart during rush hour.

The arrows along Azalp Plaza itself have been left unnumbered. What, if anything, can we say about the number of cars passing along each of those four blocks during rush hour? For each of the four, specify the maximum and minimum number of cars.



26. In the year 656 AD, a collection of ten Chinese mathematical texts were designated as preparation for the examinations required of anyone who would enter the Tang dynasty's civil service. One of these texts, by Zhang Qiujian, ends with the following problem: "One rooster is worth 5 coins, one hen 3 coins and 3 chicks 1 coin. It is required to buy 100 fowls with 100 coins. Find the number of roosters, hens and chicks bought."

So, can *you* qualify for the 7th-century Chinese bureaucracy?

27. If you are interested in seeing how systems of linear equations (and thus linear algebra) can arise in various fields, search online or in books for information on one or more of the following and report your findings.

a) *Leontif input-output models* in economics

b) electrical circuits and systems of linear equations

c) balancing chemical equations by solving a system of linear equations

The Matrix Inversion Algorithm

Man's sensibility to trifles and insensibility to great things indicates a strange inversion.
- Blaise Pascal, *Pensées*

You met inverse matrices in Chapter 3's Exercise 23, but we haven't yet discussed *how* to invert a matrix. We'll do that in this section. Recall that nonsquare matrices aren't invertible, and that only certain square matrices are. Given an invertible matrix A, we want to find an algorithm that yields its inverse matrix A^{-1}, which "undoes" A's action, so that the product of A and A^{-1} (in either order) is the identity matrix, I.

We'll begin our quest for an inversion algorithm with some preliminary observations.

On Chapter 4's first page, we noted that if A is invertible, then for each vector $\mathbf{b}$, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution. One way to find that unique solution (i.e. to solve for $\boldsymbol{x}$) is to left-multiply both sides by A^{-1}, thereby obtaining $\mathbf{x} = A^{-1}\mathbf{b}$. Of course, if we don't know A^{-1}, we can still solve the system in the usual way: Use Gaussian elimination to reduce $(A \mid \mathbf{b})$ to an equivalent augmented matrix in rref, $(I \mid \mathbf{s})$. The vector $\mathbf{s}$ on the right of the bar will be the unique solution to $A\mathbf{x} = \mathbf{b}$.

So far so simple. But now suppose we wanted to solve two systems: $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{x} = \mathbf{c}$. We could, of course, do this by reducing $(A \mid \mathbf{b})$ and $(A \mid \mathbf{c})$ to $(I \mid \mathbf{s})$ and $(I \mid \mathbf{s}')$ respectively. Then $\mathbf{s}$ and $\mathbf{s}'$ would be the solutions to the two systems. This approach, however, would entail needlessly repeating ourselves, since the row operations we'd use to reduce A to I in the first instance would be precisely the same that we'd use in the second instance. A more efficient use of our time would be to apply Gaussian elimination to the *doubly*-augmented matrix $(A \mid \mathbf{b} \;\; \mathbf{c})$, with two vectors on the right-hand side. The result will end up being $(\, I \mid \mathbf{s} \;\; \mathbf{s}')$, from which we can simultaneously read the solutions to both systems.

With those preliminary observations in hand, we can work our way towards the inversion algorithm. I will describe it first in terms of a 3×3 matrix, but the logic will hold for any square matrix whatsoever. Given an invertible matrix A, we want to find its inverse, which will be some matrix

$$A^{-1} = \begin{pmatrix} | & | & | \\ \mathbf{x_1} & \mathbf{x_2} & \mathbf{x_3} \\ | & | & | \end{pmatrix}$$

with the property that $AA^{-1} = I$. That is, we must find three column vectors $\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}$ such that

$$A \begin{pmatrix} | & | & | \\ \mathbf{x_1} & \mathbf{x_2} & \mathbf{x_3} \\ | & | & | \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

How do we do this? Well, by definition of matrix-matrix multiplication (from the column perspective), we can rewrite that last equation as

$$\begin{pmatrix} | & | & | \\ A\mathbf{x_1} & A\mathbf{x_2} & A\mathbf{x_3} \\ | & | & | \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For this to hold, the corresponding columns of the two matrices must all be equal. That is, we must have $A\mathbf{x_1} = \mathbf{e_1}$, $A\mathbf{x_2} = \mathbf{e_2}$, and $A\mathbf{x_3} = \mathbf{e_3}$. By our preliminary observation above, we can solve all three equations simultaneously by doing Gaussian elimination on the triply augmented matrix $(A \mid \mathbf{e_1} \;\; \mathbf{e_2} \;\; \mathbf{e_3})$.

The result will have the form $(I \mid \mathbf{s_1} \;\; \mathbf{s_2} \;\; \mathbf{s_3})$, where the three vectors on the right will be the solutions to $A\mathbf{x_1} = \mathbf{e_1}$, $A\mathbf{x_2} = \mathbf{e_2}$, and $A\mathbf{x_3} = \mathbf{e_3}$ respectively. In other words, the three vectors on the right will be… the three columns of the inverse matrix A^{-1} that we seek. And with that we are done. Well, almost done: We can tidy this up a bit by observing that the triply augmented matrix $(A \mid \mathbf{e_1} \;\; \mathbf{e_2} \;\; \mathbf{e_3})$ can be rewritten as $(A \mid I)$. Thus ends our derivation of the inversion algorithm.

Matrix Inversion Algorithm.

If A is a square matrix, form the augmented matrix $(A \mid I)$ and perform Gaussian elimination until the left-hand side is I. The right-hand side will then be A^{-1}.

Recall from Chapter 3's Exercise 23 that not every square matrix is invertible. If you try to perform the inversion algorithm on a noninvertible matrix, you'll soon discover that you can't carry the algorithm out because A can't be reduced to I through row operations.*

Example. Find the inverse of $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 8 & 2 \end{pmatrix}$.

Solution. Following the algorithm described above we create the augmented matrix $(A \mid I)$:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ 3 & 8 & 2 & 0 & 0 & 1 \end{array}\right).$$

Carrying out Gaussian elimination (the details of which I will leave to you), this reduces to

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 10 & -6 & 1 \\ 0 & 1 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & -7 & 5 & -1 \end{array}\right).$$

We therefore conclude that

$$A^{-1} = \begin{pmatrix} 10 & -6 & 1 \\ -2 & 1 & 0 \\ -7 & 5 & -1 \end{pmatrix},$$

which you may verify by computing $A^{-1}A$ and AA^{-1}. You'll find that both products equal I. ♦

---

* **Theorem.** If A is any square matrix, A is invertible $\Leftrightarrow$ $\text{rref}(A) = I$.

Sketch of Proof. We'll show that both statements are equivalent to a third: $A\mathbf{x} = \mathbf{b}$ has a *unique* solution for every vector $\mathbf{b}$.

It's easy to see that this third statement is equivalent to A being invertible. After all, if A is invertible, then we can obtain the unique solution to $A\mathbf{x} = \mathbf{b}$ by left-multiplying both sides by A^{-1}. Conversely, if $A\mathbf{x} = \mathbf{b}$ has a *unique* solution for every $\mathbf{b}$, then A's action can *always* be "undone", which is, of course, what it means for A to be invertible.

To see that the third statement is equivalent to $\text{rref}(A) = I$, begin with a familiar fact: Using Gaussian elimination, we can boil every system $A\mathbf{x} = \mathbf{b}$ down to $(\text{rref}(A) \mid \mathbf{b}')$, for some vector $\mathbf{b}'$. When we do so…

(1) If $\text{rref}(A) = I$, the system obviously has a **unique solution**. (Namely, $\mathbf{b}'$.)

(2) If $\text{rref}(A) \neq I$, then $\text{rref}(A)$ must either have a nonzero entry to the right of a pivot, or a zero row – or perhaps both. A nonzero entry to a pivot's right corresponds, of course, to a free variable in the linear system; hence, if the system is consistent (has at least one solution), it has *infinitely many* solutions. Similarly, if $\text{rref}(A)$ has a zero row, the system is either inconsistent or has infinitely many solutions (depending on whether the corresponding entry of $\mathbf{b}'$ is 0 or not). At any rate, if $\text{rref}(A) \neq I$, the system **does *not* have a unique solution**.

It follows that ($A\mathbf{x} = \mathbf{b}$ has a *unique* solution for every vector $\mathbf{b}$) $\Leftrightarrow$ $\text{rref}(A) = I$.

Since both statements in our theorem are equivalent to a third, they are equivalent to one another, as claimed. ■

Exercises.

28. Use the inversion algorithm to invert the following 3×3 matrices by hand:

$$A = \begin{pmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 & 2 \\ -3 & 1 & 4 \\ 2 & -3 & 4 \end{pmatrix}.$$

29. There's a simple formula for the inverse of a 2×2 matrix:

$$\text{If } A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \quad \text{then} \quad A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}.$$

(If $ad - bc = 0$, then it means that matrix A isn't invertible.)

a) Prove that this formula works as advertised by showing that $A^{-1}A = I$.

b) Commit this formula to memory! Note the pattern: When we invert, the two "off-diagonal" entries sprout negatives, the two entries on the main diagonal switch places, and the whole thing is divided by $ad - bc$, an important quantity in its own right that we'll meet again in the next chapter.

c) Using the formula you've just memorized, find the inverses of the following 2×2 matrices – if they exist:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}, \qquad C = \begin{pmatrix} 8 & 3 \\ 7 & 2 \end{pmatrix}, \qquad D = \begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix}.$$

30. Inverting large matrices by hand is appallingly tedious, but here's one exception to that rule:
A **diagonal matrix** is a square matrix whose only nonzero entries (if it has any) lie on its main diagonal.
Diagonal matrices of any size are easy to invert (and to recognize as non-invertible when that is the case).

a) Explain geometrically why any diagonal matrix that includes a zero on the diagonal is *not* invertible.

b) Explain geometrically why any diagonal matrix whose diagonal entries are all nonzero is invertible.

c) If $A = \begin{pmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$, find A^{-1} by thinking geometrically about what A does to $\mathbb{R}^5$.

(You could do this with the inversion algorithm, but thinking geometrically is much better.)

d) What does the inverse of a diagonal matrix (with no zeros on the main diagonal) look like?

e) Another nice feature of diagonal matrices: raising them to the n^{th} power is easy. Explain geometrically why we can obtain the n^{th} power of a diagonal matrix simply by raising all its diagonal entries to the n^{th} power. Then compute A^3, where A is the matrix from Part C.

31. For any $n \times n$ matrix A, the five statements below are logically equivalent, meaning they stand or fall together. That is, if A satisfies any of these five conditions, it satisfies them all. If it fails to satisfy one, it satisfies none.
Your problem: *Convince yourself that this is so.*

a) A is invertible. b) $\text{rref}(A) = I$. c) $A\mathbf{x} = \mathbf{b}$ has a *unique* solution for every vector $\mathbf{b}$.
d) A's columns are linearly independent. e) A's columns span $\mathbb{R}^n$. f) A's columns constitute a basis for $\mathbb{R}^n$.

This list is the beginning of a huge list of equivalent statements sometimes called **The Invertible Matrix Theorem**. This theorem tells us that square matrices come in two basic varieties: invertible matrices, which satisfy *all* the list's conditions, and noninvertible ones, which satisfy *none* of them. We'll extend the list in future exercise sets. The next extension will occur in Exercise 40 of this chapter.

Linear Relations Among the Columns

I am the family face;
Flesh perishes. I live on.
- Thomas Hardy, "Heredity"

Gaussian elimination isn't just for *augmented* matrices. For the rest of this chapter, we'll see how to glean important information about a linear map by applying Gaussian elimination to its matrix representation.

We'll begin with an important technical result: Just as doing row operations on an augmented matrix preserves the underlying linear system's *solutions*, doing row operations on an ordinary matrix preserves *linear relationships among the columns*. For example, if we denote a matrix A's columns by $\mathbf{a_1}, \mathbf{a_2}, \dots, \mathbf{a_5}$, and it just so happens that $2\mathbf{a_1} + 3\mathbf{a_2} - \mathbf{a_4} = 6\mathbf{a_5}$, I claim that this linear relationship will still hold among the columns of any matrix B that we obtain from A with row operations. That is, even if the row operations scramble each of A's columns beyond recognition, the resulting matrix B will still exhibit the property that twice its first column plus thrice its second minus its fourth will be… six times its fifth column. It's an old family trait like a cleft chin or broad forehead.

To see why, push everything in that relationship to one side, making it $2\mathbf{a_1} + 3\mathbf{a_2} - \mathbf{a_4} - 6\mathbf{a_5} = \mathbf{0}$. Next, reformulate this last equation (matrix-vector multiplication to the rescue!) in another way:

$$\begin{pmatrix} 2 \\ 3 \\ 0 \\ -1 \\ -6 \end{pmatrix} \text{ is a solution to the matrix-vector equation } A\mathbf{x} = \mathbf{0}.$$

Now suppose we do some row operations to the augmented matrix $(A|\mathbf{0})$, turning it into $(B|\mathbf{0})$.* As we know, those row operations preserve the underlying system's solutions, so the column vector above must also be a solution of the corresponding equation $B\mathbf{x} = \mathbf{0}$. In other words, it must be the case that

$$\begin{pmatrix} | & | & | & | & | \\ \mathbf{b_1} & \mathbf{b_2} & \mathbf{b_3} & \mathbf{b_4} & \mathbf{b_5} \\ | & | & | & | & | \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 0 \\ -1 \\ -6 \end{pmatrix} = \mathbf{0}.$$

If we multiply the left-hand side out and then shift its last term over to the right-hand side of the equation, we obtain $2\mathbf{b_1} + 3\mathbf{b_2} - \mathbf{b_4} = 6\mathbf{b_5}$. That is, twice B's first column plus thrice its second column minus its fourth column will be… six times its fifth column. Hence, as claimed, the "family trait" exhibited among A's columns is still present in the columns of B, the matrix we obtained from A via row operations.

There is nothing special about the particular linear relationship with which we've just been playing. The same argument clearly holds for *any* linear relationship among a matrix's columns. Now that you've seen *why* it holds, we can state our important result, all boxed up like a birthday present:

Row operations preserve linear relationships among a matrix's columns.

* It's easy to see that row operations preserve $\mathbf{0}$: Swapping rows just exchanges two of $\mathbf{0}$'s zeros, preserving $\mathbf{0}$. Multiplying a row by a constant multiplies one of $\mathbf{0}$'s zero entries by a constant. The entry remains 0, so $\mathbf{0}$ is preserved. Finally, adding a multiple of one row to another just adds a multiple of one of zero to another, which of course, changes nothing, preserving $\mathbf{0}$ as claimed.

An example will make the idea more concrete.

Example 1. Consider the "telephone matrix",

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

Observe that the sum of the outer columns is twice the middle column. Or, to describe this linear relationship more crisply: the middle column is the *average* of the outer columns. We'll now carry out some Gaussian elimination, pausing along the way to verify that this linear relationship persists throughout the process. We'll begin by clearing out the first column:

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{matrix} \\ -4R_1 \\ -7R_1 \end{matrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix}.$$

The matrix on the right is already quite different from A, but it has retained the telltale family trait: its middle column is the average of the outer columns. Boiling the matrix all the way down to reduced row echelon form, we obtain

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix} \begin{matrix} \\ \div(-3) \\ \div(-6) \end{matrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \begin{matrix} -2R_2 \\ \\ -R_2 \end{matrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

As you should verify, in each matrix, the middle column is the average of the outer columns. ♦

We often call linear relationships among vectors linear *dependencies*, because a linear relationship holds among a vector set if and only if its vectors are linearly *de*pendent.* The boxed result on the previous page helps us identify linearly dependent groups of columns in a matrix, because once we boil a matrix down to rref, any linear dependencies among its columns will become obvious. This will be important for us because, as we saw in Chapter 3, linear dependencies among the columns of matrix imply some form of dimensional collapse in the underlying linear map's output. I'll have more to say about this in the next section. For now, I'll just repeat this section's main result, rephrasing it slightly to reflect the terminology and the setting in which we'll most commonly use it:

Gaussian elimination preserves linear dependencies (or lack thereof) among a matrix's columns.

Exercises.

32. Are linear relationships among the *rows* of a matrix preserved by row operations?
If yes, provide a proof. If no, provide a counterexample.

33. Let A be a 100×100 matrix whose 42^{nd} column is three times its 6^{th} column plus twice its 7^{th} column.
a) What, if anything, can we conclude about the 42^{nd} column of $\text{rref}(A)$?
b) Is A invertible?

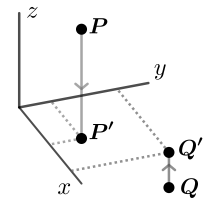
* This is nothing deep. It's just saying that given any linear relationship (say, $2\mathbf{a} + 3\mathbf{b} = \mathbf{c} - 4\mathbf{d}$), we can obviously take any one of the involved vectors and isolate it on one side of the equation, thus demonstrating that it lies in the other vectors' span. Hence the vectors are linearly dependent. On the other hand, to say that a set of vectors is linearly *in*dependent is to say that there are *no* linear relationships among them.

Image and Kernel

God created man in his image.
- Genesis 1:27

Every linear map – or matrix – determines two special subspaces: the map's **image**, which is defined as the set of *all the map's output vectors* (its "global output"), and the map's **kernel**, which is defined as the set of *all of the input vectors that the map sends to* $\mathbf{0}$.*

Example 1. Consider the linear transformation of $\mathbb{R}^3$ that orthogonally projects each point in space onto the xy-plane. This map's image is the xy-plane itself, because every point therein is clearly an output of the projection (and vice-versa). The map's kernel is the z-axis, since all points on that axis – and no other points – are orthogonally projected to the origin. ♦



The image and kernel are subspaces, so each takes one of the usual geometric forms: the origin alone (a 0-dimensional subspace), a line through the origin (1-dimensional subspace), a plane through the origin (2-dimensional subspace), or more generally, an n-dimensional hyperplane through the origin.

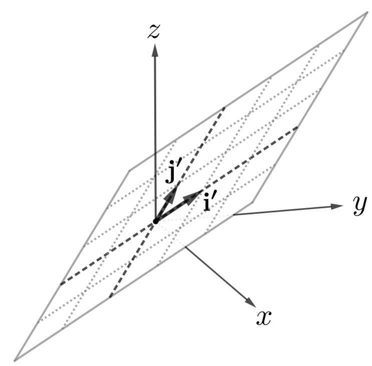
A map's **rank** is the dimension of its image. A map's **nullity** is the dimension of its kernel. (Thus, the projection in Example 1 has rank 2 and nullity 1.) The rank and nullity indicate, respectively, how many of the domain's dimensions survive or collapse when subjected to the transformation.

Example 2. The 3×2 matrix

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

defines a linear map from $\mathbb{R}^2$ to $\mathbb{R}^3$. Its columns are obviously linearly independent, so they span a plane in $\mathbb{R}^3$, as at right. This plane is A's image. On the other hand, A's kernel is $\mathbf{0}$, since any *non*zero vector $a\mathbf{i} + b\mathbf{j}$ in $\mathbb{R}^2$ is sent to a nonzero vector in $\mathbb{R}^3$; namely, to $a\mathbf{i}' + b\mathbf{j}'$.

Thus, this matrix has rank 2 and nullity 0. ♦



Given a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$, all n of the domain's dimensions may survive the transformation, or some of them, say k of them, may be "killed" (making the map's nullity k and its rank $(n-k)$). Either way, the rank and nullity must sum to n, the dimension of the map's domain. This fact is known as

> **The Rank-Nullity Theorem.**
> Given any linear map, the sum of its rank and nullity is the dimension of the map's domain.

* Why is a linear map's image a *subspace*? Well, if A represents the map, any two vectors in its image have the form $A\mathbf{v}$ and $A\mathbf{w}$. Any linear combination of them has the form $c(A\mathbf{v}) + d(A\mathbf{w})$, which is equivalent to $A(c\mathbf{v} + d\mathbf{w})$ (by Chapter 3's Exercise 14c), which obviously lies in A's image. Thus, A's image is closed under linear combinations, so it is a subspace as claimed.

Similarly, if $\mathbf{v}$ and $\mathbf{w}$ are any two vectors in A's kernel, any linear combination of them, $c\mathbf{v} + d\mathbf{w}$, will be mapped to $\mathbf{0}$ by A, since $A(c\mathbf{v} + d\mathbf{w}) = cA\mathbf{v} + dA\mathbf{w} = c\mathbf{0} + d\mathbf{0} = \mathbf{0}$. Since A's kernel is thus closed under linear combinations, it a subspace, too.

The examples from the previous page were easy to visualize, but linear maps need not be so obliging. How, in general, can we find a linear map's image and kernel? Given, say, a random-looking 60×70 matrix (a map from $\mathbb{R}^{70}$ to $\mathbb{R}^{60}$), some "dimensional collapse" (at least 10 dimensions) must occur since the contents of the 70-dimensional domain are being stuffed into the 60-dimensional range, but precisely how much? Will the map's image be *all* of $\mathbb{R}^{60}$, or will its image be some subspace of $\mathbb{R}^{60}$? And what will the map's kernel, a subspace of $\mathbb{R}^{70}$, look like? The rank-nullity theorem implies that the kernel must be a hyperplane (of 10+ dimensions) through the origin (since it is a subspace), but can we describe its configuration in space more precisely? Can we, for example, provide a *basis* for the kernel?

Before we dig into a specific example, let's view the "kernel and image problem" from 10,000 feet. Let A be any matrix. Finding A's *kernel* is simple: By definition, we need only solve the system $A\mathbf{x} = \mathbf{0}$, a routine task for Gaussian elimination. Finding a streamlined description of A's *image* (in terms of a basis) is a different story. We don't find it by solving a system. We find it by *weeding*. Observe that A's image is the span of its columns. (This follows directly from matrix-vector multiplication's column perspective.) Hence, if we start with A's columns, which span A's image, and weed out those that are linearly dependent on their predecessors, the columns that remain will be a *linearly independent* spanning set, and thus a *basis* for A's image. That's a simple enough plan in principle, but how do we spot A's redundant columns? With a very clever idea: We turn our gaze, surprisingly, away from A and look instead at its row-reduced descendant, $\text{rref}(A)$; the columns of this descendant matrix will exhibit the same linear dependencies as those of its ancestor A (as we saw in the previous section), but they will be *much* easier to spot on $\text{rref}(A)$ than on A itself. Having spotted them on $\text{rref}(A)$, we'll immediately know where to look for them on A. We then weed those redundant columns out of our spanning set, and are left with a basis for A's image.

Let's see how these ideas play out in practice on a specific extended example. If you understand it, you'll understand how to solve any such problem.

Example 3. Find the image and kernel of the linear map represented by the following matrix. Express each subspace in terms of a basis.

$$M = \begin{pmatrix} 1 & -2 & 16 & 9 & 0 & 5 \\ 2 & 6 & -38 & -2 & 1 & -70 \\ 3 & 1 & -1 & 13 & 14 & -41 \\ 3 & -2 & 20 & 19 & 1 & -17 \end{pmatrix}.$$

Solution. This represents a map from $\mathbb{R}^6$ to $\mathbb{R}^4$, so its image can be at most four-dimensional. Therefore, according to the rank-nullity theorem, its kernel must be *at least* two-dimensional. This much is true of *any* 4×6 matrix, but what more can we say about this particular matrix?

Let's begin with M's kernel, which we denote $\mathbf{ker}(\boldsymbol{M})$. It is defined as all the solutions to the system $M\mathbf{x} = \mathbf{0}$. We can find these solutions by setting up the augmented matrix $(M|\mathbf{0})$ and doing Gaussian elimination. I'll leave those boring details to you, but the result, as you should verify, is

$$\text{rref}(M|\mathbf{0}) = \left(\begin{array}{cccccc|c} 1 & 0 & 2 & 5 & 0 & -11 & 0 \\ 0 & 1 & -7 & -2 & 0 & -8 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Earlier, you learned how to extract the solutions from an augmented matrix in rref and write the solutions as linear combinations of constant vectors. Let's do that here, using the time-saving trick I demonstrated near the end of the "Endgames and Algorithm" section. We'll begin by marking up the augmented matrix a bit, labelling its columns and circling its pivots, now that it's in reduced row echelon form:

$$\begin{array}{c} \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{matrix} \\ \left(\begin{array}{cccccc|c} \textcircled{1} & 0 & 2 & 5 & 0 & -11 & 0 \\ 0 & \textcircled{1} & -7 & -2 & 0 & -8 & 0 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right). \end{array}$$

The three pivotless columns indicate that x_3, x_4, and x_6 are free variables. Calling them s, t, u, respectively and sticking those parameters into our solution vector first, we then fill the remaining slots in with the usual mental algebra. We'll then decompose the result to clarify its geometry. In the end, this process yields, as you should verify,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} -2s - 5t + 11u \\ 7s + 2t + 8u \\ s \\ t \\ 0 \\ u \end{pmatrix} = s\begin{pmatrix} -2 \\ 7 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -5 \\ 2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + u\begin{pmatrix} 11 \\ 8 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

We therefore conclude that

$$\mathbf{ker}(\boldsymbol{M}) = s\begin{pmatrix} -2 \\ 7 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -5 \\ 2 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + u\begin{pmatrix} 11 \\ 8 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

where s, t, u can be any real numbers.

So far so easy. Now let's find M's image, which we denote $\mathbf{im}(\boldsymbol{M})$. This takes more thought. As I described it in our "view from 10,000 feet" just before this example, finding a streamlined description of a matrix's image is usually a matter of weeding. The columns of M span the image, so to get a *basis* for the image, we just weed this spanning set, throwing away any columns that are linearly dependent on their predecessors (and thus contribute nothing new to the image). When we look at M, it's not at all obvious which – if any – columns should be weeded. But when we look at $\text{rref}(M)$, it will be. Happily, we already produced $\text{rref}(M)$ while computing the kernel, because $\text{rref}(M|\mathbf{0})$, which we found on the way to the kernel, is, of course, $(\text{rref}(M)|\mathbf{0})$. Thus,

$$\text{rref}(M) = \begin{pmatrix} 1 & 0 & 2 & 5 & 0 & -11 \\ 0 & 1 & -7 & -2 & 0 & -8 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Note well: $\text{rref}(M)$ is not M. It is a different matrix, corresponding to a different map. But as M's descendant via Gaussian elimination, it retains crucial ancestral information about M itself. Specifically, we know that any linear dependencies – or lack thereof – among the columns of $\text{rref}(M)$ were *inherited*, and thus must hold in the columns of M as well.

In $\text{rref}(M)$, the columns with pivots (columns 1, 2, and 5) are obviously linearly independent, so M's corresponding columns (columns 1, 2, and 5) must be linearly independent, too.

Moreover, $\text{rref}(M)$'s pivotless columns (columns 3, 4, and 6) clearly lie in the span of its first two columns. (Its 3rd column, for example, is two times its 1st column minus seven times its 2nd.) These same linear dependencies must also hold among the corresponding columns of M itself. I encourage you to verify this; watching the abstract magic play out concretely helps us grasp it.

We've learned that M sends three of $\mathbb{R}^6$'s standard basis vectors ($\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_5$) to a trio of *linearly independent* vectors in $\mathbb{R}^4$: columns 1, 2, and 5 of M. Each of those columns contributes a dimension to M's image, and collectively, they span a 3-dimensional hyperplane in $\mathbb{R}^4$. Meanwhile, M sends the other three standard basis vectors ($\mathbf{e}_3, \mathbf{e}_4$, and $\mathbf{e}_6$) to vectors (columns 3, 4, and 6 of M) that already lie in the hyperplane. Hence, those columns contribute nothing new to M's global output. As far as M's image is concerned, they are redundant. It follows that M's image is the three-dimensional hyperplane spanned by its 1st, 2nd, and 5th columns.

Specifically, since M's 1st, 2nd, and 5th columns constitute a basis for $\text{im}(M)$, it follows that:

$$\mathbf{im}(\boldsymbol{M}) = t_1\begin{pmatrix}1\\2\\3\\3\end{pmatrix} + t_2\begin{pmatrix}-2\\6\\1\\-2\end{pmatrix} + t_3\begin{pmatrix}0\\1\\14\\1\end{pmatrix},$$

where t_1, t_2, t_3 can be any real numbers. (M's rank is thus 3.) ♦

I'll summarize our results a box. Be sure that you understand the ideas behind them.

> **Summary** Given any matrix A,
>
> To find A's **kernel**: Solve $\boldsymbol{A}\mathbf{x} = \mathbf{0}$.
>
> To find A's **image**: Identify any columns that are linearly dependent on their predecessors and "weed them out". Those that remain constitute a basis for $\text{im}(A)$.
>
> *In practice, this means finding the pivot columns of* $\text{rref}(A)$. *The corresponding columns of A are a basis for* $\text{im}(A)$.

In the next section, we'll see why the peculiar name "kernel" is appropriate. But first, some exercises.

Exercises.

34. For each of the following transformations, describe the kernel and image as specifically as possible. State the rank and nullity, too. These should all be done just by thinking geometrically. No computations.

a) Rotation of $\mathbb{R}^2$ around the origin by $30°$.
b) The zero map defined on $\mathbb{R}^3$. (That is, every point in $\mathbb{R}^3$ gets mapped to the origin).
c) The map in $\mathbb{R}^2$ that sends every point to its orthogonal projection onto the line $y = x$.
d) Reflection of $\mathbb{R}^4$ over the hyperplane $w = 0$ (where we're calling $\mathbb{R}^4$'s four variables x, y, z, and w).
e) The identity map on $\mathbb{R}^n$.
f) The map from $\mathbb{R}^3$ to $\mathbb{R}$ that sends each vector in $\mathbb{R}^3$ to its first coordinate.

35. Row operations preserve *linear dependencies among the columns* (and *solutions of systems of linear equations*), but they change most other things about a matrix. To reinforce this fact, convince yourself that

a) Row operations can change a matrix's geometric effect (i.e. what the matrix does as a linear map).
[Ex: *If you use Gaussian elimination to put a rotation matrix into rref, does the result still rotate points*?]

b) Row operations can change a matrix's image.
[Ex: *Make up a* 2×2 *matrix* A *with linearly dependent columns. It's geometrically clear that* $\mathrm{im}(A)$ *is a line*. *Draw a picture of it. Now multiply one of* A*'s rows by a constant. What is the image of this new matrix? A line, yes, but is it the same line?*]

Still, the preservation of linear dependencies among columns and solutions to linear systems ensure that other (ostensibly more complicated) things are preserved by row operations, too. Convince yourself that

c) Although row operations can change a matrix's image (as we just saw in Part B), they do preserve its dimension. That is, row operations preserve a matrix's *rank*.

d) Although row operations can change a matrix's image, they preserve the matrix's kernel. Explain why.

36. For each of the following, *compute* the kernel and image using the techniques described in this section's boxed summary. Then go back and see which images and kernels you could have found by thinking geometrically.

a) $\begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$ b) $\begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix}$ c) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ d) $\begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$ e) $\begin{pmatrix} 4 & 1 \\ 3 & 1 \\ 1 & 2 \\ 2 & 0 \end{pmatrix}$ f) $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$

37. True or false – and explain. For any matrix A,

a) $\ker(A) = \ker(\mathrm{rref}(A))$ b) $\mathrm{im}(A) = \mathrm{im}(\mathrm{rref}(A))$ c) A's rank equals the number of pivots in $\mathrm{rref}(A)$.*
d) The pivot columns of $\mathrm{rref}(A)$ constitute a basis for $\mathrm{im}(A)$.
e) The columns of A corresponding to $\mathrm{rref}(A)$'s pivot columns constitute a basis for $\mathrm{im}(A)$.
f) The columns of A corresponding to $\mathrm{rref}(A)$'s *non*-pivot columns constitute a basis for $\ker(A)$.

38. A matrix's image and kernel are also called its **column space** and **null space**. Why do these names make sense?

39. The columns in every 3×3 matrix whose kernel is a plane have something in common. What is it? And why?

40. In Exercise 31, you first met the **Invertible Matrix Theorem**. In this exercise, we'll extend its list of equivalent statements about an $n \times n$ matrix A. You've already seen in Exercise 31 why the first five below are equivalent. Now convince yourself that the last three can join the list. (We'll extend the list again in Chapter 5, Exercise 8.)

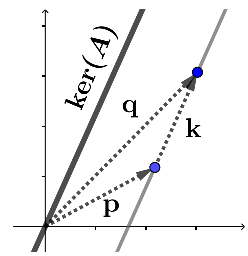
a) A is invertible. b) $\mathrm{rref}(A) = I$ c) $A\mathbf{x} = \mathbf{b}$ has a *unique* solution for every vector $\mathbf{b}$.
d) A's columns are linearly independent. e) A's columns span $\mathbb{R}^n$. f) A's columns constitute a basis for $\mathbb{R}^n$.
g) $\ker(A) = \mathbf{0}$. h) $\mathrm{im}(A) = \mathbb{R}^n$ i) $\mathrm{rank}(A) = n$.

* Some textbooks *define* A's rank as the number of pivots in $\mathrm{rref}(A)$, thus putting the algebraic cart before the geometric horse.

Why 'Kernel' is an Apt Name

What's in a name? That which we call a rose
By any other name would smell as sweet.
- Juliet, *Romeo and Juliet*, Act II, Scene 2.

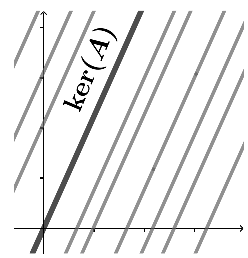
When a map's kernel is more than just $\mathbf{0}$ (i.e. if the kernel is a line, plane, or hyperplane), the rank-nullity theorem implies that the map must collapse at least one dimension of the domain, yielding an image with fewer dimensions than the domain had. When this happens, the kernel offers us a surprisingly precise blueprint of how the transformation operates on the domain's points.



Consider a map A whose domain is $\mathbb{R}^2$ and whose kernel is a line, as at right. By definition, all $\ker(A)$'s points get mapped to $\mathbf{0}$. More interestingly, if we take any "parallel copy" of the kernel (such as the figure's other line), all of its points get mapped to the same point, too - just not to $\mathbf{0}$. To understand why, take any two points $\mathbf{p}$ and $\mathbf{q}$ on the parallel copy. Because the copy is parallel to the kernel, it's geometrically clear that $\mathbf{q} = \mathbf{p} + \mathbf{k}$, for some vector $\mathbf{k}$ in $\ker(A)$. It follows that

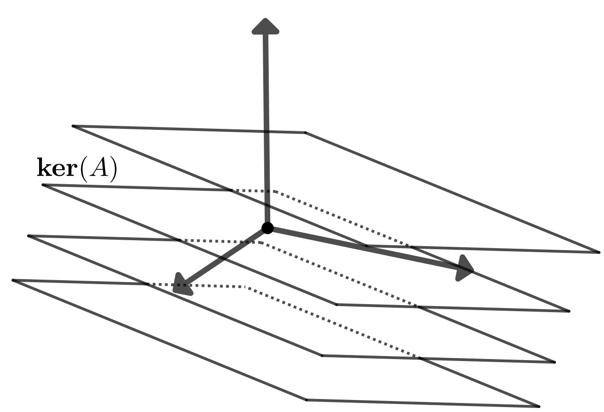
$$\begin{aligned} A\mathbf{q} &= A(\mathbf{p}+\mathbf{k}) = A\mathbf{p} + A\mathbf{k} \\ &= A\mathbf{p} \qquad \text{(since } \mathbf{k} \text{ is in the kernel).} \end{aligned}$$

That is, A maps $\mathbf{q}$ to the same point in the range as it maps $\mathbf{p}$. And since $\mathbf{q}$ was just an arbitrary point on the line, it follows that A maps *every* point on the parallel copy to the same point in the range, as claimed.

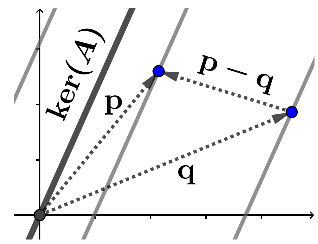


Of course, there are infinitely many copies of the kernel in the map's domain, and A will map all the points on any chosen copy to the same point in the range. These copies of the kernel are sometimes called, quite evocatively, the map's *fibers*, as if they were the metaphorical fabric from which the map's domain was woven. Every point in the domain belongs to one and only one fiber, and the map sends all points of any given fiber to the same point in the range. We can even think of the map as acting on *the fibers themselves* rather than on their individual points. We won't need this perspective in this course, but it bears fruit in the more advanced subject of *abstract algebra*.

The fibers need not be lines. If, say, the map's kernel is two-dimensional, it and all its fiberly brethren will be *planes*. But the same idea holds: The map's domain will be partitioned into a family of planes, all parallel to the kernel, with the property that the map sends all the points on each plane to the same point in the map's image. Such a transformation, effectively turns these particular *planes* (in the domain) into *points* (in the range). You should take a minute to convince yourself that the geometric argument we used when the kernel was a line still holds when the kernel is a plane. (That is, explain to yourself why A maps any two points $\mathbf{p}$ and $\mathbf{q}$ on the same "parallel copy of the kernel" to the same point.) Of course, a map whose domain is $\mathbb{R}^3$ may also have a *line* as its kernel, in which case all of space will be partitioned into a bundle of parallel lines. And if the domain has 4 or more dimensions, then the map's fibers could even be hyperplanes.



Do distinct fibers always get mapped to distinct points in the map's range? Or can A map two points from distinct fibers to the same point in the range? A little playing around will reveal the answer. Suppose $\mathbf{p}$ and $\mathbf{q}$ are points on *distinct* fibers. Then the line on which the vector $\mathbf{p}-\mathbf{q}$ lies clearly cuts *across* the fibers. In other words, vector $\mathbf{p}-\mathbf{q}$ (translated to put its tail at the origin) isn't in $\ker(A)$. Thus, $A(\mathbf{p}-\mathbf{q}) \neq \mathbf{0}$. But this is equivalent to $A\mathbf{p} - A\mathbf{q} \neq \mathbf{0}$, which implies that $A\mathbf{p} \neq A\mathbf{q}$. That is, if $\mathbf{p}$ and $\mathbf{q}$ are points on distinct fibers, then A maps them to distinct points.



To sum up, every matrix A represents a linear map. The map's domain can be partitioned into infinitely many parallel copies of $\ker(A)$. These are the map's fibers. All points on any given fiber get mapped to the same point in the range. Moreover, no other points in the domain get mapped to that point. Hence, just as $\ker(A)$ consists of all solutions to $A\mathbf{x} = \mathbf{0}$, each fiber – each parallel copy of the kernel – consists of all the solutions to $A\mathbf{x} = \mathbf{b}$ for some particular point $\mathbf{b}$ in the transformation's range.

This tells us something remarkable and useful about linear systems in general. To find all the solutions to a linear system $A\mathbf{x} = \mathbf{b}$, it suffices to find one particular solution $\mathbf{p}$ and then "add the kernel" to $\mathbf{p}$. By "adding the kernel", I mean producing the set of all vectors of the form $\mathbf{p} + \mathbf{k}$, where $\mathbf{k}$ is in the kernel. It is in this sense that the name "kernel" (as in *core*, not corn) is apt. For we now see that the kernel of a matrix A lies at the core of the solution set to *every* linear system $A\mathbf{x} = \mathbf{b}$.

This result is worth boxing up and presenting formally as a theorem.

> **Theorem.** The set of *all* solutions to $A\mathbf{x} = \mathbf{b}$ consists of all vectors of the form $\mathbf{p} + \mathbf{k}$, where $\mathbf{p}$ is any particular solution, while $\mathbf{k}$ ranges over $\ker(A)$.

Example. Let $A = \begin{pmatrix} 2 & 6 & -4 \\ 4 & 0 & -2 \\ 2 & 4 & -3 \end{pmatrix}$, and $\mathbf{b} = \begin{pmatrix} 4 \\ 8 \\ 4 \end{pmatrix}$.

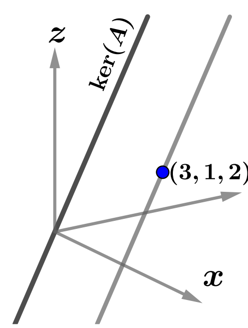
As you can verify with Gaussian elimination,

$$\ker(A) = t\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \text{ where } t \text{ ranges over the reals.}$$

Thus A's kernel is a line through $\mathbb{R}^3$'s origin, as shown in the figure.

Suppose we happen to notice that

$$\mathbf{p} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \text{ is one particular solution of } A\mathbf{x} = \mathbf{b}.$$



That is, suppose we notice that A maps $\mathbf{p}$ to $\mathbf{b}$. It follows from our work in this section that the *full* solution set to $A\mathbf{x} = \mathbf{b}$ consists of all points on the line passing through $\mathbf{p}$ parallel to the kernel. In other words, by "adding the kernel" to our particular solution, we obtain the full solution set:

$$\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + t\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \text{ where } t \text{ ranges over the reals.}$$ ♦

Exercises.

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

a) Find $\text{im}(A)$ and $\ker(A)$, and graph both of these subspaces on a set of axes.

b) Observe that A maps $2\mathbf{i}$ to $4\mathbf{i} + 2\mathbf{j}$. Without doing row-reduction, find *all* solutions to $A\mathbf{x} = 4\mathbf{i} + 2\mathbf{j}$. Add a graph of these solutions to your axes from Part A. Label this new line "solutions to $A\mathbf{x} = 4\mathbf{i} + 2\mathbf{j}$".

c) Observe that A maps $-\mathbf{j}$ to $-6\mathbf{i} - 3\mathbf{j}$. Using the preceding section's theorem, find the set of all vectors that A maps to $-6\mathbf{i} - 3\mathbf{j}$. Add them to your graph. Label this new line "solutions to $A\mathbf{x} = -6\mathbf{i} - 3\mathbf{j}$".

d) Draw the line through point $(1, 2)$ parallel to $\ker A$. Every point on this line is mapped to the same place by A. To which point are they all sent? Verify this analytically by writing down a vector expression for the line and showing that A maps every point on the line (i.e. every vector whose tip is on the line) to the same place.

42. (Eyeballing the kernel) You've learned how to find a matrix's kernel through row reduction. In this exercise, I'll introduce an alternate method that works only rarely, but it's very convenient when it does! It will work for any matrix whose columns' linear dependencies we can spot *without* doing row reduction.

The technique, which we might call "eyeballing the kernel" is best explained through an example: Let

$$M = \begin{pmatrix} 1 & 2 & 0 & 1 & 3 \\ 2 & 4 & 0 & 1 & 4 \\ 3 & 6 & 0 & 1 & 5 \\ 4 & 8 & 0 & 1 & 6 \end{pmatrix}.$$

Columns two and three are scalar multiples of column one, but column four is independent of its predecessors. The fifth column is the first column plus twice the fourth. To sum up, three of M's columns are "redundant", lying in the span of the other two. It follows that M's rank is 2. Hence, its nullity must be 3. Consequently, any three linearly independent vectors that M maps to $\mathbf{0}$ will constitute a basis for $\ker(M)$.

a) Be sure you understand why each deduction in the preceding paragraph makes sense.

b) Now be sure you understand the rest of the technique, and why it works:
If column i was found to be redundant, create a new vector as follows: Let its i^{th} entry be 1; then, thinking about the dependency that makes column i redundant (and about matrix-vector multiplication's definition), cook up the rest of the new vector's entries so that M will map the new vector to $\mathbf{0}$. Doing this here yields:

$$\mathbf{b_1} = \begin{pmatrix} -2 \\ \mathbf{1} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{b_2} = \begin{pmatrix} 0 \\ 0 \\ \mathbf{1} \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{b_3} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ -2 \\ \mathbf{1} \end{pmatrix}$$

Since these are linearly independent, $\ker M = t_1\mathbf{b_1} + t_2\mathbf{b_2} + t_3\mathbf{b_3}$ (where t_1, t_2, t_3 range over the reals).

c) Finding the image in an "eyeballing" case like this is easy. Here, what is $\text{im}\, M$? Justify your answer.

43. "Eyeball" the image and kernel. Express each as the span of a basis for the subspace.

a) $\begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$ b) $\begin{pmatrix} 3 & 1 & 6 \\ 5 & 2 & 10 \\ 5 & 3 & 10 \end{pmatrix}$ c) $\begin{pmatrix} 1 & 4 & 5 & 2 \\ 4 & 1 & 5 & 8 \end{pmatrix}$ d) $\begin{pmatrix} 4 & 0 & 1 \\ 0 & 1 & 2 \\ 2 & 0 & 4 \\ 0 & 0 & 5 \end{pmatrix}$ e) $(1 \quad 1 \quad -8)$

44. Stare at each linear system until a solution occurs to you. Then eyeball the kernel of the relevant matrix and add it to the particular solution, thus yielding *all* solutions to the system – without having to do Gaussian elimination.

a) $\begin{pmatrix} -4 & 12 \\ -7 & 21 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 8 \\ 14 \end{pmatrix}$ b) $\begin{pmatrix} 2 & 1 & 3 \\ 3 & 0 & 3 \\ 7 & 2 & 9 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 3 \\ 3 \\ 9 \end{pmatrix}$ c) $\begin{pmatrix} 6 & 6 & 1 & 4 \\ 2 & 2 & 0 & 2 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$.

Chapter 5
The Determinant

The Determinant: Definition

Once upon a midnight dreary, while I pondered, weak and weary,
over many a quaint and curious volume…

- Edgar Allen Poe, "The Raven"

The determinant's definition will initially look mysterious, but the clouds will soon clear.

Definition. If A is a square $n \times n$ matrix, its **determinant** ($\det A$) is the real number whose…

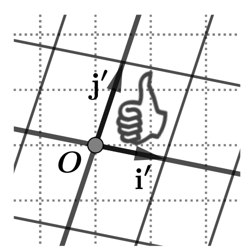
- *Magnitude* is the n-dimensional volume of the "box" in $\mathbb{R}^n$ that A's columns determine.
- *Sign* is positive if A preserves orientation, and negative if A reverses orientation.

I've used the term "box" to encompass parallelograms (in $\mathbb{R}^2$), parallelepipeds (in $\mathbb{R}^3$), and their higher-dimensional analogues (parallelotopes). Similarly, I've used "volume" to cover parallelograms' *areas* (in $\mathbb{R}^2$), parallelepipeds' volumes (in $\mathbb{R}^3$), and parallelotopes "hypervolumes" (in the appropriate spaces).

As for "orientation", recall that in Chapter 3 we considered a drawing of a right hand in $\mathbb{R}^2$, and we observed that certain maps (such as rotations) preserve the right hand's right-handedness, while others (such as reflections) transform it into a *left* hand. An orientation-*reversing* map of this latter sort sends the entire plane through the looking glass. This can occur in spaces of any number of dimensions. How does such a thing arise? Recall that any linear map of $\mathbb{R}^n$ is determined by its action on the standard basis vectors, since they specify the space's axes and their associated "grid". The linear map then transforms the space by transforming the grid, rotating and stretching the basis vectors and the axes they determine. If, while undergoing this transformation, one axis crosses through the others' span, this will cause an orientation reversal of the space. If two such crossings occur, the two reversals will undo one another, with a net effect that orientation is preserved. Indeed, the orientation is preserved or reversed according to whether there are an even or odd number of such "axis crossings" during the transformation.

Example 1. The figure at right shows the map whose matrix is

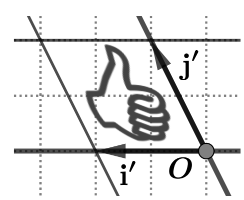
$$A = \begin{pmatrix} 5/4 & 1/2 \\ -1/4 & 3/2 \end{pmatrix}.$$



The area of the "box" in which the hand is drawn turns out to be 2 units^2. (Don't worry about how to determine that for now.) This map obviously preserves orientation (the right hand remains a right hand), so $\det A = 2$. ♦

Example 2. The figure at right shows the map whose matrix is

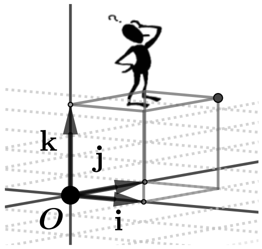
$$A = \begin{pmatrix} -2 & -1 \\ 0 & 2 \end{pmatrix}.$$

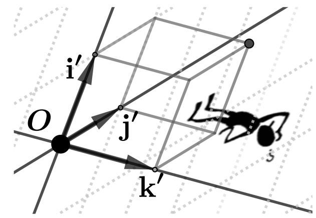


In this case, the area of the "box" is clearly 4 units^2. Moreover, this map *reverses* orientation (the right hand is now a left hand), so $\det A = -4$. ♦

Note the reason for the orientation reversal in Example 2: Dragging $\mathbf{i}$ and $\mathbf{j}$ to their new positions $\mathbf{i}'$ and $\mathbf{j}'$ requires us to pass one basis vector through the other's span – or to put it another way, it requires us to pass axis through the other.

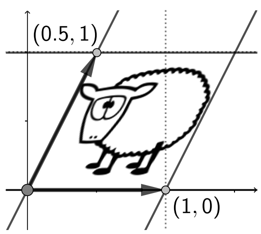
Example 3. The puzzled man standing on the box at left and scratching his head with his left hand wonders if there are linear maps that would make him scratch with his *right* hand instead. Yes, there are: Any orientation-reversing map will do the trick. The figure at right shows one. And if the volume of the box at right is, say, 30% greater than that of the box at left, then this map's determinant must be $-\mathbf{1.3}$. ♦

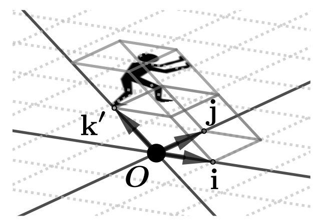




A good visualization exercise is to imagine the original axes from Example 3's left figure morphing into those at right. One way is to first imagine $\mathbf{i}$ and $\mathbf{j}$ being transformed (through rotations and stretches) into their new positions, carrying along the plane that they span. Then, once they've arrived, imagine $\mathbf{k}$ being moved down to *its* new position. At some point during this two-stage transformation, the third axis will have to pass through the plane spanned by the first two, yielding the orientation reversal.

Example 4 (Shears). A **shear** is a map that moves just *one* standard basis vector (call it the m^{th}), and in a way that its m^{th} coordinate remains 1. (It fixes all the other standard basis vectors.) The left figure shows a shear in $\mathbb{R}^2$ that moves only $\mathbf{j}$. The right figure depicts a shear in $\mathbb{R}^3$ that moves only $\mathbf{k}$, pushing the tip of $\mathbf{k}'$ over to the point $(-0.5,\ 0.2,\ 1)$. Observe that the third coordinate is still 1, as required.





Shears will be important later in this chapter, where they'll help us find a method for computing *any* matrix's determinant. For now, however, we just want to recognize what a shear's matrix looks like, and understand what the determinant of any shear matrix must be.

The first part is easy: It follows immediately from a shear's definition that a shear matrix looks like an identity matrix in which someone has tampered with the zeros in one column, changing at least one of them to a nonzero value. Thus, the two shears depicted above are represented by these matrices:

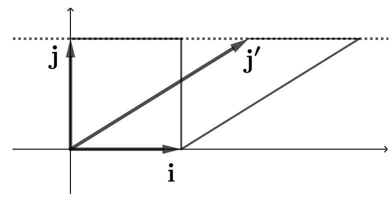
$$\begin{pmatrix} 1 & .5 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & -.5 \\ 0 & 1 & -.2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Similarly, the following matrices represent other shears:

$$S_1 = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \qquad S_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad S_3 = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

You can draw pictures of how S_1 and S_2 transform the standard grids of their respective spaces. Shear S_3 is in $\mathbb{R}^4$, so obviously we can't draw it, but you might still enjoy thinking about it. (For example, if you were a three-dimensional being living in the hyperplane $w = 0$ with no concept of the fourth spatial dimension in which your world is embedded, would you notice it if the four-dimensional world were subjected to that shear?)

Shear matrices are thus easy to recognize. But what can we say about a shear's *determinant*?

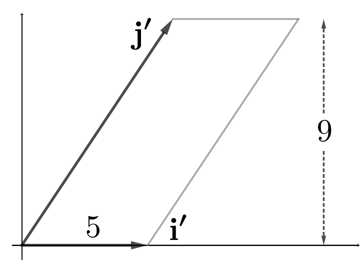


The answer: **The determinant of every shear matrix is 1**. Proof: The original box and the sheared box have the *same base* (the one determined by the fixed standard basis vectors) and, relative to that base, they also have the *same height* (1 unit). Moreover, all the boxes' corresponding cross sections (taken at the same height and parallel to their common base) are equal, since these cross sections are both obviously equal to the boxes' shared base. Thus, by Cavalieri's Principle, the boxes have the same volume. Moreover, shears obviously preserve orientation, so it follows that the determinant of every shear is 1, as claimed. ♦

Let's make a quick geometric observation. The boxes with which we're concerned in this section should, in principle, lend themselves to simple volume calculations because any box of this sort can be understood as a slanted stack of copies of the box's base, piled up into another dimension. A parallelogram is a stack of equal line segments; a parallelepiped is a stack of equal parallelograms; a four-dimensional parallelotope is a stack of equal parallelepipeds; and so forth. This being so, any such box's "volume" is simply its height times its *base's* "volume". We can use this simple observation to derive a formula for the determinant of any *upper triangular matrix*, as we'll see next.

An **upper triangular matrix** is a square matrix whose entries below the main diagonal are all zeros. (*Lower* triangular matrices are defined analogously.) Here are some examples:

$$A = \begin{pmatrix} 5 & 6 \\ 0 & 9 \end{pmatrix}, \qquad B = \begin{pmatrix} 4 & 1 & 5 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}, \qquad C = \begin{pmatrix} 3 & 1 & 4 & 1 \\ 0 & 5 & 9 & 2 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$



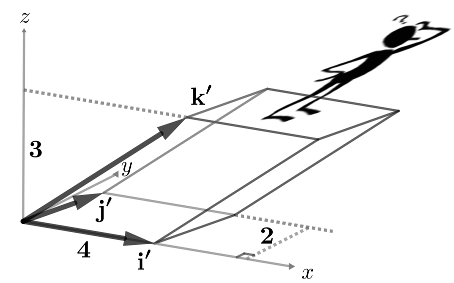
Every upper triangular matrix generates a grid of boxes whose volumes are easily computed. For example, consider matrix A. Its columns generate a parallelogram whose base, lying on the x-axis, is 5 units, and whose height is 9 units. Its *area* is thus $\mathbf{5} \cdot \mathbf{9} = 45$ units². (Note that the *first* entry in A's second column has no effect on the parallelogram's area; had it been 307 instead of 6, the parallelogram's area would still be 45 units².) Since this map clearly preserves orientation ($\mathbf{i}$ and $\mathbf{j}$ need not cross one another to reach their new positions $\mathbf{i}'$ and $\mathbf{j}'$), we may conclude that

$$\det\begin{pmatrix} \mathbf{5} & 6 \\ 0 & \mathbf{9} \end{pmatrix} = \mathbf{5} \cdot \mathbf{9} = 45.$$

Will the determinant of *every* 2×2 upper triangular matrix be the product of its two diagonal entries? Yes! If both diagonal entries are positive, the same argument we just used will still hold. (Make up some examples to convince yourself of this!) If one diagonal entry is negative, then $\mathbf{i}'$ and $\mathbf{j}'$ will be oriented in such a way that the map *reverses* orientation, and hence the determinant will be negative; but the product of the diagonal entries will also be negative, so all's well (see Exercises 3a & 3b). If both diagonal entries are negative, vectors $\mathbf{i}'$ and $\mathbf{j}'$ will be oriented in such a way that the map *preserves* orientation; of course, in that case, the diagonal entries' product will be positive, so all's well in that case, too (see Exercise 3c). Finally, if either diagonal entry is zero, the "parallelogram" collapses into something *without area* (either a line segment or a point), so the determinant will be zero, as will the diagonal entries' product.

So far so simple. But can the determinant by computed so simply for *all* upper triangular matrices, or only for 2×2 ones?

Let's turn our attention to a 3×3 example, matrix B above. Its first two columns lie conveniently in the xy-plane. (It's easy to see that this will hold for *every* 3×3 upper triangular matrix.) These two columns generate a parallelogram whose area, by the logic we employed for our 2×2 matrix A, must be $4 \cdot 2$ units^2. This parallelogram is the base of the parallelepiped generated by all three columns, the third of which gives the parallelepiped's height: 3 units. This height depends, of course, exclusively on the third column's third entry. Consequently, the parallelepiped's volume must be $(4 \cdot 2) \cdot 3$ units^3, and since orientation is clearly preserved here, the determinant is positive. Thus,



$$\det\begin{pmatrix} \mathbf{4} & 1 & 5 \\ 0 & \mathbf{2} & 1 \\ 0 & 0 & \mathbf{3} \end{pmatrix} = \mathbf{4} \cdot \mathbf{2} \cdot \mathbf{3} = 24.$$

Thus, our determinant of a 3×3 upper triangular matrix is still just the product of its diagonal entries. And indeed, the same basic argument shows that the determinant of *every* 3×3 upper triangular matrix is simply the product of its diagonal entries.

We can't depict the 4×4 case, but we can still understand it. Consider the upper triangular matrix C above. Being upper triangular, its first column lies on the x-axis, and its second lies in the xy-plane. In that plane, the first two columns generate a parallelogram of base 3 and height 5 – and thus of volume $3 \cdot 5$. That parallelogram, in turn, is the base of the parallelepiped in the xyz-hyperplane that is generated by C's first *three* columns. It has a height of 6, so its volume is $(3 \cdot 5) \cdot 6$. This parallelepiped is also the base of the *four-dimensional parallelotope* generated by all four of C's columns. The parallelotope has a height of 3, the fourth column's fourth entry. Consequently, its hypervolume is $(3 \cdot 5 \cdot 6) \cdot 3$. It follows that

$$\det\begin{pmatrix} \mathbf{3} & 1 & 4 & 1 \\ 0 & \mathbf{5} & 9 & 2 \\ 0 & 0 & \mathbf{6} & 0 \\ 0 & 0 & 0 & \mathbf{3} \end{pmatrix} = \mathbf{3} \cdot \mathbf{5} \cdot \mathbf{6} \cdot \mathbf{3} = 270,$$

so the determinant of our 4×4 matrix is indeed just the product of its diagonal entries.

Clearly, the argument above can easily be extended to 5×5 matrices, 6×6 matrices and so forth. And with minor adjustments (see exercise 4), we can use it to show that the determinant of any *lower* triangular matrix is also just the product of its diagonal entries. We summarize our hard-won results about triangular matrices (upper and lower) in a box:

Theorem. The determinant of any triangular matrix is the product of its diagonal entries.

Later in this chapter, this will help us find a method for finding the determinant of *any* square matrix M. The general method, as we'll see, will be to row-reduce M until it becomes an upper triangular matrix T, whose determinant will, thanks to our theorem, be obvious. We'll then be able to deduce $\det(M)$ from two things: $\det(T)$, and the sequence of row operations that we used to reduce M to T.

Exercises.

1. Explain geometrically (i.e. without appealing to the theorem at the end of the section) why

$$\det\begin{pmatrix}3 & 0 & 0\\ 0 & 5 & 0\\ 0 & 0 & 2\end{pmatrix} = 30.$$

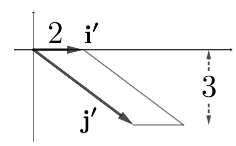
2. True or false (and explain):

a) Every diagonal matrix is an upper triangular matrix.
b) Every matrix has a determinant. c) Every square matrix has a determinant.
d) Every shear matrix is a triangular matrix. e) Every triangular matrix is a shear matrix.
f) Some matrices are both upper triangular and lower triangular.
g) $\det I = 1$.

3. a) Consider the upper triangular 2×2 matrix

$$\begin{pmatrix}2 & 4\\ 0 & -3\end{pmatrix},$$

whose effect is shown at right. By our theorem on triangular matrices, its determinant should be $2(-3) = -6$. Explain geometrically (in terms of the figure) why that negative determinant makes sense. (And do you see why this will always happen for every upper triangular 2×2 matrix with a negative bottom right entry and a positive upper left one?)



b) Make up an upper triangular 2×2 matrix with a negative top left entry, but a positive bottom right one. Draw a corresponding picture and explain geometrically why the determinant should be negative. Do you see that this will be the case for every matrix of this sort?

c) Finally, consider the case when both diagonal entries are negative. Our theorem tells us that the determinant of such a matrix should be positive. That means that orientation should be preserved, so we should be able to get $\mathbf{i}$ and $\mathbf{j}$ into their new positions $\mathbf{i}'$ and $\mathbf{j}'$ in such a way that neither must cross the line containing the other. Explain, in broad terms, how to accomplish this for a specific matrix or two of your choosing.

4. a) Give a careful definition of a *lower* triangular matrix.
b) Make up your own 2×2 lower triangular matrix with positive entries on the diagonal and explain geometrically *why* its determinant is the product of its diagonal entries.
c) Same story, but with a matrix containing a negative (or two) on the diagonal.
d) Same story, but with a 3×3 lower triangular matrix. You can stick to all positive diagonal entries.
e) Again, but with a 4×4 lower triangular matrix.

5. How can we understand a 4-dimensional hypercube? One way is to build up to it dimension by dimension, through repeated perpendicular "extrusions".

Start with a 0-dimensional object (a point) and push it one unit east, tracing out a 1-dimensional object: a "stack of points", a *line segment*. We then push this segment one unit north, tracing out a 2-dimensional object: a "stack of segments", a unit *square*. When we push this a unit upwards (perpendicular to the plane it lies in), it generates a 3-dimensional object: a "stack of squares", a unit *cube*. Now we must imagine pushing that unit cube one unit in a mysterious spatial direction that is somehow perpendicular to our entire three-dimensional space. Doing so yields a "stack of cubes" (stacked along a fourth dimension, with successive cubes touching at *all* their corresponding points). This is a unit *four-dimensional hypercube*, sometimes called a *tesseract*.

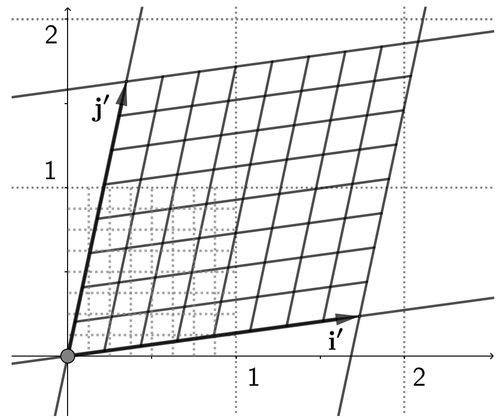
a) How many 0-dimensional vertices, 1-dimensional edges, and 2-dimensional faces does a tesseract have? Explain your answers *geometrically*.
b) Four dimensional figures have not only 0-dimensional vertices, 1-dimensional edges, and 2-dimensional faces, but also 3-dimensional "cells". The tesseract's cells are cubes. How many such cells does it have?

The Determinant: Properties

For there are many properties in it that, if universally known, would habituate its use and make it more in request with us than with the Turks themselves.

- Awister, quoted by Thomas De Quincey in *Confessions of an English Opium Eater* ("To the Reader").

If we transform $\mathbb{R}^n$ with an $n \times n$ matrix A, it's not too hard to see that the n-dimensional volume of *every* object in $\mathbb{R}^n$ (not just the "boxes") will be scaled by a factor of $|\det A|$. To see why, we'll need to don our integral calculus glasses.



Begin by noting that each "standard box" in the standard grid (generated by the standard basis vectors) has volume 1. By the determinant's definition, A transforms each standard box into a box of volume $|\det A|$. In the figure at right, for example, the box with sides $\mathbf{i}'$ and $\mathbf{j}'$ has volume $|\det A|$. (Of course, in this two-dimensional case, "volume" is *area*.)

So far so simple. Next, imagine chopping each standard box into k equal "boxlets". (In the figure, I've made $k = 64$.) Each boxlet has volume $\mathbf{1/k}$, and since the k *transformed* boxlets (whose total area is $|\det A|$) are equal, each transformed boxlet must have volume $\mathbf{|\det A|/k}$. It follows that $|\det A|$ serves as a volume-scaling factor not just for full standard boxes, but for our boxlets, too. This will hold regardless of whether we use 64 boxlets, 1000 boxlets, 10^{100} boxlets, or – letting the spirit of calculus guides us – *infinitely many* infinitesimally small boxlets. In all cases, $|\det A|$ acts as a volume-scaling factor when A acts on the boxlets. We'll can now think of any old figure in $\mathbb{R}^n$ (a hand, a sheep, a 6-dimensional hypersphere, or whatnot) in a pixelated manner – chopped into a collection of infinitesimal boxlets. Since A scales the volume of each infinitesimal boxlet by a factor of $|\det A|$, it clearly scales *the entire figure's* volume by $|\det A|$, too.

Property 1. The determinant is a volume-scaling factor: When an an $n \times n$ matrix A transforms objects in $\mathbb{R}^n$, it scales their n-dimensional volumes by $|\det A|$.

The idea of the determinant as a volume-scaling factor is tremendously important in vector calculus. In ordinary freshman calculus, we work with functions from $\mathbb{R}$ to $\mathbb{R}$. Locally (on an infinitesimal scale), such functions' graphs are straight, so locally, the functions themselves are *linear*. We can thus describe any such function's local behavior with one number: the slope of the line that it resembles at that point. And as everyone knows, we call that number the function's *derivative* at that point. In vector calculus, however, we consider functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. Locally, such functions resemble *linear transformations*. Thus, describing the local behavior of a nonlinear function from $\mathbb{R}^n$ to $\mathbb{R}^m$ requires more than a number; it requires a matrix. Accordingly, the derivative of such a function is, at any point in its domain, a *matrix*. When $m = n$, the matrix is square, and thus it has a determinant. This determinant plays a role when we change variables in integration, doing the higher-dimensional analogue of the familiar "u-substitution". You can look forward to studying how this works in a future course.

This volume-scaling property helps us understand a vital algebraic property of determinants: The determinant of a product is… the product of the determinants.

> **Property 2.** If A and B are any two $n \times n$ matrices, then
> $$\det(AB) = \det(A)\det(B).$$

For example, suppose A and B are square matrices that scale volumes by factors of 5 and 3 respectively. I claim that AB must scale volumes by a factor of 15. Why? Well, multiplying matrices corresponds to composing linear maps, so AB first scales volumes by 3 (through B's action), then scales the results by 5 (through A's action). Hence, the net effect is that AB scales volumes by a factor of $3 \cdot 5 = 15$, as claimed. This same analysis clearly holds for *any* two matrices with nonnegative determinants. A little thought shows that it also holds when a negative determinant (or two) is involved. For example, if M scales volumes by 2 *and reverses orientation* (so that $\det(M) = -2$) while N scales volumes by 5 and preserves orientation (so that $\det(N) = 5$), then MN obviously scales volumes by 10 *and reverses orientation*. That is, $\det(MN) = -10$, which is, of course, the product of $\det(M)$ and $\det(N)$. I'll leave it to you to convince yourself that this property also holds when *both* matrices have negative determinants.

Our next two properties concern some relationships between determinants and **inverse matrices.** Recall from Chapter 3, Exercise 23 that the inverse of A, which we denote A^{-1}, is the matrix that "undoes" A's action. (That is, if $\boldsymbol{A}\mathbf{v} = \mathbf{w}$, then $A^{-1}\mathbf{w} = \mathbf{v}$.) It follows that the product of A and A^{-1} (in either order) is the identity matrix I. In that same exercise, we saw that not every matrix is invertible. For a linear map (or the matrix representing it) to be invertible, it must be "one-to-one"; that is, it must always take distinct points from the domain to distinct points in the range. (After all, if A were to map both $\mathbf{v_1}$ and $\mathbf{v_2}$ to $\mathbf{w}$, then what would $A^{-1}\mathbf{w}$ be?)

See if you can justify this next property on your own before reading that explanation that follows it.

> **Property 3.** If A is any invertible matrix, then
> $$\det(A^{-1}) = \frac{1}{\det A}.$$

If A scales volumes by a factor of 7, then A^{-1}, undoing that action, must scale volumes by a factor of $1/7$. Similarly, if B scales volumes by $2/3$ and reverses orientation, then B^{-1} will need to scale volumes by $3/2$ and reverse orientation. If you understand that much, you should see why Property 3 always holds.

But wait a minute. What if $\det A = 0$? In that case, our formula involves division by 0. But fear not: This can't happen. If $\det A = 0$ (and A is $n \times n$), A scales every n-dimensional volume by a factor of zero. That doesn't mean that A crushes all of $\mathbb{R}^n$ into $\mathbf{0}$, but it does mean that some dimensional collapse must occur. (For instance, a 3×3 matrix that crushes $\mathbb{R}^3$ down into a plane has determinant 0; yes, *something* comes out the other end of the mapping, but no 3-dimensional volume survives.) Whenever dimensional collapse occurs, many distinct points in the domain end up being mapped to the same point in the range.* Thus the map isn't one-to-one, so its matrix can't be invertible, so Property 3 wouldn't be applicable. We've discovered something important in passing: *If a matrix has determinant 0, it isn't invertible*.

* "Dimensional collapse" means that the map's rank less than its domain's dimension. So by the rank-nullity theorem, the kernel's dimension is *at least 1*. Hence, the map's kernel has infinitely many points, all of which get mapped to $\mathbf{0}$.

On the other hand, it's easy to see that any $n \times n$ matrix with a *non*zero determinant *is* invertible. If the determinant is nonzero, then although n-dimensional volume may be scaled, it at least *survives* as n-dimensinal volume, which implies that there's no dimensional collapse. All n dimensions remain in the map's output, which will be pervaded by a "clean grid" generated by n linearly independent vectors. It should be intuitively clear in your mind's eye that any such transformation maps distinct points to distinct points. In other words, any such transformation is one-to-one, and therefore the matrix that represents it is invertible.

Combining the results of the previous two paragraphs yields our next trophy: A matrix is invertible if and only if its determinant is nonzero.

Property 4. Matrix A is invertible $\Leftrightarrow$ $\det A \neq 0$

So far, we have built up the determinant's most important properties just by thinking geometrically. But how do we actually compute a determinant? We'll begin with one special case: The determinant of a 2×2 matrix.

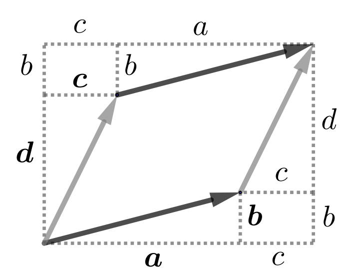
Property 5. (2×2 Determinant formula)

$$\det\begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - cb$$

Proof. The figure at right shows the parallelogram determined by the 2×2 matrix's columns. The determinant is the parallelogram's area, which is the large rectangle's area minus the areas of the four right triangles and the two small rectangles in the corners. Or in symbols,

$$\det\begin{pmatrix} a & c \\ b & d \end{pmatrix} = (a+c)(d+b) - 2\left(\frac{1}{2}ab\right) - 2\left(\frac{1}{2}cd\right) - 2(bc).$$

The right-hand side reduces, as you should verify, to $ad - cb$. ■



For the preceding proof to be fully rigorous, we'd need to consider some other possible configurations. (For example, what if the vectors make an obtuse angle? What if the vectors are arranged so that the determinant is negative?) You'll dispose of some such cases in Exercise 9.

Exercises.

6. We justified Property 3 geometrically. Now explain it with algebra. [**Hint**: By definition, $A^{-1}A = I$.]

7. Consider the following matrices:

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} 8 & 3 \\ 7 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 6 \\ 2 & 4 \end{pmatrix}, \quad D = \begin{pmatrix} -4 & 3 \\ -5 & 4 \end{pmatrix}, \quad E = \begin{pmatrix} 9 & 7 \\ 3 & 9 \end{pmatrix}.$$

a) Compute their determinants. b) Which of the preceding matrices are invertible?
c) Remind yourself *why* a matrix whose determinant is zero cannot be inverted.
d) Which of the preceding matrices represent linear transformations that preserve area?
e) Which of them reverse the orientation of $\mathbb{R}^2$?
f) Find the determinants of the *inverses* of the invertible matrices. Be sure that you understand why your answers make geometric sense. [*Note: You need not find the inverse matrices themselves, just their determinants*.]
g) Find B^2 and use it to compute $\det(B^2)$. Be sure that you understand *why* the relationship between $\det B$ and $\det(B^2)$ makes geometric sense.
h) Without finding AB, find $\det(AB)$. Then find AB and compute its determinant directly to verify your answer.

8. (Extending the **Invertible Matrix Theorem**) In Chapter 4, Exercise 40, you found that for any $n \times n$ matrix A, statements A - I below are logically equivalent, meaning that they all stand or all fall together. Convince yourself that statement J can join this growing list, too. (The list will grow again in Exercise 19.)

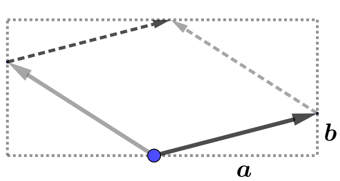
a) A is invertible. b) $\text{rref}(A) = I$. c) $A\mathbf{x} = \mathbf{b}$ has a *unique* solution for every vector $\mathbf{b}$.
d) A's columns are linearly independent. e) A's columns span $\mathbb{R}^n$. f) A's columns constitute a basis for $\mathbb{R}^n$.
g) $\ker(A) = \mathbf{0}$. h) $\text{im}(A) = \mathbb{R}^n$. i) $\text{rank}(A) = n$. j) $\det A \neq 0$.

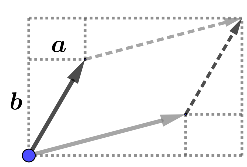
9. We justified our 2×2 determinant formula

$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc$$

with a nice orientation-preserving matrix whose columns determined a parallelogram lying in the first quadrant. Given a less nice setup – such as one of those indicated in the figures below (the first of which has a parallelogram stretching into the second quadrant, the second of which corresponds to an orientation-*reversing* map) – does the 2×2 determinant formula still hold? Yes it does. Demonstrate this. Once you've grasped those two cases, you'll likely be convinced that the formula does indeed hold for all possible 2×2 matrices.

[**Hint for Part A:** *Entries in a matrix can be negative, but lengths must, of course, be positive*.]

a)


b)


10. In Exercise 29 of Chapter 4, you established a quick formula for the inverse of a 2×2 matrix

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

You can now rewrite that quick formula for A^{-1} in a manner that is slightly easier to remember. Do so.

11. Property 2 extends to products of three or more matrices. For example, it's also true that

$$\det(ABC) = \det(A)\det(B)\det(C).$$

However many matrices there are, *the determinant of a product is the product of the determinants*. Explain why. Then use this idea to compute $\det(D^{1000})$, where D is the matrix of that name in Exercise 7.

12. Determinants often give us surprising leverage in proofs - even when we are proving things unrelated to volume. In this exercise, for instance, determinants will help us prove that a square matrix A's "**left inverse"** (i.e. a matrix B such that $BA = I$) necessarily works as a **"right inverse"** as well (that is, $AB = I$).

The preceding fact about "one-sided inverses" might seem obvious, but recall that when I introduced inverse matrices (Chapter 3, Exercise 23), I did so as follows:

*The matrix that undoes the action of a square matrix A is called the **inverse matrix** of A, and is denoted A^{-1}. Thus, by definition, $A^{-1}A = I = AA^{-1}$.*

In a footnote for that sentence, I added the following:

If we wish to show that a matrix B is in fact A^{-1}, we must – by this definition – verify two separate things: $BA = I$ and $AB = I$. But we'll prove later (Ch. 5, Exercise 12) that each of these things implies the other, so to verify both, we just need to verify one.

Well, here we are. Let's get to work.

a) Convince yourself that each step in the following argument holds, justifying each step where appropriate.

Claim. If A is an $n \times n$ matrix and $BA = I$, then $AB = I$, too.

Proof. If $BA = I$, then $\det(BA) = \det I$. Thus, $\det(B)\det(A) = 1$. From this it follows that $\det B \neq 0$. Hence, B doesn't collapse any dimensions. In particular, $\ker(B) = \mathbf{0}$, a fact that we'll use shortly.

Since $BA = I$, it follows that $BAB = B$, or equivalently, $BAB - B = Z$, where Z is the zero matrix, representing the "zero map". Rewriting the last equation as $B(AB - I) = Z$, we deduce that $B(AB - I)$ sends *all* vectors in $\mathbb{R}^n$ to $\mathbf{0}$. But since B sends only $\mathbf{0}$ to $\mathbf{0}$, it must be the case that all outputs of $(AB - I)$ are already $\mathbf{0}$. In other words, matrix $(AB - I)$ must represent the zero map. Accordingly, $AB - I = Z$. Adding I to both sides of this equation, we find that $AB = I$, as claimed. ■

b) Prove the converse claim: If A is an $n \times n$ matrix and $AB = I$, then $BA = I$, too. (And hence, $B = A^{-1}$.)

The moral of this exercise is that if we ever show that some square matrix A has a "one-sided inverse", then it must in fact be the full inverse A^{-1}. We are therefore mercifully free from the burden of having to distinguish between "left inverses" and "right inverses" of square matrices.

Elementary Matrices (and their Determinants)

Elementary, my dear Watson!

- not Sherlock Holmes[*]

An **elementary matrix** is one that can be obtained from the *identity* matrix by applying one row operation. Accordingly, the following are all examples of elementary matrices:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}.$$

The first matrix was obtained by swapping I's first two rows. The second was obtained by scaling I's middle row by 3. The third was obtained by adding two copies of I's middle row to its bottom row.

Row operations are violent but effective surgery in which we slice open a matrix to alter its innards. Elementary matrices can deliver all the results of "row operation surgery" but in a less invasive manner. The idea: Instead of doing a row operation on a given matrix, we left-multiply it by the elementary matrix obtained from I by that row operation. For example, instead of scaling a matrix's 2nd row by 3 like this,

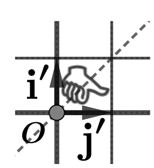
$$\begin{pmatrix} 2 & 5 & 7 \\ 1/3 & 2/3 & 3 \\ 4 & 1 & 4 \end{pmatrix} \quad \times 3 \quad \begin{pmatrix} 2 & 5 & 7 \\ 1 & 2 & 9 \\ 4 & 1 & 4 \end{pmatrix},$$

we can left-multiply the given matrix by the elementary matrix obtained by scaling I's 2nd row by 3:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 5 & 7 \\ 1/3 & 2/3 & 3 \\ 4 & 1 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 5 & 7 \\ 1 & 2 & 9 \\ 4 & 1 & 4 \end{pmatrix},$$

Obviously, doing the row operation directly is simpler in a narrow pragmatic sense, but the elementary matrix approach has a crucial advantage for theoretical work: It replaces a messy ad hoc procedure with pristine matrix *algebra*. It replaces bookkeeping scratchwork with an *equation* – an object to which we can then apply algebraic rules. And this algebra, in turn, will lead us to a simple algorithm for computing the determinant of *any* square matrix, not just those special few whose determinants we've found so far (shear, triangular, and 2×2 matrices). But before we can do that, we'll need to learn one last thing about elementary matrices: their determinants. This turns out to be easy, since, as we'll show in the next three paragraphs, every elementary matrix represents one of three simple geometric operations.

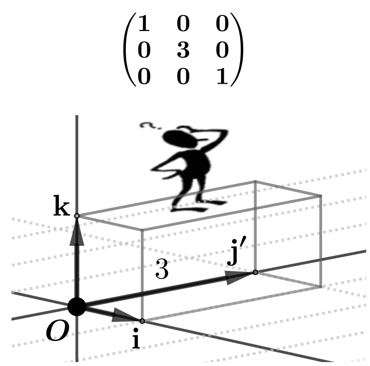
$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$



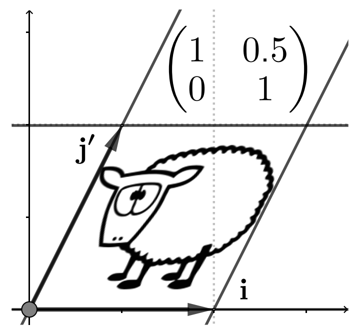
First, any elementary matrix obtained by swapping rows of I represents a *reflection*. Why? Swapping I's j^{th} and k^{th} *rows* is equivalent to swapping its j^{th} and k^{th} *columns* – an operation with clear geometric meaning: It swaps the j^{th} and k^{th} standard basis vectors' positions, *reflecting* them (in their common plane) across the line that bisects the right angle between them. Hence, all elementary matrices of this first type are reflections, as claimed.[†] Hence, the determinants of all "row swap elementary matrices" are $-\mathbf{1}$, since reflections preserve volumes but reverse orientation.

[*] This famous expression never occurs in any of Arthur Conan Doyle's many Sherlock Holmes stories and novels. One steeped in Sherlockiana might call it "the curious incident of the Holmes exclamation in the night".

[†] Strictly speaking, the linear map corresponding to an $n \times n$ matrix of this sort is a reflection in an $(n-1)$-dimensional "mirror". If we swap columns j and k, the "mirror" is the subspace spanned by the vector $\mathbf{e}_j + \mathbf{e}_k$ *and* the $(n-2)$ fixed basis vectors.



Second, any elementary matrix that we obtain by scaling one of I's rows is a *stretch along an axis* (with a reflection if the scalar is negative). Why? Scaling I's j^{th} row by c is equivalent to scaling its j^{th} *column* by c – an operation with a clear geometric interpretation: It alters the standard grid by stretching the j^{th} basis vector by a factor of $|c|$ (and reversing its direction if $c < 0$), while the others stay put. The transformed grid thus consists of rectangular boxes, each of volume $(|c| \cdot 1 \cdot 1 \cdots 1) = |c|$. Orientation will be preserved or reversed according to c's algebraic sign. Thus, all elementary matrices of this second type represent stretches (sometimes with an accompanying reflection). Hence, the determinant of any "row scale elementary matrix" is the row scale factor $\boldsymbol{c}$ itself.



Third and last, any elementary matrix that we obtain by adding a multiple of one of I's rows to another of its rows represents a *shear*. To see why, suppose that we add c copies of I's j^{th} row to its k^{th} row. The resulting elementary matrix will look like an identity matrix in which someone has tampered with one column, changing one of its zeros to some other number. As we saw in Example 4 in this chapter's first section, a matrix of that form is a *shear matrix*. And as we discussed in that same example, the determinant of any shear matrix is 1. Thus, any elementary matrix of this third kind has a determinant of $\mathbf{1}$.

To sum up: Any elementary matrix that we can use to…

- *Swap rows* represents a **reflection**. Hence, its determinant is $-\mathbf{1}$.
- *Scale a row* (by c) represents a **stretch** (by a factor of c; if $c < 0$, there's a **reflection**, too). Hence, its determinant is $\boldsymbol{c}$.
- *Add a multiple of one row to another* is a **shear**. Hence, its determinant is $\mathbf{1}$.

After a few exercises, we'll finally be ready to turn to our last major problem of the chapter: deriving a method for computing the determinant of *any* square matrix.

Exercises.

13. For each of the following row operations, find the elementary matrix that carries it out. Verify your matrices by left-multiplying them against the telephone matrix at right.

$$\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}$$

a) Swap rows 2 and 3.
b) Scale row 2 by 4.
c) Subtract 4 copies of row 1 from row 2.

14. What are the determinants of the three elementary matrices you found in the previous problem?

15. Suppose E_1, E_2, E_3 are elementary matrices; E_1 swaps rows, E_2 scales a row by -3, and E_3 adds a multiple of one row to another. Now suppose that $E_3E_2E_1A$ is a triangular matrix T, whose determinant is 5. What, if anything, can we conclude about $\det(A)$? [**Hint**: *Recall Exercise 11*.]

Computing Determinants (by Row Reduction)

At long last, we can establish a simple algorithm for computing any square matrix's determinant.

Determinant Algorithm.

Use Gaussian elimination to reduce the given matrix A to a *triangular* matrix T, and keep track of…

- the number s of row swaps you use
- the product p of all the factors by which you multiply rows.

We then have

$$\det A = \frac{\det T}{(-1)^s p}$$

(As discussed earlier in this chapter, $\det T$ is the product of T's diagonal entries.)

If you solved Exercise 15 on the previous page, you've already understood, in essence, why this algorithm works. But let's spell out the details.

Proof. To reduce A to a triangular matrix T, we perform row operations. Let $E_1, E_2, \dots, E_m$ be their corresponding elementary matrices so that $E_m \cdots E_2 E_1 A = T$. Taking determinants of both sides (and recalling Exercise 11) yields $\det(E_m) \cdots \det(E_2)\det(E_1)\det(A) = \det(T)$. Or equivalently,

$$\det A = \frac{\det T}{\det(E_m) \cdots \det(E_2)\det(E_1)}.$$

By our work in the previous section, we can evaluate that denominator. There are s "row swap elementary matrices" down there, so we know that the product of their determinants is $(-1)^s$. Next, the "row scale elementary matrices": We know that the product of their determinants is p. This leaves the elementary matrices that add multiples of one row to another. These are shear matrices, so their determinants are all 1. It follows that $\det A = \det T / (-1)^s p$, as claimed. ■

An example will make the idea clearer.

Example 1. Find the determinant of $\begin{pmatrix} 2 & -3 & 7 \\ 4 & 9 & -3 \\ 2 & 4 & -2 \end{pmatrix}$.

Solution. We'll reduce the matrix to triangular form, keeping track of row swaps and the product of any factors we use to scale rows. Here's one way to accomplish this row reduction:

$$\begin{pmatrix} 2 & -3 & 7 \\ 4 & 9 & -3 \\ 2 & 4 & -2 \end{pmatrix} \overset{R_1 \leftrightarrow R_3}{} \begin{pmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ 2 & -3 & 7 \end{pmatrix} \overset{\times 1/2}{} \begin{pmatrix} 1 & 2 & -1 \\ 4 & 9 & -3 \\ 2 & -3 & 7 \end{pmatrix} \begin{matrix} \\ -4R_1 \\ -2R_1 \end{matrix} \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 0 & -7 & 9 \end{pmatrix} \begin{matrix} \\ \\ +7R_2 \end{matrix} \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 7 \\ 0 & 0 & 16 \end{pmatrix}.$$

This row reduction entailed 1 row swap. The product of all scaling factors (only one) was $1/2$. Since the process yielded a triangular matrix whose determinant is 16, we conclude that

$$\det A = \frac{16}{(-1)^1(1/2)} = \mathbf{-32}.$$ ♦

And that's that. With that technique in hand, you can now compute the determinant of any square matrix whatsoever. All you need is Gaussian elimination and some careful bookkeeping. Incidentally, we can use this algorithm to recover the quick formula for 2×2 determinants that we derived geometrically.

Example 2. Use row-reduction to rederive the "quick formula" for a 2×2 determinant.

Solution. Start with an expression for a general 2×2 matrix, and row reduce it to triangular form:

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \overset{\times (1/a)}{} \begin{pmatrix} 1 & c/a \\ b & d \end{pmatrix} \underset{-bR_1}{} \begin{pmatrix} 1 & c/a \\ 0 & (ad-bc)/a \end{pmatrix}.$$

This reduction required no row swaps, and the product of all scaling factors was $1/a$. It yielded a triangular matrix whose determinant is $(ad - bc)/a$, so we conclude that

$$\det\begin{pmatrix} a & c \\ b & d \end{pmatrix} = \frac{(ad-bc)/a}{1/a} = ad - bc. \qquad \blacklozenge$$

We'll end with a surprising theorem: Any matrix and its transpose have the same determinant.

Theorem. $\det(A^T) = \det(A)$ for every square matrix A.

Proof. Reduce A to triangular form T. Let $E_1, E_2, \dots, E_m$ be the elementary matrices corresponding to the row operations, so that $E_m \cdots E_2 E_1 A = T$. Taking determinants of both sides (and recalling Exercise 11) yields $\det(E_m) \cdots \det(E_2)\det(E_1)\det(A) = \det(T)$. Equivalently,

$$\det A = \frac{\det T}{\det(E_m) \cdots \det(E_2)\det(E_1)}.$$

We'll now demonstrate that $\det A^T$ has this same form. Go back to $E_m \cdots E_2 E_1 A = T$ and take the *transpose* of both sides. By a property you proved in Chapter 3 (Exercise 26e), this yields

$$A^T E_1^T E_2^T \cdots E_m^T = T^T.$$

Now take determinants of both sides, using the property that the determinant of a product is the product of the determinants, then solve for $\det(A^T)$. Doing so, we find that

$$\det(A^T) = \frac{\det(T^T)}{\det(E_1^T)\det(E_2^T) \cdots \det(E_m^T)}.$$

In fact, we'll soon be able to erase all those transpose superscripts. To see why, first observe that since T is an upper triangular matrix, T^T is *lower* triangular. The determinants of both T and T^T are thus the products of their diagonal entries. Moreover, as transpose "mates", T and T^T have the *same* diagonal entries. Hence, $\det(T^T) = \det T$. Next, consider those elementary matrices. Transposition leaves the first two types (row scale and row swap) unchanged, so it doesn't change their determinants. As for the third type of elementary matrix, these are triangular with 1s on their main diagonal, and transposing any such matrix turns it into a matrix of the same sort. Since all such matrices have determinants of 1, transposition clearly preserves their determinants, too. We've now shown that $\det(E_i^T) = \det(E_i)$ for all i. The two highlighted equations indicate that we can indeed erase all the transpose superscripts from the right-hand side of our expression for $\det(A^T)$. Doing so and then rearranging the denominator's factors, we obtain the right-hand side of our earlier expression for $\det A$. Thus, $\det(A^T) = \det A$, as claimed. ∎

Exercises.

16. Use row-reduction to compute the determinants of the following matrices.

a) $\begin{pmatrix} 3 & 2 & 1 \\ 0 & 3 & -6 \\ 0 & 2 & -2 \end{pmatrix}$ b) $\begin{pmatrix} 5 & 3 & 7 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}$ c) $\begin{pmatrix} 2 & 4 & 6 \\ 3 & -4 & 8 \\ 0 & 2 & 5 \end{pmatrix}$ d) $\begin{pmatrix} 1 & 7 & 1 \\ -1 & 0 & 2 \\ 3 & -1 & -3 \end{pmatrix}$ e) $\begin{pmatrix} 3 & 2 & 6 \\ 4 & 8 & 4 \\ 1 & 0 & 2 \end{pmatrix}$

f) $\begin{pmatrix} 0 & 3 & 2 & 4 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 5 & 2 \\ 1 & 1 & 3 & 3 \end{pmatrix}$ g) $\begin{pmatrix} 1 & -1 & 2 & -2 \\ -1 & 2 & 1 & 6 \\ 2 & 1 & 14 & 10 \\ -2 & 6 & 10 & 33 \end{pmatrix}$ h) $\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 5 \\ 1 & 1 & 3 & 6 \\ 1 & 1 & 4 & 7 \end{pmatrix}$

17. Explain geometrically why the matrix in Exercise 16h has a determinant of 0.

18. (Quick Formula for 3×3 determinants) In Example 2, we re-derived the quick formula for a 2×2 determinant. As it happens, we can use the same idea that we used there to derive a quick formula for 3×3 determinants. The algebra involved is basic but tedious (try it), so I'll dispense with the details and just tell you the punchline:

$$\det\begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix} = aei + dhc + gbf - ahf - dbi - gec.$$

To find a mnemonic device for this result, imagine lines through the matrix, acting like objects in old video games: When a line exits the matrix through one side, it re-enters on the opposite side, proceeding in the same direction. Thus, a line beginning at d and travelling southeast passes through h and then… c. Bearing this in mind, look again at the determinant's six terms: The first three (aei, dhc, gbf) are "spelled out" by three lines starting at each top-row entry, sloping southeast. The next three terms (which are *subtracted*) are spelled out by three lines starting at each top-row entry, but now sloping south*west*. Having linked those six terms to those six lines, we can quickly compute a 3×3 matrix's determinant by mentally following the lines, writing down the six products (with plusses and minuses in the right places), and then adding them up. For example,

$$\det\begin{pmatrix} 1 & 3 & -1 \\ 2 & 4 & 0 \\ 5 & 1 & 6 \end{pmatrix} = 24 + 0 + (-2) - 0 - 36 - (-20) = \mathbf{6}.$$

Note well: This quick formula works only for the special case of 3×3 determinants! Don't try to use an analogue of it for a matrix of any other size. It won't work.

a) Verify that the determinant we just computed really is 6 by *re*computing it via row reduction.

b) Use this quick formula to re-compute the determinants of the 3×3 matrices in Exercise 16.

19. (Extending the **Invertible Matrix Theorem**) In Exercise 8, you saw that statements A - J in the list below were equivalent statements about an $n \times n$ matrix A. Explain why we can add K, L, and M to the list:

a) A is invertible. b) $\text{rref}(A) = I$. c) $A\mathbf{x} = \mathbf{b}$ has a *unique* solution for every vector $\mathbf{b}$.
d) A's columns are linearly independent. e) A's columns span $\mathbb{R}^n$. f) A's columns constitute a basis for $\mathbb{R}^n$.
g) $\ker(A) = \mathbf{0}$. h) $\text{im}(A) = \mathbb{R}^n$. i) $\text{rank}(A) = n$. j) $\det A \neq 0$.
k) A's *rows* are linearly independent. l) A's rows span $\mathbb{R}^n$. m) A's rows constitute a basis for $\mathbb{R}^n$.

To reiterate, the moral of the invertible matrix theorem is that square matrices come in two types: *invertible* matrices (which satisfy all twelve of those conditions) and *noninvertible* matrices (which satisfy none of them).

20. Do the *rows* of $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 1 & 1 & 2 \end{pmatrix}$ span $\mathbb{R}^3$? [**Hint**: *Use the previous exercise*.]

21. Suppose a square matrix has a column consisting entirely of zeros. What can we say about its determinant? Why? (And what, if anything, could we say about a matrix containing a *row* of all zeros?)

Computing Determinants (by Cofactor Expansion)

In this last section, we'll discuss *cofactor expansion*, an alternate algorithm for computing determinants. Though significantly less efficient than the row-reduction algorithm, cofactor expansion is, bizarrely, given pride of place in most linear algebra textbooks. Despite its relative inefficiency (which I'll discuss at the end of this section), cofactor expansion is at least algebraically interesting, and it's useful in certain special conditions – particularly when a matrix has a row or column consisting mainly of zeros.

We'll sneak up on it by reexamining the quick formula for 3×3 determinants we met in Exercise 18:

$$\det\begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix} = aei + dhc + gbf - ahf - dbi - gec.$$

First, let's rearrange the expression on the right as follows: $\boldsymbol{a}(ei - hf) - \boldsymbol{d}(bi - hc) + \boldsymbol{g}(bf - ec)$. Consider those three terms. Each has two factors. What can we say about the two factors of each term? After some hard staring, we can discern a pattern: Each term's first factor (in boldface) is an element from the matrix's top row. Much more subtly, each term's second factor (in parentheses) is the determinant of the 2×2 matrix that we "expose" by mentally deleting the row and the column containing the *first* factor. (Verify this.) Interesting! But why is the middle term alone *subtracted*? Is this part of a pattern as well? We can discover the answer by experimenting a bit more.

Our first experiment is inspired by a nagging aesthetic blemish: Why should a matrix's *first* row have special status with respect to the determinant? Could we rearrange our original six-term expression for the determinant to emphasize the *second* row? Well, let's play around (a cherished mathematical activity) and see what we discover. Given enough play time, you'd find that we can rearrange the original six-term determinant expression from our "quick formula" as follows: $-\boldsymbol{b}(di - gf) + \boldsymbol{e}(ai - gc) - \boldsymbol{h}(af - dc)$. The same basic pattern holds: Each term's first factor is drawn from a particular row, and its second factor is the determinant of the 2×2 matrix we get by crossing out the row and column of the first factor. Nice! But curiously, the sign pattern has been reversed. Now it's the two *outer* terms that are subtracted. Why? Maybe we'll gain further insight if we play this same game with the *third* row? Only one way to find out.

Rearranging the original six-term expression again, we get $\boldsymbol{c}(dh - ge) - \boldsymbol{f}(ah - gb) + \boldsymbol{i}(ae - db)$. Reassuringly, our basic pattern continues to hold: Each term's second factor is still the determinant of the matrix that we get by nixing the first factor's row and column. But the signs have reverted to the original $+ \; - \; +$ pattern that we saw when we "expanded" our determinant along the first row. Why?

After pondering this for a while, you may begin to feel another source of aesthetic unease. We've now seen, with satisfying symmetry, that one row is as good as another as far as our basic pattern is concerned, and yet… why should *rows* matter more than *columns* with respect to determinants? After all, we proved in the previous section that transposing a matrix – turning its rows into columns and vice versa – doesn't change its determinant. Well, can we do unto the *columns* what we've done to the rows? Let's see.

We can in fact rewrite our original determinant in the form $\boldsymbol{a}(ei - hf) - \boldsymbol{b}(di - gf) + \boldsymbol{c}(dh - ge)$. Now each term's first factor is from the first *column*, and the second is the expected 2×2 determinant. In fact, we can play this same game with the second column, $-\boldsymbol{d}(bi - hc) + \boldsymbol{e}(ai - gc) - \boldsymbol{f}(ah - gb)$, or the third, $\boldsymbol{g}(bf - ec) - \boldsymbol{h}(af - dc) + \boldsymbol{i}(ae - db)$. Each time, everything works out as we'd expect. Not only have we vindicated our aesthetic sense of symmetry, but we've now also gathered enough data (our six different "expansions" of the determinant along each row and column) to crack the code of the alternating signs… after some more hard staring and thinking.

We've seen that the following are all equivalent to our six-term expression for a 3×3 determinant:

$$\begin{gathered}
\boldsymbol{a}(ei - hf) - \boldsymbol{d}(bi - hc) + \boldsymbol{g}(bf - ec) \\
-\boldsymbol{b}(di - gf) + \boldsymbol{e}(ai - gc) - \boldsymbol{h}(af - dc) \\
\boldsymbol{c}(dh - ge) - \boldsymbol{f}(ah - gb) + \boldsymbol{i}(ae - db) \\
\boldsymbol{a}(ei - hf) - \boldsymbol{b}(di - gf) + \boldsymbol{c}(dh - ge) \\
-\boldsymbol{d}(bi - hc) + \boldsymbol{e}(ai - gc) - \boldsymbol{f}(ah - gb) \\
\boldsymbol{g}(bf - ec) - \boldsymbol{h}(af - dc) + \boldsymbol{i}(ae - db)
\end{gathered}$$

$$\begin{pmatrix} \boldsymbol{a} & \boldsymbol{d} & \boldsymbol{g} \\ \boldsymbol{b} & \boldsymbol{e} & \boldsymbol{h} \\ \boldsymbol{c} & \boldsymbol{f} & \boldsymbol{i} \end{pmatrix} \qquad \begin{pmatrix} + & - & + \\ - & + & - \\ + & - & + \end{pmatrix}$$

In the 18 terms above, we see that the sign is intimately connected with the first factor. Observe that both terms whose first factor is $\boldsymbol{a}$ are *added*. Both terms whose first factor is $\boldsymbol{b}$ are *subtracted*. Both terms whose first factor is $\boldsymbol{c}$ are *added*. And so on and so forth, all the way down to $\boldsymbol{i}$. In fact, what we see is that a term's sign is determined by its first factor's *position* in the matrix. At right, I've rewritten the original matrix, supplemented by another containing $+$ and – signs in the corresponding slots. The result, a simple checkerboard pattern, captures the sign that goes with each position in the matrix. Or, more formally, we can say that if the term's first factor is in row j, column k, then the term is added if $(j + k)$ is even, and subtracted if $(i + j)$ is odd.

We've now unearthed the full pattern, which generalizes to determinants of every size. With patience, we could follow the strategy that led us to our "quick formulas" for 2×2 and 3×3 determinants (Example 2 in the previous section & Exercise 18) and find formulas for 4×4, 5×5, or higher order determinants. Doing so, we'd see that any $n \times n$ determinant can be expressed as a sum of $n!$ terms. With some algebraic shenanigans, we can rearrange those $n!$ terms into a sum of just n terms, each of which has two factors (and a choice of sign). We can arrange matters so that the first factors come from any chosen row or column of the given matrix. The corresponding second factors will then be the determinants of the $(n-1) \times (n-1)$ matrices that we get by deleting the first factor's row and column. Finally, the sign is given by the first factor's position in the $n \times n$ "checkerboard matrix" with a $+$ in its top left corner.

This leads us to a recursive procedure for computing determinants. A 5×5 determinant, for example, reduces to a computation involving 4×4 determinants, each of which reduces to computations involving 3×3 determinants, which then reduce to computations involving 2×2 determinants, which are easy.

This process of reducing a determinant to determinants of lower degree is called **cofactor expansion**.* Since it involves picking a row or column along which to "expand", we refer more specifically to, say, cofactor expansion *along the 1st row*, or cofactor expansion *along the 5th column*, or what have you.

In practice, cofactor expansion is most convenient when a row or column consists mainly of zeros, because in that case, each zero entry yields a zero term in the cofactor expansion. A few examples will make the idea clear.

---

* The name is explained by some terminology traditionally associated with the technique. First, the matrix we get by crossing out row j and column k is called the **minor** associated with the j, k^{th} entry; next, the minor's determinant multiplied by $(-1)^{j+k}$ (i.e. the $+$ or $-$ dictated by the "checkerboard matrix") is called the the j, k^{th} entry's **cofactor**. With that terminology in place, we can summarize the technique of cofactor expansion as follows:

A matrix's determinant is a weighted sum of any row's (or column's) entries, where the weights are the entries' cofactors.

Example 1. Use cofactor expansion to compute the determinant of $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \\ 1 & 2 & -2 \end{pmatrix}$.

Solution. Let's expand along the second row to take advantage of that zero:

$$\det\begin{pmatrix} 1 & 2 & 3 \\ \mathbf{4} & \mathbf{5} & \mathbf{0} \\ 1 & 2 & -2 \end{pmatrix} = -\mathbf{4}\det\begin{pmatrix} 2 & 3 \\ 2 & -2 \end{pmatrix} + \mathbf{5}\det\begin{pmatrix} 1 & 3 \\ 1 & -2 \end{pmatrix} - \mathbf{0}$$
$$= -4(-10) + 5(-5) = \mathbf{15}.$$

Observe that the first and third terms in the initial expansion are *subtracted* owing to the positions of entries 4 and 0 in the checkerboard matrix of plusses and minuses. Also, I didn't bother writing down the determinant associated with the 0 in the third term, because there would be no point: whatever it is, it will be multiplied by zero. ♦

Now let's try this with a larger matrix.

Example 2. Use cofactor expansion to compute the determinant of the matrix

$$B = \begin{pmatrix} -8 & 2 & 3 & 7 \\ 1 & 2 & 0 & 3 \\ 4 & 5 & 0 & 0 \\ 1 & 2 & 0 & -2 \end{pmatrix}.$$

Solution. Using cofactor expansion on the third column to profit from all those zeros, we find that

$$\det\begin{pmatrix} -8 & 2 & \mathbf{3} & 7 \\ 1 & 2 & \mathbf{0} & 3 \\ 4 & 5 & \mathbf{0} & 0 \\ 1 & 2 & \mathbf{0} & -2 \end{pmatrix} = \mathbf{3}\det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 0 \\ 1 & 2 & -2 \end{pmatrix} - \mathbf{0} + \mathbf{0} - \mathbf{0} = \mathbf{3}(15) = \mathbf{45}.$$

Why was the first term *added*? Well, the "checkerboard matrix" of plusses and minuses always has a $+$ in the upper left corner, so moving two spots to the right brings us to another $+$. Incidentally, the one 3×3 matrix in our cofactor expansion was matrix A from Example 1. ♦

Computer programmers turn up their noses at cofactor expansion – with reason. Roughly speaking, computing an $n \times n$ determinant by cofactor expansion along on a random row or column requires about $n!$ arithmetic operations, whereas computing it by means of row reduction requires about n^3 operations. Hence, cofactor expansion is a much more "expensive" way to compute a determinant (when $n > 5$). A computer program using row-reduction to compute a 20×20 determinant matrix will grind through the roughly $20^3 = 8000$ arithmetic operations in no time, but another that uses cofactor expansion will never reach the end of its task, which requires $20! \approx 2$ million trillion operations.

Exercises.

22. Verify the result of Example 1 by recomputing that determinant in the following ways:
a) With the 3×3 "quick formula" b) cofactor expansion on column 3 c) cofactor expansion on row 1

23. Use cofactor expansion to find the determinants of the matrices in Exercise 16.

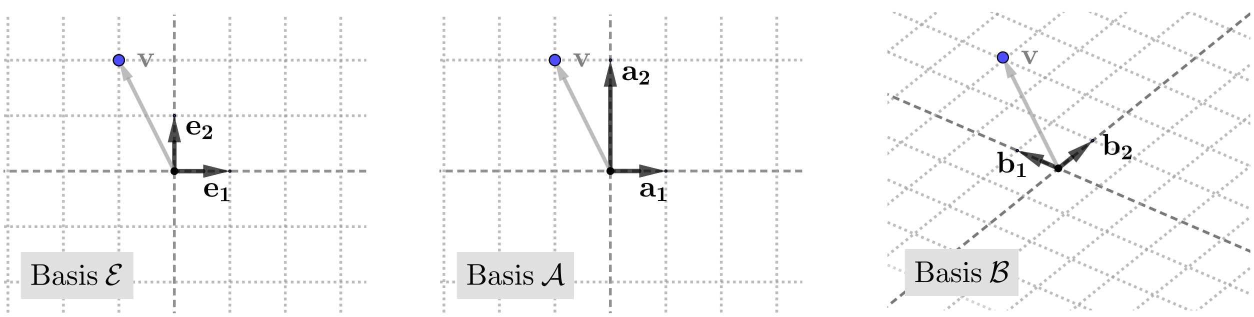
Chapter 6
Change of Basis

Coordinates (and Changing Bases)

> The names of the cerros and the sierras and the deserts exist only on maps. We name them that we do not lose our way. Yet it was because the way was lost to us already that we have made those names. The world cannot be lost. We are the ones. And it is because these names and these coordinates are our own naming that they cannot save us. They cannot find for us the way again.
>
> - Cormac McCarthy, *The Crossing*

So far, we've thought of linear maps as transformations that move or distort all objects in a vector space (vectors, right hands, men on boxes) by moving or distorting the skeleton-like grid that supports them. However, in some circumstances it's useful to think of the space's objects as fixed immobile things, while thinking of the grid no longer as a skeleton, but rather as a ghostly mesh superimposed over the space. Accordingly, we can transform this "ghost grid" as we please without moving or distorting the objects. Such a grid lets us give names (i.e. coordinates) to the space's points, with which we can then calculate. As we'll see in this chapter, we sometimes will want to adjust the axes in this ghost grid so that the resulting coordinates will simplify our subsequent computations.

Let's be more specific. A basis for a vector space determines a grid, which in turn endows each point in the space with coordinates. Change the basis and you change the coordinates of the space's points. To illustrate this idea, each of the figures below shows the same fixed point (and its position vector $\mathbf{v}$), immobile against the shifting background of different "ghost grids" induced by three different bases.



Relative to the grid determined by the standard basis $\mathbf{e_1}, \mathbf{e_2}$ (we'll call the standard basis $\mathcal{E}$ in this chapter), the point's coordinates are $(-1, 2)$. Next, relative to the basis $\mathcal{A}$, which consists of the vectors $\mathbf{a_1}, \mathbf{a_2}$, the point's coordinates are $(-1, 1)$. Finally, relative to basis $\mathcal{B}$ (consisting of $\mathbf{b_1}, \mathbf{b_2}$) its coordinates are $(3, 2)$. As for the position vector $\mathbf{v}$, that symbol refers to the thing itself – the fixed arrow. To express its coordinates as a column vector, we'll need notation specifying the basis to which the coordinates refer: We'll enclose the vector's symbol in brackets and specify the basis with a subscript. (If no specific basis is mentioned, we'll assume, by default, that the standard basis $\mathcal{E}$ is meant.) Thus, in the present case,

$$[\mathbf{v}]_{\mathcal{E}} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \qquad [\mathbf{v}]_{\mathcal{A}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \text{and} \quad [\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

are three different expressions of the coordinates of $\mathbf{v}$, according to three different coordinate systems, determined by three different bases for $\mathbb{R}^2$.

A question naturally arises: Given a vector's coordinates relative to one basis, how can we translate them into its coordinates relative to a different basis? This will be easy to answer after we've pointed out two simple properties of our new subscript notation, which we'll state and prove in the following lemma:

Lemma. For any vectors $\mathbf{v}$ and $\mathbf{w}$, any scalar k, and any basis $\mathcal{A}$, the following properties hold:

$$[\mathbf{v}+\mathbf{w}]_{\mathcal{A}} = [\mathbf{v}]_{\mathcal{A}} + [\mathbf{w}]_{\mathcal{A}} \quad \text{and} \quad [k\mathbf{v}]_{\mathcal{A}} = k[\mathbf{v}]_{\mathcal{A}}$$

Proof. Although these properties are obvious if one thinks about what they mean geometrically (draw some pictures and convince yourself of this), I want to give an algebraic proof to reinforce the algebraic meaning of coordinates. To that end, let the vectors in basis $\mathcal{A}$ be $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. Then we have

$$\mathbf{v} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n \quad \text{and} \quad \mathbf{w} = d_1\mathbf{a}_1 + d_2\mathbf{a}_2 + \cdots + d_n\mathbf{a}_n$$

for some constants c_i and d_i. Consequently,

$$\mathbf{v}+\mathbf{w} = (c_1+d_1)\mathbf{a}_1 + (c_2+d_2)\mathbf{a}_2 + \cdots + (c_n+d_n)\mathbf{a}_n,$$

and

$$k\mathbf{v} = kc_1\mathbf{a}_1 + kc_2\mathbf{a}_2 + \cdots + kc_n\mathbf{a}_n.$$

These last four equations imply that

$$[\mathbf{v}]_{\mathcal{A}} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}, \quad [\mathbf{w}]_{\mathcal{A}} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}, \quad [\mathbf{v}+\mathbf{w}]_{\mathcal{A}} = \begin{pmatrix} c_1+d_1 \\ c_2+d_2 \\ \vdots \\ c_n+d_n \end{pmatrix}, \quad \text{and} \quad k\mathbf{v} = \begin{pmatrix} kc_1 \\ kc_2 \\ \vdots \\ kc_n \end{pmatrix}.$$

Looking at these expressions, the two claimed properties are now obvious. ■

We'll now turn to our main question: Given bases $\mathcal{A}$ and $\mathcal{B}$, how do we translate from $[\mathbf{v}]_{\mathcal{A}}$ to $[\mathbf{v}]_{\mathcal{B}}$? To set the stage, suppose bases $\mathcal{A}$ and $\mathcal{B}$ consist of vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ and $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ respectively, and that $\mathbf{v} = c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n$. Thus, $[\mathbf{v}]_{\mathcal{A}}$ is the column vector consisting of the c_i coefficients. To relate it to $[\mathbf{v}]_{\mathcal{B}}$, we'll apply the $[\quad]_{\mathcal{B}}$ operation to both sides of that expression for $\mathbf{v}$, thus obtaining

$$[\mathbf{v}]_{\mathcal{B}} = [c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + \cdots + c_n\mathbf{a}_n]_{\mathcal{B}}\,.$$

Using our lemma, we can rewrite the right-hand side so that

$$[\mathbf{v}]_{\mathcal{B}} = c_1[\mathbf{a}_1]_{\mathcal{B}} + c_2[\mathbf{a}_2]_{\mathcal{B}} + \cdots + c_n[\mathbf{a}_n]_{\mathcal{B}}\,.$$

The column perspective on matrix-vector multiplication lets us rewrite the right-hand side again:

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} | & | & & | \\ [\mathbf{a}_1]_{\mathcal{B}} & [\mathbf{a}_2]_{\mathcal{B}} & \cdots & [\mathbf{a}_n]_{\mathcal{B}} \\ | & | & & | \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}.$$

But as noted earlier, that column vector on the right is simply $[\mathbf{v}]_{\mathcal{A}}$. Accordingly, we've deduced that

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} | & | & & | \\ [\mathbf{a}_1]_{\mathcal{B}} & [\mathbf{a}_2]_{\mathcal{B}} & \cdots & [\mathbf{a}_n]_{\mathcal{B}} \\ | & | & & | \end{pmatrix} [\mathbf{v}]_{\mathcal{A}}.$$

We've now found what we sought: The matrix above is our "translator" from basis $\mathcal{A}$ to basis $\mathcal{B}$.

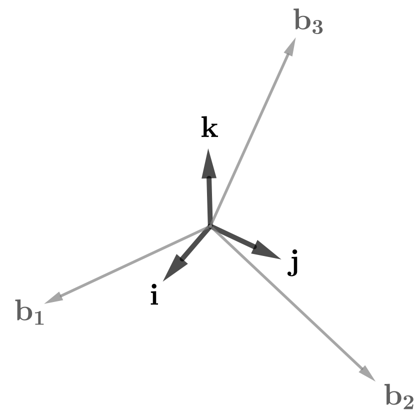
Let's summarize our findings in a box.

> **Theorem.** If $\mathcal{A}$ and $\mathcal{B}$ are bases of a vector space, then the ***$\mathcal{A}$*-to-*$\mathcal{B}$* change of basis matrix** (i.e. the matrix that changes vectors' $\mathcal{A}$-coordinates into their $\mathcal{B}$-coordinates) is… the matrix whose columns are *$\mathcal{A}$'s vectors expressed in $\mathcal{B}$-coordinates*.

It's easy to prove that every change of basis matrix is invertible, and that its inverse changes bases in the *reverse* direction (i.e. if C translates from $\mathcal{A}$ to $\mathcal{B}$, then C^{-1} translates from $\mathcal{B}$ to $\mathcal{A}$), just as we'd expect.*

Example 1. Consider two bases of $\mathbb{R}^3$: the standard basis $\mathcal{E}$ and another, $\mathcal{B}$, consisting of three vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ such that

$$[\mathbf{b}_1]_\mathcal{E} = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \quad [\mathbf{b}_2]_\mathcal{E} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \quad [\mathbf{b}_3]_\mathcal{E} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.$$



Because we have $\mathcal{B}$'s vectors expressed in $\mathcal{E}$-coordinates, the theorem above tells us that the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix (note the order!) is

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

Inverting C (either with the inversion algorithm you learned in Chapter 4 or with a computer), we obtain the $\mathcal{E}$-to-$\mathcal{B}$ change of basis matrix:

$$C^{-1} = \frac{1}{28}\begin{pmatrix} 12 & -4 & -1 \\ 4 & 8 & -5 \\ 0 & 0 & 7 \end{pmatrix}.$$

With these matrices in hand, we can easily translate between the two systems of coordinates. For instance, If $\mathbf{v} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{j}$, then we know $[\mathbf{v}]_\mathcal{E}$, but what is $[\mathbf{v}]_\mathcal{B}$? Well, C^{-1} is our $\mathcal{E}$-to-$\mathcal{B}$ change of basis matrix, so it follows that

$$[\mathbf{v}]_\mathcal{B} = C^{-1}[\mathbf{v}]_\mathcal{E} = \frac{1}{28}\begin{pmatrix} 12 & -4 & -1 \\ 4 & 8 & -5 \\ 0 & 0 & 7 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1/28 \\ 5/28 \\ 21/28 \end{pmatrix}.$$

Or, to go the other way, if $\mathbf{w} = 2\mathbf{b}_1 + \mathbf{b}_3$, then we know $[\mathbf{w}]_\mathcal{B}$, but what is $[\mathbf{w}]_\mathcal{E}$? We can easily compute this using our $\mathcal{B}$-to-$\mathcal{E}$ change of matrix, C. Namely:

$$[\mathbf{w}]_\mathcal{E} = C[\mathbf{w}]_\mathcal{B} = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 3 & 2 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 4 \end{pmatrix}.$$ ♦

*__Proof__: To say that C translates from $\mathcal{A}$ to $\mathcal{B}$ means that $C[\mathbf{v}]_\mathcal{A} = [\mathbf{v}]_\mathcal{B}$ for all $\mathbf{v}$. If C is invertible, we can multiply both sides by C^{-1} to obtain $[\mathbf{v}]_\mathcal{A} = C^{-1}[\mathbf{v}]_\mathcal{B}$, which shows that C^{-1} translates from $\mathcal{B}$ to $\mathcal{A}$ as claimed. But is C invertible? Yes! By our theorem above, we know that C's columns constitute a basis (namely, basis $\mathcal{A}$), so C is invertible by the Invertible Matrix Theorem.

Change of basis matrices can befuddle beginning linear algebra students to such an extent that some textbook authors introduce special idiosyncratic notation for it. (I've seen at least four different types.) I'm of two minds about introducing such notation, so when I teach linear algebra, I strike a compromise: I don't use it "officially", but when explaining something on the board involving change of basis matrices, I sometimes scribble unofficial bookkeeping notes under my matrices to help the class (including me) maintain our collective bearings as we navigate through a thorny problem. I'll show you here what my classroom scratchwork looks like. To signal its unofficial status, I will write it by out by hand, even in the book. If you find it helpful, as I suspect you will, you should incorporate it into your own scratchwork.

I'll start with a simple instance to convey the basic idea. In Example 1, we encountered this equation:

$$[\mathbf{w}]_{\mathcal{E}} = C[\mathbf{w}]_{\mathcal{B}}.$$

If you know that C is the $\mathcal{B}$-to-$\mathcal{E}$ change of matrix (and know what that means), then this equation is clear. But when I want to remind myself (or my students) what a given change of basis matrix does, I supply that information below the matrix – and in parentheses, which signal that it isn't part of the actual notation:

$$[\vec{w}]_{\mathcal{E}} = \underset{(\mathcal{E}\leftarrow\mathcal{B})}{C}[\vec{w}]_{\mathcal{B}}$$

The parenthetical stuff reminds us, by following the arrow, that C is the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix. "OK," you'll probably say. "This reminds me that C will turn $[\mathbf{w}]_{\mathcal{B}}$ into $[\mathbf{w}]_{\mathcal{E}}$, which explains the equation, but why does the arrow point *backwards*?"

The purpose of the backwards arrow will be clear if we consider a *product* of change-of-basis matrices. For example, suppose we have three bases, $\mathcal{A}, \mathcal{B}, \mathcal{E}$, and two change-of-basis matrices:

C: the $\mathcal{A}$-to-$\mathcal{B}$ change of matrix,
D: the $\mathcal{A}$-to-$\mathcal{E}$ change of matrix,

Question: What does DC^{-1} do? The matrices in a product always act on input vectors *from right to left*, so this product applies C^{-1} to an input vector, then applies D to the result. Recall what these matrices do: C^{-1} translates from $\mathcal{B}$ to $\mathcal{A}$, and then D translates from $\mathcal{A}$ to $\mathcal{E}$, so the net effect of their product DC^{-1} is to translate from $\mathcal{B}$-coordinates to $\mathcal{E}$-coordinates. That is, **DC^{-1} is the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix**. Reasoning that out wasn't difficult, but our bookkeeping scratchwork lets you *see* it at a glance:

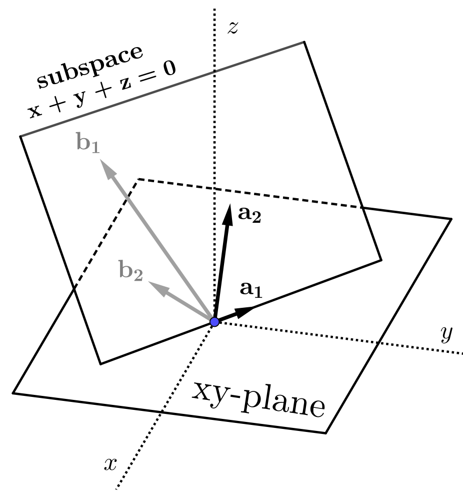
$$\underset{(\mathcal{E}\leftarrow\mathcal{A})}{D}\;\underset{(\mathcal{A}\leftarrow\mathcal{B})}{C^{-1}}$$

We see that any input vector with $\mathcal{B}$-coordinates "enters" on the right, and emerges with $\mathcal{E}$-coordinates. This explains why the arrows are backwards: they are oriented to "go with the flow" of the order in which we apply the matrices to an input vector. Hence, we see immediately that for any vector $\mathbf{v}$,

$$DC^{-1}[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}]_{\mathcal{E}}.$$

Such is my version of change-of-basis bookkeeping scratchwork. I'll use it once more in the next section. But let's put it aside for now and look at another example of finding a change of basis matrix.

This one will involve changing bases in a *sub*space of $\mathbb{R}^n$.



Example 2. The plane $x+y+z=0$ is a two-dimensional subspace of $\mathbb{R}^3$. Here are two bases for the subspace, both expressed with respect to $\mathbb{R}^3$'s *standard* basis $\mathcal{E}$:

$$\mathcal{A}: \quad [\mathbf{a}_1]_\mathcal{E} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad [\mathbf{a}_2]_\mathcal{E} = \begin{pmatrix} -2 \\ 0 \\ 2 \end{pmatrix}.$$

$$\mathcal{B}: \quad [\mathbf{b}_1]_\mathcal{E} = \begin{pmatrix} -1 \\ -2 \\ 3 \end{pmatrix}, \quad [\mathbf{b}_2]_\mathcal{E} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Before reading on, take a moment to verify that these *are* in fact bases for the subspace.

Now let's find the $\mathcal{A}$-to-$\mathcal{B}$ change of basis matrix.

By our theorem, its columns will be basis $\mathcal{A}$'s vectors expressed in $\mathcal{B}$-coordinates.

So to find the first column, we just need to express $\mathbf{a}_1$as a linear combination of $\mathbf{b}_1$ and $\mathbf{b}_2$; the coefficients in that linear combination will, of course, give us $\mathbf{a}_1$'s $\mathcal{B}$-coordinates. This is easy. We just need to solve the equation $x\mathbf{b}_1 + y\mathbf{b}_2 = \mathbf{a}_1$, which is equivalent to an augmented matrix:

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

Carrying out the row reduction, we get

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

so $x = 1$ and $y = -3$. Thus, we've found that $1\mathbf{b}_1 - 3\mathbf{b}_2 = \mathbf{a}_1$, which means that $[\mathbf{a}_1]_\mathcal{B} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$.

We've now found the first column of our change of basis matrix. We can find the second similarly. I'll leave it to you to think and compute your way through that one. In the end, you'll find that $2\mathbf{b}_1 - 4\mathbf{b}_2 = \mathbf{a}_2$. Hence, our $\mathcal{A}$-to-$\mathcal{B}$ change of basis matrix is:

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

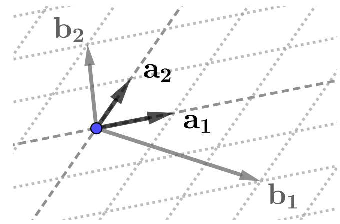
By inverting this, we obtain the $\mathcal{B}$-to-$\mathcal{A}$ change of basis matrix:

$$C^{-1} = \frac{1}{2}\begin{pmatrix} -4 & -2 \\ 3 & 1 \end{pmatrix}.$$

We initially described the basis vectors with three components because we were describing them, rather luxuriantly, with respect to $\mathbb{R}^3$'s standard basis. We were ignoring the fact that they all live in a two-dimensional subspace, which means that when we employ a basis for that subspace, we can describe them with only two coordinates. ♦

Exercises.

1. The figure at right shows two bases for $\mathbb{R}^2$: $\mathcal{A}$ ($\mathbf{a}_1$ and $\mathbf{a}_2$) and $\mathcal{B}$ ($\mathbf{b}_1$ and $\mathbf{b}_2$).



a) Find the $\mathcal{B}$-to-$\mathcal{A}$ change of basis matrix.

b) Find the $\mathcal{A}$-to-$\mathcal{B}$ change of basis matrix.

c) If $[\mathbf{v}]_{\mathcal{A}} = \binom{8}{12}$, what is $[\mathbf{v}]_{\mathcal{B}}$? d) If $[\mathbf{w}]_{\mathcal{B}} = \binom{8}{12}$, what is $[\mathbf{w}]_{\mathcal{A}}$?

e) In the previous part, you saw that $\mathbf{w}$ has a curious property: Switching from $\mathcal{A}$ to $\mathcal{B}$ coordinates (or vice-versa) merely swaps the two numbers (8 and 12) in their coordinate slots. Do any other vectors have this property? If not, why not? If so, which ones? Where do they lie on the graph?

2. Let $\mathbf{b}_1$ and $\mathbf{b}_2$ be an alternate basis $\mathcal{B}$ for $\mathbb{R}^2$, such that $[\mathbf{b}_1]_{\mathcal{E}} = \binom{3}{2}$ and $[\mathbf{b}_2]_{\mathcal{E}} = \binom{1}{2}$.

a) Against the background of the standard coordinate axes, draw vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ and the grid they determine. Mark these points' locations: P, whose $\mathcal{E}$-coordinates are $(4, -1)$, and Q, whose $\mathcal{B}$-coordinates are $(4, -1)$.

b) How would we represent $\mathbf{b}_1$ and $\mathbf{b}_2$ in $\mathcal{B}$-coordinates? [This is easy and requires no calculations.]

c) Find the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix.

d) Find the $\mathcal{E}$-to-$\mathcal{B}$ change of basis matrix.

e) Find the $\mathcal{B}$-coordinates of point P.

f) Find the standard coordinates of point Q.

3. Consider the following three vectors in $\mathbb{R}^3$: $\mathbf{a}_1 = (\mathbf{i} + \mathbf{j})$, $\mathbf{a}_2 = (\mathbf{j} + \mathbf{k})$, $\mathbf{a}_3 = (\mathbf{i} + \mathbf{j} + 2\mathbf{k})$.

a) Do these vectors constitute a basis for $\mathbb{R}^3$? Give at least two different arguments proving your answer. [**Hint**: *The Invertible Matrix Theorem might be useful here*.]

b) Having answered the previous part "yes", call this basis $\mathcal{A}$, and find the $\mathcal{E}$-to-$\mathcal{A}$ change of basis matrix.

c) Use your matrix from the previous part to transform $3\mathbf{i} - \mathbf{j} + 5\mathbf{k}$ into $\mathcal{A}$-coordinates.

4. The points in $\mathbb{R}^4$ whose standard coordinates (x, y, z, w) satisfy the equation $x + 2y + 3z + 4w = 0$ constitute a 3-dimensional subspace. Here are two bases for the subspace, expressed in standard coordinates:

$$\mathcal{A}: \quad [\mathbf{a}_1]_{\mathcal{E}} = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad [\mathbf{a}_2]_{\mathcal{E}} = \begin{pmatrix} 3 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad [\mathbf{a}_3]_{\mathcal{E}} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ -1 \end{pmatrix},$$

$$\mathcal{B}: \quad [\mathbf{b}_1]_{\mathcal{E}} = \begin{pmatrix} 4 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad [\mathbf{b}_2]_{\mathcal{E}} = \begin{pmatrix} -2 \\ 0 \\ 2 \\ -1 \end{pmatrix}, \quad [\mathbf{b}_3]_{\mathcal{E}} = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

a) Verify that these are in fact bases for the subspace.

b) Before computing anything, how many rows and columns will the change of basis matrix from $\mathcal{B}$ to $\mathcal{A}$ have? [**Hint**: Just think about what the matrix does.]

c) Find the $\mathcal{B}$-to-$\mathcal{A}$ change of basis matrix.

[**Note**: As you think your way through this, you'll find that you must solve three linear systems of the same form: $(M|\,\mathbf{v})$, where the matrix on the left is the same in all three cases, while the vector on the right varies. Rather than going through the same tedious process of row-reducing M on three separate occasions to solve $(M|\,\mathbf{v}_1)$, $(M|\,\mathbf{v}_2)$, and $(M|\,\mathbf{v}_3)$, you can do all three at once by row-reducing $(M|\,\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$. We discussed this trick in Chapter 4, at the beginning of the section called "The Matrix Inversion Algorithm".]

5. In Chapter 1, you learned that if we are given any two vectors

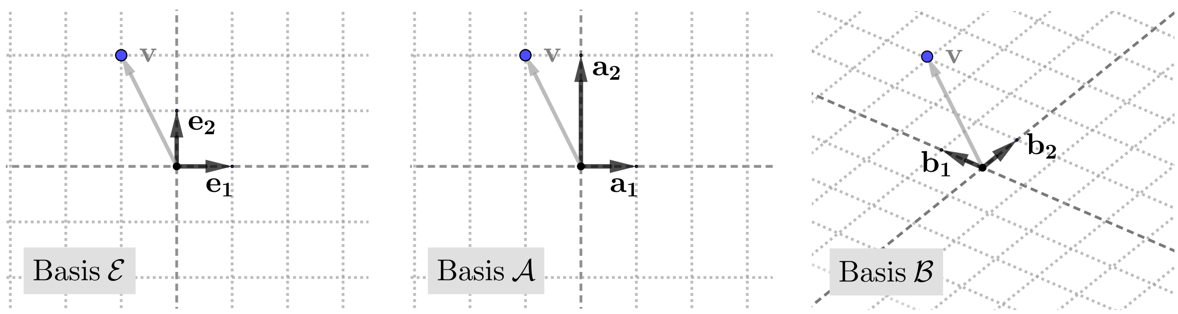
$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \quad \text{and} \quad \mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

in $\mathbb{R}^n$, we can compute their lengths and their dot product in terms of their components. Namely,

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \quad \text{and} \quad \mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \cdots v_n w_n.$$

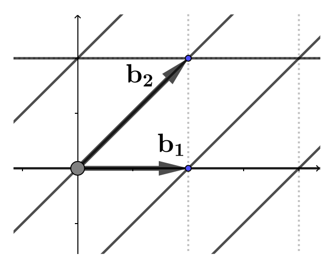
However, we derived these two formulas under the assumption that those coordinates are *Cartesian* – or as we'd say now, under the assumption that those column vectors have been expressed in terms of the *standard* basis $\mathcal{E}$. Those two formulas usually do **not** hold in other coordinate systems. You'll see some counterexamples in this exercise.

a) The three figures from this chapter's first page show a vector $\mathbf{v}$ against three different coordinate grids:



Vector $\mathbf{v}$ remains constant as the backgrounds change behind it. It does not change, so neither does its length. However, the square root of the sum of its squared coordinates *does* change when we express it in $\mathcal{A}$ or $\mathcal{B}$ coordinates. Show this, and then remember the important lesson: If a vector's coordinates aren't Cartesian, then the usual length formula (square root of the sum of a vector's squared coordinates) typically *doesn't* work. The number it produces usually *isn't* the vector's length; it is usually devoid of geometric meaning.

b) Recall that we did not define the dot product via a formula, as many books do. Instead, we defined the dot product of two vectors $\mathbf{v}$ and $\mathbf{w}$ geometrically – as *the product of their scalar projections onto* $\mathbf{w}$. From that coordinate-free definition, we derived almost every important fact about dot products, which means that almost everything we proved about dot products holds regardless of our coordinate system. The one exception to this was the coordinate-based formula for dot products above. Our derivation of that formula was *not* coordinate free; it specifically presumed that the coordinates were *Cartesian*.



It should therefore be no surprise that when we use nonstandard coordinates, the "dot product formula" (the sum of the products of corresponding coordinates) does *not* actually yield the dot product (i.e. the product of scalar projections). To construct a counterexample, let's consider the basis $\mathcal{B}$ of $\mathbb{R}^2$ shown in the figure, consisting of $\mathbf{b_1} = \mathbf{i}$ and $\mathbf{b_2} = \mathbf{i} + \mathbf{j}$. By our geometric definition of the dot product, we know that $\mathbf{i} \cdot \mathbf{j} = 0$. But if we express $\mathbf{i}$ and $\mathbf{j}$ in $\mathcal{B}$-coordinates and then we sum up the products of these vectors' corresponding $\mathcal{B}$-coordinates, we will *not* obtain their dot product (which is 0). Verify this, and then draw the important moral: If two vectors' coordinates aren't Cartesian, then the usual dot product formula (sum of the products of the respective coordinates) typically *doesn't* work. The number it produces usually *isn't* the vectors' dot product, as we've defined it. It is usually devoid of geometric meaning.

Matrices of Maps (via Nonstandard Bases)

It is not down in any map. True places never are.
- Herman Melville, *Moby Dick* (Ch. 12, "Biographical")

When learning to represent linear maps as matrices in Chapter 3, you worked only with the standard basis. Ah, for those carefree days of youth! In fact, every linear map has *infinitely many* matrix representations, one for each choice of basis. These different representations are like photos of the same subject (the map) taken from different perspectives. Sometimes a nonstandard representation yields such a flattering photo – a matrix with especially nice algebraic properties – that we favor it over the standard representation. And as we'll soon see, nonstandard representations can help us find standard representations, too.

Let's introduce some symbols. Suppose T is a linear transformation of $\mathbb{R}^n$ and that $\mathcal{B}$ is a basis for $\mathbb{R}^n$. Then $[T]_{\mathcal{B}}$ denotes ***T*'s matrix relative to basis $\mathcal{B}$**, defined as the matrix that changes $\mathbf{v}$'s $\mathcal{B}$-coordinates into $T(\mathbf{v})$'s $\mathcal{B}$-coordinates for every vector $\mathbf{v}$. (Or in symbols, $[T]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{B}}$ for all $\mathbf{v}$.)

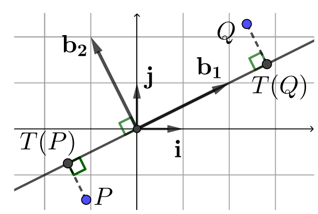
But how do we construct such a matrix? We know that in T's matrix relative to the *standard* basis $\mathcal{E}$, the columns are the images of $\mathcal{E}$'s vectors… expressed in $\mathcal{E}$-coordinates. This statement generalizes in the obvious way to any basis whatsoever, making it easy to remember:

Matrix Representation of T Relative to Basis $\mathcal{B}$.

The columns are the images of $\mathcal{B}$'s vectors… expressed in $\mathcal{B}$-coordinates.

(Or in symbols, but don't memorize it this way, $[T]_{\mathcal{B}}$'s i^{th} column is $[T(\mathbf{b}_i)]_{\mathcal{B}}$, where $\mathbf{b}_i$ is basis $\mathcal{B}$'s i^{th} vector.)

Example 1. Let T be an *orthogonal projection* of $\mathbb{R}^2$ onto $y = x/2$. This means that from every point P in $\mathbb{R}^2$, we drop a perpendicular to the line $y = x/2$. The perpendicular's foot is P's image, $T(P)$.



Finding T's *standard* matrix directly would be a bit of a slog here, requiring some analytic geometry to determine the projected images of $\mathbf{i}$ and $\mathbf{j}$.

In contrast, T's *non*standard matrix relative to basis $\mathcal{B}$ (vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ in the figure) is easy. The first column of $[T]_{\mathcal{B}}$ is $\mathbf{b}_1$'s image *expressed in $\mathcal{B}$-coordinates*. It's geometrically obvious that $T(\mathbf{b}_1) = \mathbf{b}_1$. To express this image in $\mathcal{B}$-coordinates, we note that $T(\mathbf{b}_1) = \mathbf{b}_1 = 1\mathbf{b}_1 + 0\mathbf{b}_2$, so

$$[T(\mathbf{b}_1)]_{\mathcal{B}} = \begin{pmatrix}1\\0\end{pmatrix}.$$

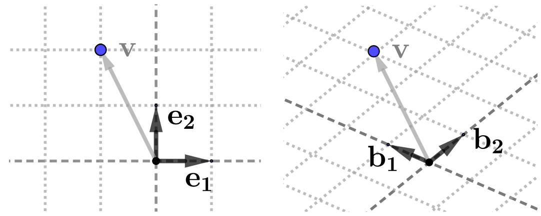
This will be the *first* column of our nonstandard matrix $[T]_{\mathcal{B}}$. Its second column will be $\mathbf{b}_2$'s image *expressed in $\mathcal{B}$-coordinates*. It's geometrically clear that $T(\mathbf{b}_2) = \mathbf{0}$, which equals $0\mathbf{b}_1 + 0\mathbf{b}_2$, so

$$[T(\mathbf{b}_2)]_{\mathcal{B}} = \begin{pmatrix}0\\0\end{pmatrix}.$$

And now we have our nonstandard matrix representation of T relative to $\mathcal{B}$:

$$[T]_{\mathcal{B}} = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}. \quad \blacklozenge$$

Next, let's see how a nonstandard representation of a map can lead us to its standard representation. It helps to think of a vector's coordinate representations as names for the vector in different languages. For instance, the pictures at right show a vector $\mathbf{v}$ against "ghost grid" backgrounds generated by two different bases, $\mathcal{E}$ and $\mathcal{B}$, which we might playfully think of as $\mathcal{E}$nglish and $\mathcal{B}$ulgarian. Vector $\mathbf{v}$ has different names in these two languages:



$$[\mathbf{v}]_{\mathcal{E}} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \quad \text{and} \quad [\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Now, let's suppose that our goal is to find the *standard* matrix representation of some linear map T of $\mathbb{R}^2$. That is, we want a matrix that will let us apply the map T to vectors "in English", turning $[\mathbf{v}]_{\mathcal{E}}$ into $[T(\mathbf{v})]_{\mathcal{E}}$. Let's suppose further that we already know a *non*standard matrix representation of the map, $[T]_{\mathcal{B}}$. This "Bulgarian matrix" would let us apply T to $\mathbf{v}$ *in Bulgarian*, because we know that $[T]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{B}}$. If only we knew Bulgarian!* Actually, we can hack this matrix: all we'll need is a reliable Bulgarian/English translation program, for then we could achieve our goal (turning $[\mathbf{v}]_{\mathcal{E}}$ to $[T(\mathbf{v})]_{\mathcal{E}}$) in three easy steps:

(1) Translate $[\mathbf{v}]_{\mathcal{E}}$ into $[\mathbf{v}]_{\mathcal{B}}$. (English to Bulgarian)
(2) Use the "Bulgarian matrix" to apply the map, sending $[\mathbf{v}]_{\mathcal{B}}$ to $[T(\mathbf{v})]_{\mathcal{B}}$. (Map in Bulgarian)
(3) Translate $[T(\mathbf{v})]_{\mathcal{B}}$ into $[T(\mathbf{v})]_{\mathcal{E}}$. (Bulgarian to English)

Excellent: The net result would indeed be the English-to-English transformation we seek, but… it depends vitally on a "Bulgarian/English translation program". Do we have a mathematical "translation program"? Yes! That would be a *change of basis matrix* (and its inverse). If C is the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix, then our three-step plan for mapping $[\mathbf{v}]_{\mathcal{E}}$ to $[T(\mathbf{v})]_{\mathcal{E}}$ can be accomplished as follows:

(1) Apply C^{-1} (translate from $\mathcal{E}$ to $\mathcal{B}$)
(2) Apply $[T]_{\mathcal{B}}$ (apply the map)
(3) Apply C (translate back from $\mathcal{B}$ to $\mathcal{E}$)

Successively applying those matrices to a vector's standard "English" coordinates $[\mathbf{v}]_{\mathcal{E}}$ maps it to $[T(\mathbf{v})]_{\mathcal{E}}$. It might help to see this expressed in the unofficial bookkeeping style I mentioned in the last section:

$$\Big(\underset{(\mathcal{E} \leftarrow \mathcal{B})}{C}[T]_{\mathcal{B}}\underset{(\mathcal{B} \leftarrow \mathcal{E})}{C^{-1}}\Big)[\vec{v}]_{\mathcal{E}} = [T(\vec{v})]_{\mathcal{E}}$$

Follow the flow of the left side's subscripts (reading from right to left): We see a vector's $\mathcal{E}$-coordinates translated to $\mathcal{B}$-coordinates, so that T (expressed relative to $\mathcal{B}$) can act on them; once T maps them, the result gets translated back to $\mathcal{E}$-coordinates. The net result will thus be the $\mathcal{E}$-coordinates of the input vector's image under T. And that, of course, is what we see on the equation's right side: $[T(\mathbf{v})]_{\mathcal{E}}$.

The matrix product on the left, $C[T]_{\mathcal{B}}C^{-1}$, is thus T's standard matrix representation. In symbols,

$$[T]_{\mathcal{E}} = C[T]_{\mathcal{B}}C^{-1}.$$

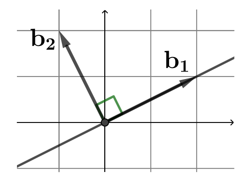
Placing C and C^{-1} around $[T]_{\mathcal{B}}$ like a pair of headphones lets us "hear" the Bulgarian matrix in English.

* Благодаря на Бога за онлайн програмите за превод, нали?

To make this technique more concrete, let's apply it to our orthogonal projection from Example 1.

Example 2. Let T be the *orthogonal projection* of $\mathbb{R}^2$ onto $y = x/2$.
Find T's *standard* matrix representation.

Solution. We begin by choosing a suitable basis – one that makes it easy to find a matrix representation. We've already seen, in Example 1, that basis $\mathcal{B}$ ($\mathbf{b}_1$ and $\mathbf{b}_2$ in the figure) fits this bill nicely. More specifically, we saw that



$$[T]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Now that we have a *non*standard matrix representation $[T]_{\mathcal{B}}$, we'll be able to convert it into T's standard matrix representation $[T]_{\mathcal{E}}$ once we have the two $\mathcal{B} \leftrightarrow \mathcal{E}$ change of basis matrices. It's clear from the figure that the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix is

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

Inverting this matrix gives us the $\mathcal{E}$-to-$\mathcal{B}$ change of basis matrix:

$$C^{-1} = \frac{1}{5}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}.$$

We can now put our "change of basis headset" around $[T]_{\mathcal{B}}$, and we'll be done. To ensure that we're doing everything in the right order, we'll take a minute to remind ourselves why this works. We want the standard matrix, which changes $[\mathbf{v}]_{\mathcal{E}}$ into $[T(\mathbf{v})]_{\mathcal{E}}$. We can get it in three easy steps: Translate $[\mathbf{v}]_{\mathcal{E}}$ into $\mathcal{B}$ulgarian (using C^{-1}), then carry out the map T in Bulgarian (using $[T]_{\mathcal{B}}$), and finally, translate the result back into $\mathcal{E}$nglish (using C). We therefore conclude that

$$[T]_{\mathcal{E}} = C[T]_{\mathcal{B}}C^{-1} = \begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\left[\frac{1}{5}\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix}\right].$$

Pulling the scalar up front, and then using associativity of matrix multiplication, we rewrite this as

$$\frac{1}{5}\left[\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right]\begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix},$$

After doing the matrix multiplication, we conclude – as you should verify – that

$$[T]_{\mathcal{E}} = \frac{1}{5}\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}.$$ ♦

It's geometrically obvious the orthogonal projection in Example 2 maps the point (2,1) to itself. Thus, as a quick check of our standard matrix representation, let's make sure that $[T]_{\mathcal{E}}$ actually does this:

$$[T]_{\mathcal{E}}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 10 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

This checks out, increasing our confidence that we didn't make an arithmetic mistake in the computations that led us to $[T]_{\mathcal{E}}$.

Exercises.

6. Let $\mathcal{B}$ be an alternate basis for $\mathbb{R}^2$ consisting of $\mathbf{b}_1$ and $\mathbf{b}_2$, where $[\mathbf{b}_1]_\mathcal{E} = \binom{3}{2}$ and $[\mathbf{b}_2]_\mathcal{E} = \binom{1}{2}$.

a) If we let $[R]_\mathcal{E}$ be the *standard* matrix representation a $90°$ **clockwise** rotation about the origin, and we let C be the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix, express $[R]_\mathcal{B}$ in terms of $[R]_\mathcal{E}$ and C.

b) Compute the matrix $[R]_\mathcal{B}$.

c) Draw a sketch showing the point P whose $\mathcal{B}$-coordinates are $(-1, -2)$. (No calculations needed; just draw the grid determined by the $\mathcal{B}$-basis.) Then mark the approximate point P' on your figure to which P is mapped by a $90°$ clockwise rotation about the origin. Find the exact $\mathcal{B}$-coordinates of P' by using a matrix that you found in the previous part of this problem. Find the *standard* coordinates of P', too.

d) Find the $\mathcal{B}$-coordinates of the point that this rotation maps *to* the point whose $\mathcal{B}$-coordinates are $(1, 1)$.

7. A useful fact when building a basis of two *perpendicular* vectors in $\mathbb{R}^2$: If a point's standard coordinates are (a, b), rotating it counterclockwise by $90°$ will take it to a point with standard coordinates $(-b, a)$. Explain why.

8. Use a change of basis to find the *standard* matrix representation of reflection across the line $y = 2x$ in $\mathbb{R}^2$.

9. Find the matrix representing orthogonal projection onto the x-axis… relative to basis $\mathcal{B}$: vectors $\mathbf{b}_1$ and $\mathbf{b}_2$, where

$$[\mathbf{b}_1]_\mathcal{E} = \binom{2}{-1} \quad \text{and} \quad [\mathbf{b}_2]_\mathcal{E} = \binom{1}{3}.$$

10. We say that matrix A is **similar** to matrix B if "surrounding" A with some matrix C and its inverse turns it into B. (More formally, A is similar to B if $CAC^{-1} = B$ for some matrix C.) Similar matrices represent the same underlying linear transformation, but with respect to different bases.*

a) Prove algebraically that similarity is a *symmetric* relation, meaning if A is similar to B, then B is similar to A.

b) Prove algebraically that similarity is *transitive*, meaning two matrices similar to a third are themselves similar.

c) Find two matrices that are similar to $M = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$.

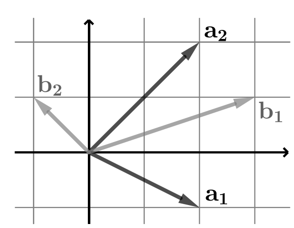
d) Similar matrices have the same determinant. Explain why. Then verify that this is true in the case of the two matrices that you found in the previous part.

e) Similar matrices have the same rank. Explain why. Then verify that this is true in the case of the two matrices that you found in the previous part.

f) The definition of similar matrices makes sense only when A and B are *square* matrices. Why?

11. Suppose T is a linear transformation of $\mathbb{R}^n$, and that $[T]_\mathcal{B}$ is a *diagonal* matrix relative to some basis $\mathcal{B}$. Explain geometrically what T does to $\mathcal{B}$'s basis vectors.
(In Chapter 7, we'll specifically seek out bases that yield diagonal matrix representations.)

12. The figure at right shows two bases of $\mathbb{R}^2$ – basis $\mathcal{A}$, which consists of $\mathbf{a}_1$ and $\mathbf{a}_2$, and basis $\mathcal{B}$, which consists of $\mathbf{b}_1$ and $\mathbf{b}_2$ – superimposed on the same standard grid.



a) Find the $\mathcal{A}$-to-$\mathcal{B}$ change of basis matrix.
[**Hint**: It may help to use a third basis as a stepping stone.]

b) Find the $\mathcal{B}$-to-$\mathcal{A}$ change of basis matrix.

c) Let T be the linear map that stretches $\mathbf{b}_1$ by a factor of 3 and $\mathbf{b}_2$ by a factor of 4. Find the matrix representations of T relative to $\mathcal{B}$, $\mathcal{A}$, and $\mathcal{E}$.

* **Proof**: Let A be an $n \times n$ matrix. Every matrix is the standard matrix of a linear map, so let T be the map such that $A = [T]_\mathcal{E}$. Let C be any *invertible* $n \times n$ matrix. By the Invertible Matrix Theorem, its columns (thought of as vectors' $\mathcal{E}$-coordinates) constitute a basis for $\mathbb{R}^n$. Call this basis $\mathcal{B}$. Then C is the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix, C^{-1} is the $\mathcal{E}$-to-$\mathcal{B}$ change of basis matrix, and $CAC^{-1} = [T]_\mathcal{B}$. Thus if A and B are similar, A and B represent the same linear map, but with respect to different bases. ■

Linear Isometries, Orthogonal Matrices

MR. FRIEDMAN: I think that issue is entirely orthogonal to the issue here because the Commonwealth is acknowledging –

CHIEF JUSTICE ROBERTS: I'm sorry. Entirely what?

MR. FRIEDMAN: Orthogonal. Right angle. Unrelated. Irrelevant.

CHIEF JUSTICE ROBERTS: Oh.

- Oral argument at the US Supreme Court in *Briscoe v. Virginia* (1/11/2010)

A **linear isometry** is a linear map that preserves all distances between points. Examples include rotations and reflections. (That these *are* isometries should be clear: If points P and Q are, say, 4 units apart, then any rotation – or reflection – will obviously send them to points P' and Q' that are *still* 4 units apart.) By definition, linear isometries preserve distances, but it's easy to prove that they also preserve *angles*.*

What does an isometry do to the standard basis vectors? An isometry preserves lengths and angles, so it must map the standard basis frame onto some other frame of *mutually perpendicular unit vectors*. Conversely, any linear map that takes the standard basis frame to a new frame of mutually perpendicular unit vectors is an isometry.† It follows that a square matrix has mutually perpendicular unit-length columns if and only if it represents a linear isometry (relative to the standard basis).

Consequently, the following matrices must all represent isometries, since their columns are unit length and mutually perpendicular. You should verify these properties for for each matrix:

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}, \quad \begin{pmatrix} \cos(22°) & -\sin(22°) \\ \sin(22°) & \cos(22°) \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{3}/2 & -1/2 \\ 0 & 1/2 & \sqrt{3}/2 \end{pmatrix}.$$

The phrase "square matrix whose columns are mutually perpendicular unit vectors" is unwieldy, so we define a new term: An **orthogonal matrix** is a square matrix whose columns are mutually perpendicular unit vectors. Using this terminology, we'd say that all the matrices above are orthogonal matrices, and we can now restate our result from the previous paragraph more crisply:

A is an **orthogonal matrix** $\Leftrightarrow$ A is the standard representation of a **linear isometry**.

Orthogonal matrices thus have geometric significance. They have a nice algebraic property, too, which makes them prized as change of basis matrices: An orthogonal matrix's inverse is simply its *transpose*. Inverting a matrix is normally a computational slog – but not with orthogonal matrices! Let's prove this.

* **Proof**: If we subject any angle $\angle PQR$ to an isometry, sending P, Q, R to P', Q', R', then by definition, line segments PQ, PR, and QR have the same lengths as their images $P'Q', P'R', Q'R'$. Hence, triangles ΔPQR and $\Delta P'Q'R'$ are congruent (by the SSS-congruence criterion), so it follows that $\angle PQR = \angle P'Q'R'$. Thus, angles are preserved by isometries, as claimed. ■

† **Proof**: Given such a linear map, call its "transformed" basis vectors $\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n$. These are mutually perpendicular unit vectors, so we can clearly "reach it" from the *standard* basis frame as follows: First, it's intuitively clear that we can rotate the standard basis frame until it is superimposed over the transformed frame. (Think about this in $\mathbb{R}^2$ and $\mathbb{R}^3$.) Having done this, check to see if each standard basis vector $\mathbf{e}_i$ "matches up" with the one that is supposed to be its transformed image, $\mathbf{e}'_i$. If this *is* the case, then we're done – the map is just the rotation, which is of course an isometry. If not, then we can use reflections to swap the locations of "mismatched" vectors until everything matches. In that case, the linear map is a composition of the initial rotation and some number of reflections, all of which are isometries. Since the composition of isometries is obviously an isometry, the map must be an isometry as claimed. ■

Theorem. If Q is an orthogonal matrix, then $Q^{-1} = Q^T$.

Proof. We'll show that $Q^T Q = I$, thus demonstrating that $Q^{-1} = Q^T$. To this end, note that

$Q^T Q$'s ij^{th} entry
$= (Q^T$'s i^{th} row) $\cdot$ (Q's j^{th} column) (matrix multiplication's "entry perspective")
$= (Q$'s i^{th} column) $\cdot$ (Q's j^{th} column) (definition of the transpose).

Matrix Q is orthogonal, which means that its columns are mutually perpendicular unit vectors. Because of that mutual perpendicularity, the last dot product above will always be 0 when $i \neq j$. For the cases when $i = j$, we're dotting a column with itself. Recall that dotting any vector with itself yields the square of that vector's length. The columns here are all unit vectors, so each has a squared length of 1. Thus, we've shown that $Q^T Q$ is a matrix whose ij^{th} entry is 0, except when $i = j$, in which case it is 1. In other words, $Q^T Q = I$,* which means that $Q^{-1} = Q^T$, as claimed. ■

In fact, the converse is also true: if some matrix's inverse is its transpose, then the matrix is orthogonal. (The proof is basically the same, but it runs in the opposite direction.)

Example. Consider the basis $\mathcal{B}$ whose vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ are such that

$$[\mathbf{b}_1]_{\mathcal{E}} = \begin{pmatrix} 1/3 \\ 2/3 \\ 2/3 \end{pmatrix}, \quad [\mathbf{b}_2]_{\mathcal{E}} = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}, \quad [\mathbf{b}_3]_{\mathcal{E}} = \begin{pmatrix} 2/3 \\ -2/3 \\ 1/3 \end{pmatrix},$$

These three vectors are, as you should verify, *mutually perpendicular unit vectors*.† Consequently,

$$Q = \begin{pmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{pmatrix}$$

Is an orthogonal matrix. It follows from the theorem above that

$$Q^{-1} = Q^T = \frac{1}{3}\begin{pmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{pmatrix}.^{\ddagger}$$

Note that Q is the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix, and Q^{-1} is the $\mathcal{E}$-to-$\mathcal{B}$ change of basis matrix. ♦

Orthogonal matrices can sometimes help us grasp the geometry of linear maps that aren't isometries themselves, but which can be "factored" into simpler maps, some of which *are* isometries. You'll see a glimpse of that in Exercise 14, but to understand it more fully, you'll need to learn about all things "eigen". You'll do this in the next chapter, but before we dive in, you have another set of exercises to complete.

---

* Must we also show that $QQ^T = I$? No. By Ch 5, Exercise 12, a square matrix's 'left inverse' is automatically its 'right inverse'.

† Here in the snug privacy of the footnotes, I'll admit that linear algebra's terminology is needlessly confusing on one point here. A basis of mutually perpendicular vectors is called, reasonably enough, an **orthogonal basis**. If all the vectors in an orthogonal basis are also *unit length* (e.g. this example's basis $\mathcal{B}$) we call it an **ortho*normal* basis**. ("Ortho" for perpendicular, "normal" for unit length - as in "normalizing" a vector, discussed in Exercise 8 of Chapter 1). So far so simple. But then... What do we call an $n \times n$ matrix whose columns constitute an orthonormal basis for $\mathbb{R}^n$? We should call it an ortho*normal* matrix, right? Well of course we should, but we don't. We call it an **orthogonal matrix**, as you've seen already. Unfortunately, it's too late to change this awkward but well-established convention, but like so many people before you, you'll get used to it. Such is life.

‡ Here I've used the fairly obvious fact that $(cM)^T = cM^T$ for any matrix M and scalar c. Convince yourself that this *is* obvious! (You might as well do this now, since you'll see it again as Exercise 15a on the next page.)

Exercises.

13. Some questions about orthogonal matrices:

a) If A is an orthogonal matrix, what can we say about its determinant? Explain your answer.
b) If $\det B = 1$, does it follow that B must be orthogonal? If so, prove it. If not, give a counterexample.
c) Suppose C represents some sort of mysterious rotation of 10-dimensional space about the origin. What can we say about $C^T C$? Explain.
d) Make up a few examples of orthogonal 2×2 and 3×3 matrices.
e) Curious fact: If a matrix A is orthogonal, then so is A^T. (Or stated another way, the *rows* of every orthogonal matrix are mutually perpendicular unit vectors, too.) Explain why.
f) Verify the truth of Part E by checking the examples you produced in Part D.
g) If A and B are orthogonal $n \times n$ matrices, explain why their product AB is also orthogonal.
h) The Devil gives you a 666 x 666 matrix and won't let you leave his evil realm until you've inverted it by hand and presented him with the correct inverse matrix. "Now you'll be here for centuries doing Gaussian elimination! What exquisite torture!" the Evil One gloats. "What does this damned thing represent anyway?" you ask, half resigned to several lifetimes worth of arithmetic. "A special reflection in 666-dimensional space!" he cackles. "Too subtle for the likes of you!" Hearing this, you smile, knowing that you'll be free sooner than Lucifer supposes. Why?
(Also, if you can write one matrix entry per second, how long will it take to invert the diabolic matrix?)

14. Consider the matrix

$$A = \frac{1}{3}\begin{pmatrix} 7 & -2 & 0 \\ -2 & 6 & -2 \\ 0 & -2 & 5 \end{pmatrix}.$$

If we think of A as representing some linear transformation T relative to the standard basis so that $A = [T]_{\mathcal{E}}$, it's hard to get an intuitive feel for what this linear map does geometrically to $\mathbb{R}^3$. But with a well-chosen change of basis, we can understand this map's effect on $\mathbb{R}^3$ much better.

a) Find $[T]_{\mathcal{B}}$, where $\mathcal{B}$ is the basis in the previous page's example. (You should find that $[T]_{\mathcal{B}}$ is a *diagonal* matrix.)
b) Now try to describe T's geometric effect in terms of its action on the vectors of basis $\mathcal{B}$.
c) **(Rhetorical questions for the road)** Given an arbitrary linear map, is there always a basis relative to which the map has a *diagonal* matrix? When such a basis exists, how do we find it? For example, given matrix A above, how could you have found the "diagonalizing" basis $\mathcal{B}$, on your own? *You'll learn the answers in Chapter 7.*

15. Explain why each of the following statements is true. For any $n \times n$ matrix A and any scalar c…

a) $(cA)^T = cA^T$
b) $(cA)^{-1} = c^{-1}A^{-1}$
c) $\det(cA) = c^n \det(A)$.

Chapter 7

Eigenstuff

Eigenvectors and their Eigenvalues

Keep on a straight line… I don't believe I can.
Trying to find a needle in a haystack –
Chilly wind, you're piercing like a dagger, it hurts me so.
Nobody needs to discover me. I'm back again.

- Peter Gabriel as a world-weary eigenvector in "Looking for Someone" (from Genesis's album *Trespass*).

German-English dictionaries will tell you that *eigen* means "characteristic", "particular", or "[one's] own", but these don't fully convey *eigen*'s mathematical significance. In linear algebra, *eigen* means something like "essence-revealing". A linear map's *eigen* things (eigenvectors, eigenvalues, eigenbasis, eigenspaces) reveal its geometric essence.

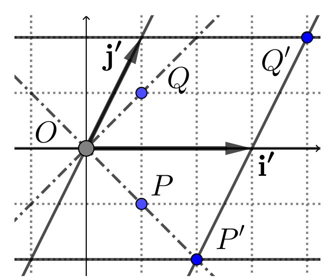
When blasted by the chilly wind of a linear map, most vectors get blown off the line on which they lie, but a few, hidden like needles in the vector haystack, manage to "keep on a straight line"; the map merely *scales* them. We call these the linear map's *eigenvectors*. Whenever we can discover a basis for a space made up of eigenvectors relative to a map, we have grasped the map's geometric action on the space: It simply scales the space by various factors along various axes. (As for vectors that *don't* lie on the axes, the map just sums their scaled *components* relative to those axes.) Now for our formal definitions.

Definitions. A nonzero vector $\mathbf{v}$ is an **eigenvector** of a linear map T (or matrix A) if there is some scalar λ such that $T(\mathbf{v}) = \lambda\mathbf{v}$. (Or $A\mathbf{v} = \lambda\mathbf{v}$.)

The scalar λ is called an **eigenvalue** of the map (or matrix).
(We also say it is the eigenvalue corresponding to the eigenvector $\mathbf{v}$.)

Some examples will clarify these two simple definitions.

Example 1. Let $A = \begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$, a matrix whose action is depicted below.



First, let's see what A does to the vector corresponding to point Q:

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

Since the output vector is *not* a scalar multiple of the input vector (equivalently, since A maps point Q to Q', knocking it off the line OQ), we conclude that

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ is } \textit{not} \text{ an eigenvector of } A.$$

Now let's consider A's effect on the vector corresponding to point P:

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

Here, the output vector is just the input vector scaled by 2. Thus,

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ is an eigenvector of } A, \text{ with eigenvalue } 2.$$

Another eigenvector of A is $\mathbf{i}$ itself; as the figure shows, it gets scaled by 3. Thus, $\mathbf{i}$ is an eigenvector of A with eigenvalue 3. ♦

Example 2. Let T be a counterclockwise rotation of $\mathbb{R}^2$ about the origin by 90°. Obviously, *every* nonzero vector changes its direction when subjected to this rotation (none of them manages to "keep on a straight line") so this rotation has no eigenvectors. ♦

Example 3. Let R be a reflection of $\mathbb{R}^3$ across the xy-plane. Thinking about this geometric operation, you should be able to convince yourself of the following three things:

- All vectors in the xy-plane are eigenvectors of R (with eigenvalue 1).
- All vectors lying along the z-axis are eigenvectors of R (with eigenvalue -1).
- No other vectors in $\mathbb{R}^3$ are eigenvectors.

Be sure you can understand the three preceding statements by visualizing the reflection. ♦

Now that you know what eigenvectors and eigenvalues are, here is our next big definition:

> **Definition**. An **eigenbasis** relative to an $n \times n$ matrix (or linear transformation of $\mathbb{R}^n$) is a basis of $\mathbb{R}^n$ consisting entirely of eigenvectors of the matrix/map.

We love eigenbases because if we represent a map relative to an eigenbasis, we get a *diagonal* matrix.* We love diagonal matrices because their geometric action is intuitive (mere scaling along the basis "axes") and their algebraic properties are nice (their inverses, determinants, and powers are all easy to compute).

Looking back at Example 1, we see that $\mathbb{R}^2$ has - relative to matrix A - the following eigenbasis:

$$\text{Eigenbasis } \mathcal{B}: \quad \mathbf{b}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Matrix A merely scales these two linearly independent vectors by factors of 2 and 3 respectively. If we let T be the map whose matrix representation relative to the *standard* basis is A, then T's representation relative to the *eigen*basis $\mathcal{B}$, will - as promised - be diagonal:

$$[T]_{\mathcal{B}} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

We were able to get this diagonal representation only because the map had enough linearly independent eigenvectors to form an eigenbasis. Not all maps are like this.

When a matrix or map does have enough linearly independent eigenvectors to form an eigenbasis, we say that it is **diagonalizable**, since we can then use that eigenbasis to represent it as a diagonal matrix. Thus, A from Example 1 is diagonalizable.

In contrast, the 90° rotation in Example 2 has no eigenvectors at all, so there clearly can't be an eigenbasis for $\mathbb{R}^2$ relative to it. Equivalently, a 90° rotation in the plane is *not* a diagonalizable map.

In Example 3, reflection R (reflection in the xy-plane) *is* diagonalizable, since it admits an eigenbasis. In fact, the *standard* basis is an eigenbasis here, so R's standard matrix representation is diagonal:

$$[R]_{\mathcal{E}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

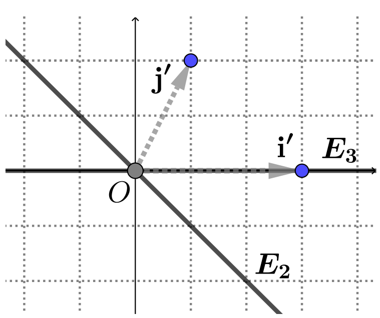
---

* **Proof**: Suppose $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is an eigenbasis $\mathcal{B}$ (with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$) relative to some linear map T. Then for each i, we have $T\mathbf{b}_i = \lambda_i \mathbf{b}_i$, so $[T\mathbf{b}_i]_{\mathcal{B}}$ is the column vector whose ith entry is λ_i and whose other entries are all zeros. Since $[T]_{\mathcal{B}}$ consists of these column vectors, it is diagonal as claimed. Moreover, the diagonal entries are T's eigenvalues.

Discovering an eigenvector is like finding gold. When we find one eigenvector, we take it as a sign that a whole "eigenvector vein" is near, spurring us to the happy task of uncovering our treasure's full extent. Each eigenvector belongs to a *subspace* consisting entirely of eigenvectors, all with the same eigenvalue. To see why, first observe that if $\mathbf{v}$ is an eigenvector of A with eigenvalue λ, then so are all its nonzero scalar multiples. (Proof: For any scalar c, we have $A(c\mathbf{v}) = cA(\mathbf{v}) = c(\lambda\mathbf{v}) = (c\lambda)\mathbf{v} = (\lambda c)\mathbf{v} = \lambda(c\mathbf{v})$.*) Moreover, if $\mathbf{v}$ and $\mathbf{w}$ are any *two* eigenvectors with the same eigenvalue λ, then their sum will be a third such eigenvector. (Proof: $A(\mathbf{v}+\mathbf{w}) = A(\mathbf{v}) + A(\mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w} = \lambda(\mathbf{v}+\mathbf{w})$.) The set of eigenvectors with eigenvalue λ is thus closed under scalar multiplication and vector addition. Hence, we may conclude that this is a subspace, as claimed.† Indeed, it has a special name, which I'll bet you can guess:

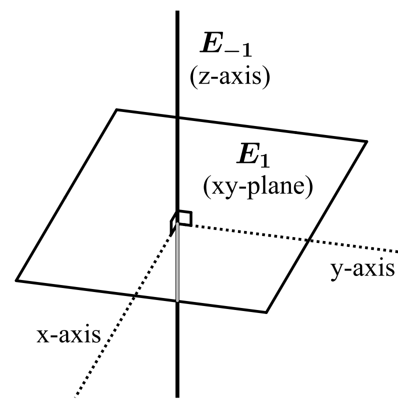
> **Definition**. Each eigenvalue λ of a map/matrix has a corresponding **eigenspace**, denoted E_λ, which is a subspace consisting of all eigenvectors with eigenvalue λ (and $\mathbf{0}$, too).

Example 1's matrix has two eigenvalues, so relative to it, $\mathbb{R}^2$ has two eigenspaces, depicted at left.



Example 2's map (90° rotation) lacks eigenvalues, so it has no associated eigenspaces. Example 3's reflection in the xy-plane has two eigenvalues, so relative to that reflection, $\mathbb{R}^3$ has two eigenspaces, depicted at right.

The figures also make it clear that we can find an eigenbasis relative to each of these two maps. For example, relative to the reflection, we'll get an eigenbasis for $\mathbb{R}^3$ by taking any two linearly independent vectors in E_1 plus any one nonzero vector in E_{-1}.



* The Footnote Pedant wishes to point out that the first equals sign in this chain is justified by a basic linearity property (Exercise 14B in Chapter 3); the second is justified because we were given that $\mathbf{v}$ is an eigenvector with eigenvalue λ; the third and fifth are justified by an associative property of scalar multiplication (mentioned in a footnote near the beginning of Chapter 1); the fourth is the commutativity of multiplication of real numbers.

† The Footnote Pedant is now standing athwart my last sentence, yelling STOP. "You said," he declares, "that the set is closed under scalar multiplication, but you've only proved that it is closed under multiplication by *nonzero* scalars! Aren't you sweeping something under the rug? Scaling any eigenvector by 0 turns it into $\mathbf{0}$, which *isn't* an eigenvector because we *defined* eigenvectors to be nonzero. Therefore, the set of eigenvectors with eigenvalue λ is *not* closed under scalar multiplication!" Strictly speaking, he's right. I could satisfy him by replacing the offending phrase "the set of all eigenvectors with eigenvalue λ" with this: "the set of all vectors $\mathbf{v}$ with the property that $A\mathbf{v} = \lambda\mathbf{v}$", which includes all the eigenvectors and *also* the zero vector, since $A(\mathbf{0}) = \lambda\mathbf{0}$, which holds, of course, because every linear map sends $\mathbf{0}$ to $\mathbf{0}$. I'll not do that, though – except down here in the footnotes – since that reformulation is much less memorable, and the discrepancy is minor.

A more interesting thinker than the Footnote Pedant would ask this: "Why did we define eigenvectors to be nonzero in the first place? Why don't we just drop that requirement?" It's a good question. The answer is that if we allowed $\mathbf{0}$ to count as an eigenvector, then *every* real number would be an eigenvalue of *every* matrix, since $A(\mathbf{0}) = \mathbf{0} = r\mathbf{0}$ for all real r. That won't do. Moreover, if $\mathbf{0}$ were considered an eigenvector, then it wouldn't have a particular eigenvalue the way every other eigenvector does; it would have infinitely many eigenvalues. This is a mess. Hence, we exclude $\mathbf{0}$ from the eigenvectors to preserve the statement "every eigenvector has a unique eigenvalue", much as we exclude 1 from the primes to preserve the statement "every number can be factored into a unique product of primes."

Exercises.

1. Is $\begin{pmatrix} 4 \\ -3 \\ 1 \end{pmatrix}$ an eigenvector of $\begin{pmatrix} 3 & 7 & 9 \\ -4 & -5 & 1 \\ 2 & 4 & 4 \end{pmatrix}$? If so, find the corresponding eigenvalue.

2. Describe the eigenvalues and eigenspaces of the following linear transformations of $\mathbb{R}^2$.

a) The zero map (which maps everything to the origin).
b) $180°$ rotation about the origin.
c) Orthogonal projection onto the line $y = 2x$.
d) Reflection across the y-axis.

3. Let A be an invertible $n \times n$ matrix. Suppose $\mathbf{v}$ is an eigenvector of A with eigenvalue λ. Determine whether $\mathbf{v}$ is also an eigenvector of each of the following matrices – and if so, with what eigenvalue. Explain your answers.

a) $2A$ b) nA c) A^2 d) A^n e) A^{-1} f) I g) $2A + 3I$.

4. "Eigenstuff" is defined only for *square* matrices. Explain why.

5. What are the only possible eigenvalues of an *orthogonal* matrix? Explain your answer.

6. (Extending the **Invertible Matrix Theorem**) In Exercise 19 of Chapter 5, you saw that statements A - M in the list below were equivalent statements about an $n \times n$ matrix A. Explain why we can add statement N to the list:

a) A is invertible. b) $\text{rref}(A) = I$. c) $A\mathbf{x} = \mathbf{b}$ has a unique solution for every vector $\mathbf{b}$.
d) A's columns are linearly independent. e) A's columns span $\mathbb{R}^n$. f) A's columns constitute a basis for $\mathbb{R}^n$.
g) $\ker(A) = \mathbf{0}$. h) $\text{im}(A) = \mathbb{R}^n$. i) $\text{rank}(A) = n$. j) $\det A \neq 0$.
k) A's *rows* are linearly independent. l) A's rows span $\mathbb{R}^n$. m) A's rows constitute a basis for $\mathbb{R}^n$.
n) A *doesn't* have 0 as an eigenvalue.

To reiterate, the moral of the invertible matrix theorem is that square matrices come in two types: *invertible* matrices (which satisfy all 14 of those conditions) and *noninvertible* matrices (which satisfy none of them).

7. To appreciate why expressing everything an eigenbasis makes computations so simple…

a) Suppose $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ is an eigenbasis $\mathcal{B}$ (with corresponding eigenvalues 3, 8, -2) relative to a linear map T on $\mathbb{R}^3$. Suppose we have a vector $\mathbf{v}$ whose $\mathcal{B}$-coordinates are $(4, -2, 5)$. What will the $\mathcal{B}$-coordinates of $T\mathbf{v}$ be?

b) Given the previous part's setup, if $\mathbf{x}$'s $\mathcal{B}$-coordinates are (x_1, x_2, x_3), what will the $\mathcal{B}$-coordinates of $T\mathbf{x}$ be?

b) More generally, suppose $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ is an eigenbasis $\mathcal{B}$ (with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$) relative to some linear map T. If $\mathbf{x}$ is a vector with $\mathcal{B}$-coordinates $(x_1, x_2, \ldots, x_n)$, then what will $T\mathbf{x}$'s $\mathcal{B}$-coordinates be?

8. As you now know, if $\mathcal{B}$ is an eigenbasis relative to a map T, then $[T]_\mathcal{B}$ is a *diagonal* matrix representation of T. This is important, so explain in your own words *why* this is true and what the diagonal entries of $[T]_\mathcal{B}$ will be.

9. (Eigendecomposition) We can sometimes decompose a given matrix into a product of several simpler factors. (Such a matrix decomposition is analogous to factoring a polynomial or factoring an ordinary whole number.) Of the various types of matrix decompositions, one of the most common and useful is called *eigendecomposition*. An eigendecomposition is like an X-ray of a matrix, exposing its normally hidden eigenstuff to plain view. As you'll learn in this exercise, a matrix can be eigendecomposed if and only if it admits an eigenbasis. Let's get to work.

We begin by noting that we can view any matrix A as the standard matrix of some linear map T (so that $A = [T]_\mathcal{E}$). Moreover, if this map has enough linearly independent eigenvectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ to form an eigenbasis $\mathcal{B}$, we know that $[T]_\mathcal{B}$ is a *diagonal* matrix Λ, whose diagonal entries are eigenvalues corresponding to the $\mathbf{b}_i$.* Now…

* The pointy symbol Λ is a Greek letter: *capital* lambda. We use it in this context to remind ourselves that this diagonal matrix's entries are eigenvalues, which, of course, we also represent by lambdas, albeit lower-case ones: λ.

a) Let $V = \begin{pmatrix} | & | & & | \\ \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \\ | & | & & | \end{pmatrix}$, the columns of which "store" the standard coordinates of A's eigenbasis vectors.

I claim that matrix V is invertible. Explain why this is so.

b) I claim that any matrix A that admits an eigenbasis can be "eigendecomposed" into a product of three matrices:

$$A = V\Lambda V^{-1},$$

where the diagonal matrix Λ stores the eigenvalues, and V stores an eigenvector corresponding to each one (in the corresponding column). Your problem: Justify my claim.

c) For instance, consider matrix A from Example 1. We found that it has an eigenbasis consisting of the vectors $\mathbf{b}_1 = \mathbf{i} - \mathbf{j}$ and $\mathbf{b}_2 = \mathbf{i}$ with corresponding eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 3$. In this case, what are the matrices V, Λ, and V^{-1}? Check your eigendecomposition by multiplying out $V\Lambda V^{-1}$ and verifying that the product is A.

d) **(Eigendecomposition helps us raise a matrix to a *power*)**
Eigendecomposition is useful when we want to apply a linear map to something, then apply the same linear map to the output, then apply the map to *that* output, and so forth. (Such iterative procedures are common in statistics, numerical analysis, machine learning, and even in Google's Page Rank algorithm.) The situation I've just described yields a computation of the form

$$A\Big(A\Big(A\big(A\cdots\big(A(\mathbf{v})\big)\cdots\big)\Big)\Big),$$

and since function composition corresponds to matrix multiplication, this reduces to $AAAA\cdots A(\mathbf{v})$, or more compactly, to $A^n(\mathbf{v})$. Alas, raising a matrix to a high power, especially if the matrix is large (as they so often are in applications) is computationally "expensive", even for a very fast computer.

For example, suppose A is a 100×100 matrix, and we wish to raise it to some high power. How many computations must a computer do to carry such a computation out by brute force? Well, just to find $A^2 = AA$ involves doing 10,000 dot products (one for each entry in the matrix product), and each dot product involves 100 products (one for each component) and 99 sums. That works out to almost 2 million arithmetic operations to multiply A by itself just *once*. Grinding out A^n by brute force would require $(n-1)$ of these computationally expensive matrix multiplications. Clearly, we'll want to minimize the number of such matrix multiplications that our computer will have to do. A way to reduce expenses is to find an eigendecomposition of A (if it has one) and raise *that* to the n^{th} power, because if $A = V\Lambda V^{-1}$, it follows that

$$A^n = (V\Lambda V^{-1})^n = \underbrace{(V\Lambda V^{-1})(V\Lambda V^{-1})(V\Lambda V^{-1})\cdots(V\Lambda V^{-1})}_{n \text{ of these trios}}.$$

The associativity of matrix multiplication lets us regroup those parentheses, pairing each trio's concluding V^{-1} with the V from the trio that follows. Such pairs cancel each other out, leaving us with this:

$$A^n = V\Lambda^n V^{-1}.$$

Your problem: Explain why this expression for A^n is much less computationally expensive (when n is large) than the brute force approach to A^n described above. [*Hint*: Raising a *diagonal* matrix to the n^{th} power is easy. Recall Exercise 30e from Chapter 4.]

e) Given matrix A from Example 1, compute A^{10} by hand using the eigendecomposition of A you found in Part C. Then try computing A by brute force, too. (It's ok if you eventually give up. The point is to *feel* the difference between the two approaches so that eigendecomposition's superiority will be palpable to you.)

10. (A fun curiosity) If a square matrix has the property that the entries on each row add up to the same number s, then s is an eigenvalue of the matrix. Explain why this is so.

Finding the Eigenstuff

Mary had a little λ, little λ, little λ…

- Mrs. Traditional

Now that you know what a matrix's eigenthings are, it's time to discuss how we *find* a matrix's eigenthings. Since each eigenvector $\mathbf{v}$ is scaled by some particular eigenvalue λ, we'll want to seek eigenthings in pairs. Namely, given any square matrix A, we want to find all pairs of *nonzero vectors* $\mathbf{v}$ and *scalars* λ such that

$$A\mathbf{v} = \lambda\mathbf{v}.$$

This is the second most important equation in all linear algebra (behind only the ubiquitous $\boldsymbol{A}\mathbf{x} = \mathbf{b}$). Let's rewrite it in an equivalent form where $\mathbf{v}$ appears only once:

$$A\mathbf{v} = \lambda\mathbf{v} \quad \Leftrightarrow \quad A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0} \quad \Leftrightarrow \quad A\mathbf{v} - (\lambda I)\mathbf{v} = \mathbf{0} \quad \Leftrightarrow \quad (A - \lambda I)\mathbf{v} = \mathbf{0}.^*$$

It follows from this last reformulation that a given nonzero vector $\mathbf{v}$ and a given scalar λ will be one of A's eigenvector/eigenvalue pairs if and only if $(A - \lambda I)$ maps the given *nonzero* vector $\mathbf{v}$ to $\mathbf{0}$.

This reformulation of the original equation may seem odd, but it delivers crucial geometric insights: First, recall that a *nonzero* vector $\mathbf{v}$ can get mapped to $\mathbf{0}$ by a matrix (here, $A - \lambda I$) only when the matrix *collapses at least one dimension* of the space on which it acts. (Otherwise, the matrix would be invertible, in which case it would only map $\mathbf{0}$ to $\mathbf{0}$.) Second, recall that a matrix induces a dimensional collapse precisely when *the matrix's determinant is zero*.

Putting this all together, we see that our original equation for eigenvector/eigenvalue pairs will be satisfied by a *nonzero* vector $\mathbf{v}$ (and its scalar mate λ) precisely when

$$\det(A - \lambda I) = 0.$$

Notably, this equation doesn't explicitly refer to the eigen*vector* half $\mathbf{v}$ of the pairs that we are seeking. It's actually a good thing that $\mathbf{v}$ is hiding, since this temporarily narrows our focus to a single unknown, λ. If we can solve the equation $\det(A - \lambda I) = 0$ for our unknown λ, we'll have all the eigen*values* of A. Once we have them, they'll confess – after some algebraic coaxing – the locations of the eigen*spaces* where their eigen*vector* mates live. We can then see if those eigenspaces contain enough linearly independent eigenvectors to build an eigenbasis, and if so, we can use it to carry out an eigendecomposition of A.

But before we go that far, let's begin with a quick example that concentrates on eigenvalues alone. That way, we'll see how the first – and most crucial – piece of the puzzle plays out in a concrete instance.

Example 1. Find the eigenvalues of $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$.

Solution. As discussed above, A's eigenvalues are the solutions of $\det(A - \lambda I) = 0$.
Working out that determinant reveals that the left side of the equation is just a polynomial in λ:

$$\det(A - \lambda I) = \det\left(\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) = \det\begin{pmatrix} 1-\lambda & 2 \\ 4 & 3-\lambda \end{pmatrix} = \lambda^2 - 4\lambda - 5.$$

The roots of that polynomial, and hence A's eigenvalues, are $\mathbf{5}$ and $-\mathbf{1}$. ♦

---

* The only potentially mysterious step here is rewriting $\lambda\mathbf{v}$ as $(\lambda I)\mathbf{v}$**.** You should convince yourself that this substitution is valid (i.e. that matrix λI's effect on $\mathbf{v}$ is the same as merely scaling $\mathbf{v}$ by λ). We took that step so that both terms in the expression would be *matrix*-vector multiplications, from which we could then factor out the common vector $\mathbf{v}$.

If A is an $n \times n$ matrix, $\det(A - \lambda I)$ always turns out to be a n^{th}-degree polynomial. A natural name for this eigenvalue-laden polynomial would be A's "eigenpolynomial", but alas, it's called something else:*

Definition. If A is any square matrix, $\det(A - \lambda I)$ is called A's **characteristic polynomial**.

To repeat, A's characteristic polynomial is important because its roots are A's eigenvalues. And once we have its eigen*values*, finding the other eigenstuff is as simple as following your nose, as we'll now see.

Example 2. (Continuing Example 1.) Find the eigenspaces corresponding to the matrix

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

Then, if A admits an eigenbasis, find one, and use it to create an eigendecomposition of A.†

Solution. In Example 1, we already found A's eigenvalues: 5 and -1.

By definition, the eigenspace E_5 consists of all vectors $\mathbf{v}$ such that $A\mathbf{v} = 5\mathbf{v}$.
Using an algebraic trick from the previous page, we'll rewrite this equation in an equivalent form:

$$(A - 5I)\mathbf{v} = \mathbf{0}.$$

Now we're in familiar territory. We'll just rewrite this last equation as an augmented matrix,

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

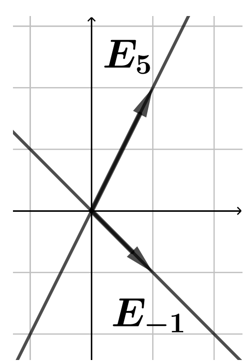
and solve the system. Its solutions, as you should verify, are all the points on the line $y = 2x$. Thus, E_5 is the line $y = 2x$.

The eigenspace E_{-1} consists of all vectors $\mathbf{v}$ such that $A\mathbf{v} = -\mathbf{v}$. Or equivalently, all vectors $\mathbf{v}$ such that $(A + I)\mathbf{v} = \mathbf{0}$. We can find them by solving

$$\left(\begin{array}{cc|c} 2 & 2 & 0 \\ 4 & 4 & 0 \end{array}\right).$$

When we do so, we find, as you should verify, that every point on the line $y = -x$ is a solution. Thus, E_{-1} is the line $y = -x$.

Any two linearly independent eigenvectors of A will form an eigenbasis, so we can clearly get one by taking one eigenvector from each eigenspace – such as these ones (expressed standard coordinates):



$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ with eigenvalue 5 and $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ with eigenvalue -1.

It follows from Exercise 9B that A's corresponding eigendecomposition is

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

as you can verify by multiplying out the right side. ♦

* Well, sort of. "Characteristic" is a translation of *eigen*, so it might as well be "eigenpolynomial". But perhaps this word, cobbled together from German (*eigen*), Greek (*poly*), and Latin (*nomen*), was deemed too outré for use by some dreary linguistic purist.

† The notion of an *eigendecomposition* was introduced in Exercise 9.

If you watch someone (your professor, a friend, or someone online) computing a matrix's eigenspaces, it can look like a mysterious algorithm, especially if he or she compresses some of the preliminary steps. Indeed, I've seen some students learn to carry out the process in a mindless algorithmic manner, which runs something like this: "If one of A's eigenvalues is 42 (or whatever), then to find E_{42}, we subtract 42 from each diagonal entry of A, then we turn the result into an augmented matrix by appending a column of zeros, and finally we solve that system. The solutions are the eigenvectors that make up E_4." That recipe will indeed produce the desired eigenspace, and there's nothing wrong with using it – provided you understand *why* it works. If you use it without understanding why it works, you aren't learning linear algebra; you are just following orders. On the other hand, if you understand why it works, there's no reason to memorize that sequence of steps in the first place. After all, it takes only a few seconds to reason through the entire process, justifying every step: "By definition, E_{42} consists of all vectors that satisfy $A\mathbf{v} = 42\mathbf{v}$. To find these vectors, we'll rewrite that equation as $(A - 42I)\mathbf{v} = \mathbf{0}$, since we can easily turn this into an augmented matrix whose underlying system we can then solve with Gaussian elimination." I encourage you to think through the process this way rather than memorize an algorithm.

Once you have the eigenspaces of an $n \times n$ matrix, it's easy to see whether the underlying map is diagonalizable: If the eigenspaces' dimensions add up to n, we can gather up enough linearly independent eigenvectors to form an eigenbasis, which we can then use to represent the map as a diagonal matrix. But if sum of the eigenspaces' dimensions falls short of n, we're out of luck: The map doesn't admit an eigenbasis, so we can't represent it as a diagonal matrix. (Incidentally, matrices that can't be diagonalized are sometimes called **defective** matrices.)

Let's do one another example of finding eigenstuff.

Example 3. Find the eigenspaces of

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

Solution. A's characteristic polynomial is

$$\det(A - \lambda I) = \det\begin{pmatrix} 1-\lambda & 0 & 2 \\ 0 & 3-\lambda & 0 \\ 2 & 0 & 1-\lambda \end{pmatrix}.$$

Doing cofactor expansion on the second row and massaging the resulting expression reveals the characteristic polynomial to be $(-1)(3-\lambda)^2(\lambda+1)$. Thus, A's eigenvalues are $\mathbf{3}$ and $-\mathbf{1}$.

The eigenspace E_3 consists of all vectors $\mathbf{v}$ such that $A\mathbf{v} = 3\mathbf{v}$. We can rewrite this equation in the equivalent form $(A - 3I)\mathbf{v} = \mathbf{0}$, so we can solve it by row-reducing an augmented matrix:

$$\left(\begin{array}{ccc|c} -2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & -2 & 0 \end{array}\right) \begin{matrix} \\ \\ +R_1 \end{matrix} \left(\begin{array}{ccc|c} -2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \begin{matrix} \div(-2) \\ \\ \\ \end{matrix} \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

From this, we see that $x = z$, while there are no constraints on y or z, so these are free variables. Hence, solutions have the form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ s \\ t \end{pmatrix} = t\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + s\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

That is, the eigenspace E_3 is a plane in $\mathbb{R}^3$: the span of the two independent vectors $\mathbf{i} + \mathbf{k}$ and $\mathbf{j}$.

As for the eigenspace E_{-1}, this consists of all vectors $\mathbf{v}$ such that $A\mathbf{v} = -\mathbf{v}$. We'll rewrite this in the form $(A + I)\mathbf{v} = \mathbf{0}$, and then solve it by row-reducing an augmented matrix:

$$\left(\begin{array}{ccc|c} 2 & 0 & 2 & 0 \\ 0 & 4 & 0 & 0 \\ 2 & 0 & 2 & 0 \end{array}\right) \begin{array}{c} \div 2 \\ \div 4 \\ \div 2 \end{array} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right) \begin{array}{c} \\ \\ -R_1 \end{array} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

From this we see that $z = -x$, and that $y = 0$. Hence, the solutions have the form

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ 0 \\ -t \end{pmatrix} = t\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

That is, the eigenspace E_{-1} is a line in $\mathbb{R}^3$: the span of the vector $\mathbf{i} - \mathbf{k}$.

To sum up, A's eigenvalues are 3 and -1, and the corresponding eigenspaces are:

$$E_3\text{: span of} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \text{and} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}; \qquad E_{-1}\text{: span of} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}. \qquad \blacklozenge$$

Although I didn't ask for it in the previous example, you should be able to see that A admits an eigenbasis: any two linearly independent vectors from E_3 plus any nonzero vector from E_{-1} will do. For example, vectors $\mathbf{i} + \mathbf{k}$, $\mathbf{j}$, and $\mathbf{i} - \mathbf{k}$ make up one eigenbasis. With respect to that basis, the linear map whose standard matrix is A will have a diagonal representation, and its diagonal entries will be 3, 3, and -1. You should be able to write down the related eigendecomposition of A, too.

Once we've found a matrix's eigen*values*, the rest of its eigenstuff comes easily. But there's a catch: Finding the eigenvalues usually isn't easy. Well… finding the eigenvalues of a 2×2 matrix is easy, since 2×2 matrices have *quadratic* characteristic polynomials, whose roots we can always find with the quadratic formula. But alas, 3×3 and 4×4 matrices have *cubic* and *quartic* characteristic polynomials. Cubic and quartic analogs of the quadratic formula do exist, but they are so appallingly complicated that no one in his or her right mind knows them.* And for larger matrices, the situation is considerably worse: No algebraic formula for the roots of 5th (or higher) degree polynomials exists.† So how, in applications, do we find large matrices' eigenvalues? We approximate them. Laborers in the numerical linear algebra mines have worked out *eigenvalue algorithms* to approximate matrices' eigenvalues as closely as we like. These sophisticated iterative algorithms – which are built into computer programs that scientists and engineers use every day – typically *don't* involve the characteristic polynomial. Yet even for large matrices, the characteristic polynomial remains a vital theoretical touchstone. For example, in Exercise 13, you'll use it to prove an important fact about the eigenvalues of triangular matrices – of any size.

---

* The cubic and quartic formulas are *historically* important. Their discovery in 16th-century Italy was the first original achievement of European mathematics since the ancient Greeks and initiated the *mathematical* Renaissance. Remarkably, their discovery involved the first real use of complex numbers in mathematics – and this was at a time when even negative numbers, much less their square roots, were considered hopelessly absurd fictions. The story of the cubic, which includes mathematical duels, flashes of insight, vows of secrecy made and broken, and a colorful cast of characters (above all Niccolò Tartaglia and Gerolamo Cardano), is engagingly told by Paul Nahin in the first chapter of his semi-popular book on complex numbers, *An Imaginary Tale.*

† To be clear, I do not mean that no one has found such a formula yet. I mean that there *isn't* one and there never will be one. This was proved in the 19th century in yet another fascinating episode of mathematical history – one that ultimately gave birth to large portions of what we now call *abstract algebra* – group theory and Galois theory in particular.

Exercises.

11. a) Remind yourself *why* a matrix A's eigenvalues are the zeros of its characteristic polynomial, $\det(A - \lambda I)$.
b) If, say, 8 is an eigenvalue of matrix A, explain how to find the eigenspace E_8, and why your method works.

12. For each of the following, find the eigenvalues, describe the corresponding eigenspaces, and determine whether the matrix admits an eigenbasis. If so, state an eigenbasis, and give the corresponding eigendecomposition.

a) $\begin{pmatrix} 3 & 2 \\ 3 & 8 \end{pmatrix}$ b) $\begin{pmatrix} -2 & 6 \\ -2 & 5 \end{pmatrix}$ c) $\begin{pmatrix} 2 & 3 \\ -4 & -2 \end{pmatrix}$ d) $\begin{pmatrix} 2 & 2 \\ -8 & -6 \end{pmatrix}$

e) $\begin{pmatrix} 5 & 2 & 3 \\ 0 & 5 & 4 \\ 0 & 0 & 1 \end{pmatrix}$ f) $\begin{pmatrix} 3 & -2 & 5 \\ 1 & 0 & 7 \\ 0 & 0 & 2 \end{pmatrix}$ g) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

13. (Triangular matrices' eigenvalues) In Part E of the previous problem, you saw that the eigenvalues of a *triangular* matrix turned out to be the entries on its main diagonal. Was that a coincidence? It was not! Prove the following:

Theorem. If M is a *triangular* matrix, then M's eigenvalues are the entries on its main diagonal.

This gives us another reason to like triangular matrices: We can see their eigenvalues at a glance.

14. (The transpose's eigenvalues)
a) Justify each equals sign in the following chain, which pertains to every square matrix A:

$$\det(A - \lambda I) = \det((A - \lambda I)^T) = \det(A^T - (\lambda I)^T) = \det(A^T - \lambda I).$$

You've just shown that every square matrix and its transpose share the same characteristic polynomial.

b) Can we conclude that A and A^T will have the same eigenvalues? What about their eigenvectors?

c) Suppose that a square matrix A has the property the entries on each *column* sum up to the same number s.*
What, if anything, can we conclude about A's "eigenstuff"? (Compare Exercise 10.)

15. *Invertibility and diagonalizability are completely unrelated concepts*.
To reinforce this fact, I've presented four matrices below that cover all possible combinations of invertibility and diagonalizability. Verify this, and fill in the table below.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

	Invertible?	Diagonalizable?
A		
B		
C		
D		

16. You've seen 2×2 matrices with no real eigenvalues, such as the matrix in 12C above, or the matrix that effects a $90°$ rotation of $\mathbb{R}^2$ about the origin. However, *every* 3×3 matrix has at least one real eigenvalue. Explain why.†

* Such matrices arise naturally in probability calculations, for example as *transition matrices* (also known as *stochastic matrices*) in Markov chains. The entries in each column of such a matrix represent probabilities that must add up to 1.

† There are such things as *complex* eigenvalues, but not in this introductory textbook.

17. (Algebraic and Geometric Multiplicities of Eigenvalues)

You might or might not be familiar with "the Fundamental Theorem of Algebra" (FTA), which states that every n^{th}-degree polynomial $p(x)$ can be factored as follows:

$$p(x) = c(x - r_1)(x - r_2)\cdots(x - r_n),$$

where c is a constant, and where some or all of the constants r_i may be complex numbers. A few examples:

$$x^2 - 6x + 9 = (x - 3)(x - 3)$$
$$2x^2 + 3x - 2 = 2(x - 1/2)(x + 2)$$
$$x^3 + x = (x - 0)(x - i)(x + i)$$
$$x^4 - 2x^3 - 7x^2 + 20x - 12 = (x - 2)^2(x - 1)(x + 3).$$

The FTA can't tell you how to find a factorization – only that a factorization *exists* (in the mind of God, as it were). The FTA is a vital theorem for theoretical work; we'll use it below - and again in exercise 18. The constants r_i that appear in the n^{th}-degree polynomial's factorization are its n roots. If the same root appears k times in this factorization, we say that root has **multiplicity** k. (When $k > 1$, we call the root a **repeated root**.)

a) State the roots – and the multiplicity of each repeated root – of each of the four polynomials above.

Now for some definitions:
The ***algebraic*** **multiplicity** of an eigenvalue of A is its multiplicity as a root of A's characteristic polynomial.
The ***geometric*** **multiplicity** of an eigenvalue of A is the dimension of its associated eigenspace.

b) Exercise 12F concerned a matrix whose characteristic polynomial was $(-1)(\lambda - 2)^2(\lambda - 1)$. Its eigenvalues were thus 2 and 1, and you found that the respective eigenspaces were both *lines* in $\mathbb{R}^3$.
That being so, state the algebraic and geometric multiplicities of each eigenvalue of that matrix.

One can show that *each eigenvalue's* ***geometric*** *multiplicity is less than or equal to its algebraic multiplicity*. The proof is somewhat involved, and would involve a serious detour, so we'll take it for granted in this exercise. With that in mind, explain why each of the following statements about an $n \times n$ matrix A's *real* eigenvalues hold.

c) The sum of the algebraic multiplicities is at most n.
d) The sum of the geometric multiplicities is at most n.
e) The matrix is diagonalizable ⇔ the geometric multiplicities add up to n.
f) If the algebraic multiplicities add up to n, then the matrix might - or might not - be diagonalizable.
g) The matrix is diagonalizable ⇔ (1) the algebraic multiplicities add up to n *and* (2) each eigenvalue's geometric multiplicity equals its algebraic multiplicity.
h) If the matrix has even one eigenvalue whose geometric multiplicity is strictly less than its algebraic multiplicity, then the matrix is defective (i.e. it *can't* be diagonalized).
i) If an eigenvalue's algebraic multiplicity is 1, the eigenvalue's corresponding eigenspace is a *line*.

18. (The Trace of a Matrix)

If A is a square matrix, then its **trace** (which we denote $\text{tr}(A)$) is the sum of the entries on A's main diagonal. Surprisingly, the trace turns up in some interesting places, a few of which you'll meet in this problem.

a) Prove the following: If A is a 2×2 matrix, then A's characteristic polynomial is

$$\lambda^2 - \big(\text{tr}(A)\big)\lambda + \det(A).$$

b) Use the preceding result to find the characteristic polynomials of the 2×2 matrices in Exercise 12 more quickly than you found them originally.

c) Prove the following: If A is a 2×2 matrix with eigenvalues λ_1 and λ_2 (which may be equal or complex!), then

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 \quad \text{and} \quad \det(A) = \lambda_1\lambda_2.$$

That is, the eigenvalues' sum is the trace; the eigenvalues' product is the determinant.
[*Hint*: *Combine the result of Part A with the Fundamental Theorem of Algebra from Exercise 17*.]

d) Use Part C to find the following matrices' eigenvalues *without* finding their characteristic polynomials:

$$A = \begin{pmatrix} 3 & 4 \\ 2 & 5 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}.$$

e) In fact, the results from Part C extend to square matrices of all sizes, not just 2×2 ones. That is, the sum and product of every square matrix's eigenvalues turn out to be its trace and determinant respectively.* This result, perhaps combined with others, can be used to impress your friends at parties by finding a matrix's eigenvalues without having to think about its characteristic polynomial. For example, what are the eigenvalues of

$$M = \begin{pmatrix} 2 & 4 & 3 \\ 1 & 2 & 6 \\ 3 & 6 & 0 \end{pmatrix}?$$

Well, all the rows add up to the same constant, so that constant, **9**, must be an eigenvalue (by Exercise 10). Next, observe that the columns are linearly *de*pendent, since column two is double column one. Hence, by the Invertible Matrix Theorem (see Exercise 6), this matrix is *not* invertible, and accordingly, it must have **0** as one of its eigenvalues. What's the third eigenvalue? Well, $\operatorname{tr}(A) = 4$, so the three eigenvalues must sum to 4, which means that the third eigenvalue must be **-5**. Hence, M's eigenvalues are 9, 0, and -5.

Use such "party tricks" to find the eigenvalues of these matrices without computing their characteristic polynomials, carefully justifying each of your steps:

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 2 & 2 & 4 \\ 3 & 1 & 4 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 5 & 1 & 3 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 4 & 2 & 0 \end{pmatrix}.$$

* But note that in this eigenvalue sum, one must include each eigenvalue a number of times equal to its own algebraic multiplicity. Thus, if A's characteristic polynomial is $(\lambda - 5)^2(\lambda + 2)(\lambda - 1)^3$, then its trace would be $5 + 5 + (-2) + 1 + 1 + 1 = 11$, and its determinant would be $(5)^2(-2)(1)^3 = -50$.

Eigenstuff and Long Run Behavior

In the long run we are all dead.

- John Maynard Keynes, *A Tract on Monetary Reform*

One way to grasp a mathematical system's long run behavior is to express the system in linear algebraic terms, and then analyze its eigenstuff, which often turns out to be related to *limits*, your old friends from calculus. To explain the idea, I'll dwell for a bit on the famous **Fibonacci sequence**: $0, 1, 1, 2, 3, 5, 8, 13$ ..., a numerical system that unfolds from a simple seed (the two initial terms, 0 and 1) plus a simple rule: Each successive term is the sum of the previous two terms.

Given *two* consecutive terms in the Fibonacci sequence, we produce the next term by adding them. But given just *one* number from the sequence (say, 832040), is there some mathematical way to find, or even approximate, its successor without having to reconstruct the entire sequence up to that point?

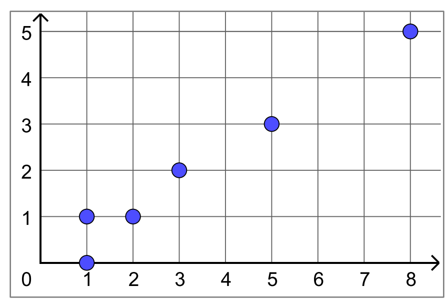
There is. The trick involves grasping the sequence in algebraic, and ultimately *linear* algebraic terms. If we let F_n be the Fibonacci sequence's n^{th} term, we can specify the full sequence recursively as follows:

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

Next, to inject some geometry, we will form vectors in $\mathbb{R}^2$ whose top component is a Fibonacci number (i.e. any number from the sequence), and whose bottom component is the *previous* Fibonacci number:

$$\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix}.$$

At right, I've plotted points corresponding to the first six of these "Fibonacci vectors". For example, the point at $(5, 3)$ tells us that 5 is a Fibonacci number… whose predecessor in the sequence is 3. The *second* component of a Fibonacci vector might initially seem pointless, but tucking it into a vector along with the first is clever, opening the door not only to geometry, but to a matrix as well. The key observation that leads us to it is that a certain *linear map* will take us from any one Fibonacci vector/point,



$$\begin{pmatrix} F_{n-1} \\ F_{n-2} \end{pmatrix}, \text{ to the next one, } \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix},$$

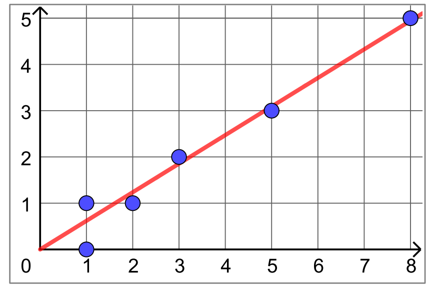
since the latter's components are linear combinations of the former's components:

$$\begin{aligned} F_n &= 1F_{n-1} + 1F_{n-2} \\ F_{n-1} &= 1F_{n-1} + 0F_{n-2}. \end{aligned}$$

We can bundle these linear relationships into a single matrix-vector equation:

$$\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{n-1} \\ F_{n-2} \end{pmatrix}.$$

Thus, we've found a 2×2 matrix that sends each Fibonacci vector/point to its *successor* on the graph. Accordingly, we can now see the entire graph as generated from the initial point $(1, 0)$ and the matrix: We feed $(1,0)$ into the matrix, which maps it to $(2, 1)$. We feed *that* point back into the matrix, which maps it to $(3, 2)$… and so on forever. The points generated this way are the Fibonacci points.



Observe that the graph's points come close to falling on a line. Or rather, the line is the limit towards which successive points tend. In the long run, points in the sequence are indistinguishable from points that lie *on* the line. Once we're far enough into the sequence, the matrix is essentially just *scaling* Fibonacci vectors, stretching them along the line without rotating them. In other words, in the long run, the Fibonacci vectors are eigenvectors, and the line is an eigenspace of the matrix. Its associated eigen*value* is, of course, the scaling factor by which the matrix, in the long run, stretches each Fibonacci vector.

Computing the matrix's eigenvalues is easy. You should do that now, using Exercise 18a's shortcut. You'll find that it has only one eigenvalue greater than 1. That's the relevant eigenvalue here, since the Fibonacci vectors are stretched *away* from the origin. This eigenvalue turns out to be the **golden ratio**,

$$\phi = \frac{1+\sqrt{5}}{2} \approx 1.618,$$

an irrational number famous for turning up in all sorts of mathematical places. You'll have the chance to explore some of ϕ's properties in the next exercise set. At any rate, we've established that if $\mathbf{f}_n$ is the n^{th} Fibonacci vector, then $\mathbf{f}_{n+1} \approx \phi\mathbf{f}_n$, with the approximation improving as n gets larger. Plucking the first components from each side of this approximation yields

$$F_{n+1} \approx \phi F_n,$$

and so, to return to our original question, if someone hands us a large Fibonacci number such as 832040 and asks us what the next one in the sequence will be, we need not reconstruct the whole sequence up to that point to find it. We can just multiply 832040 by ϕ. The product, rounded to the nearest integer, turns out to be 1346269. This integer should be close to the next term in the Fibonacci sequence. Remarkably, it turns out to be the next term exactly!* In fact, when we round ϕF_n to the nearest integer, we obtain F_{n+1} exactly for all $n \geq 2$. In other words, this "long run" tendency for the Fibonacci sequence takes hold almost immediately. This happens because the initial Fibonacci point $(1, 0)$ already lies so close to the magic line. This line, the eigenspace E_ϕ, attracts vectors like a magnet as we run Fibonacci vectors (and the resulting sequence of successive outputs) through the matrix.

Interestingly, the line's magnetic quality persists even if we initially feed the matrix a vector that *isn't* a Fibonacci vector. For example, the point/vector $(7, 1)$ isn't a Fibonacci vector, but if we feed it to the matrix, then feed its output back in, and so on, the resulting sequence of points tends towards the line, even though we're no longer working with Fibonacci numbers. These points' first components constitute a new sequence of numbers, still governed by Fibonacci's rule ("add the previous two to get the next"), but developing from a different seed:

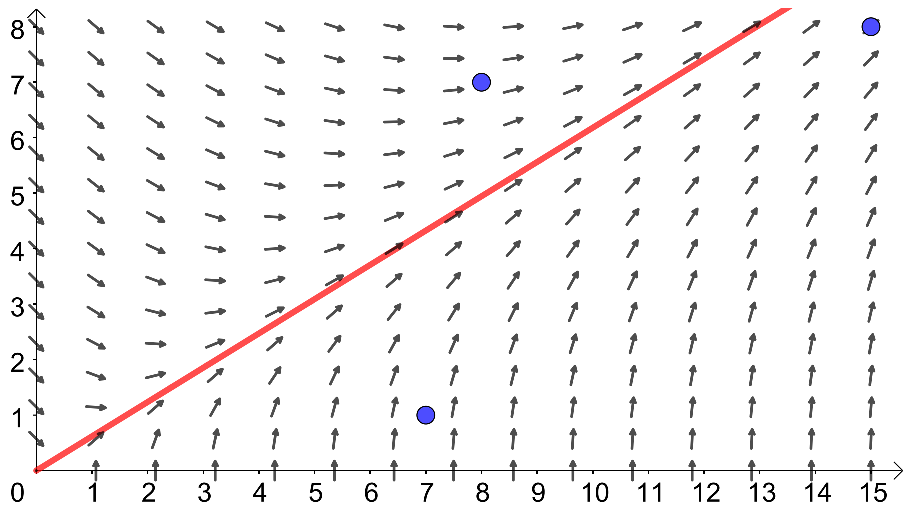
$$1, 7, 8, 15, 23, 38, 61, 99, \dots .$$

In the long run, this new sequence still grows at each step by a factor of ϕ. But since our initial point $(7, 1)$ now sits considerably further from the "magnetic" line (the eigenspace corresponding to ϕ), it takes a few more steps before the scaling factor of ϕ nails down the next term flawlessly. For example, 23 is the new

* Either patiently (by generating the sequence) or impatiently (by looking it up), you can confirm that 832040 and 1346269 are, respectively, F_{30} and F_{31}.

sequence's fifth term, and $23\phi \approx 37$, which is close… but not quite right. The sixth term is actually 38. However, by the next term, all's well: $38\phi \approx 61$, which is indeed the sequence's seventh term.

It helps to think of the matrix as generating a force field that pervades the plane, capable of pushing points (i.e. vectors) towards the eigenspace. The figure below gives the idea:*



Initially, we "dropped" a point into the field at $(7, 1)$, where it was subjected to forces pushing it north and very slightly east (the arrows in the picture give only the *direction* in which the "wind" blows at each point; they don't indicate its strength). This brought the point to $(8, 7)$, where the wind blows easterly, with a slight push to the north. These forces then pushed the point to $(15, 8)$, quite close to the line. Here, the wind blows nearly parallel to the line, but it still pushes points a bit closer to it. The overall effect, as the figure makes clear, is that no matter where we initially drop our point, the "winds" will ultimately push the point towards the line.

"But," you might reasonably ask, "What about the matrix's *other* eigenvalue,

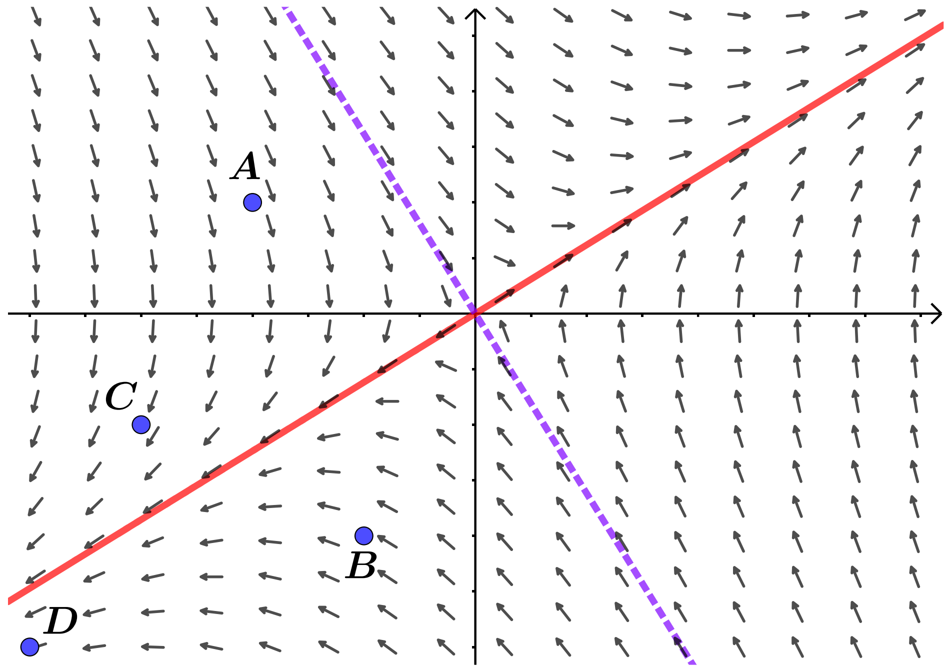
$$\frac{1-\sqrt{5}}{2} \approx 0.618,$$

and *its* eigenspace? Don't they have any effect on the geometric picture above?" They do indeed, but they will be visible only if we widen our field of view to encompass the entire plane – not just its first quadrant. When we do so, our picture will contain a pair of one-dimensional eigenspaces, like so:

* A computer generates the figure by taking many points (x, y) in the plane, and at each such point, computing the quantity

$$\begin{pmatrix}1 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} - \begin{pmatrix}x\\y\end{pmatrix},$$

which is the vector pointing *from* (x, y) to the point where the matrix sends (x, y). The computer then scales this direction vector down dramatically (to avoid traffic jams in the picture) and places it at (x, y).



The dotted eigenspace's eigenvalue is approximately 0.618. Since this is less than 1, points on it are drawn *towards* the origin, as the "force field" indicates. Even in this expanded view, it's easy to see that a point dropped anywhere in the plane will eventually find itself moving towards the first eigenspace, E_ϕ. For example, if we drop a point at $(-4, 2)$, it first gets mapped southeast to B $(-2, -4)$, then pushed northwest to C $(-6, -2)$, and then to D $(-8, -6)$. From there, the point is clearly locked forevermore into the stream that flows along E_ϕ, albeit in its "negative direction", flowing into the 3rd quadrant.

The ideas above have applications far beyond the Fibonacci sequence. We are now trespassing in the vast domain of *dynamical systems*, which you can explore in an introductory differential equations course – should you be lucky enough to take one employing a dynamical systems point of view. Still, even without the machinery of differential equations, you can sample the topic's flavor by considering the following example of a *discrete* dynamical system.

Living in the remote depths of a certain forest are some wild dachshunds and their preferred prey, feral mailmen. If D_n and M_n represent the population of each species in this region at the beginning of year n, then a simple **predator-prey model** of how these populations change in time might look like this:

$$\begin{aligned} D_{n+1} &= .86D_n + .08M_n \\ M_{n+1} &= -.12D_n + 1.14M_n. \end{aligned}$$

To see why this might make sense, consider each equation in turn. The first indicates that the number of dachshunds *next* year depend on how many dachshunds and mailmen there are this year. If for, example, there are no mailmen in the region this year, then the dachshund population will, according to this model, be reduced by 14% as they are forced to endure a year without their favorite food. On the other hand, for every 100 mailmen in the area this year, there will be 8 new dachshunds next year – a perfectly reasonable assumption, since mailman meat nourishes the local dachshunds and attracts more dachshunds from afar.

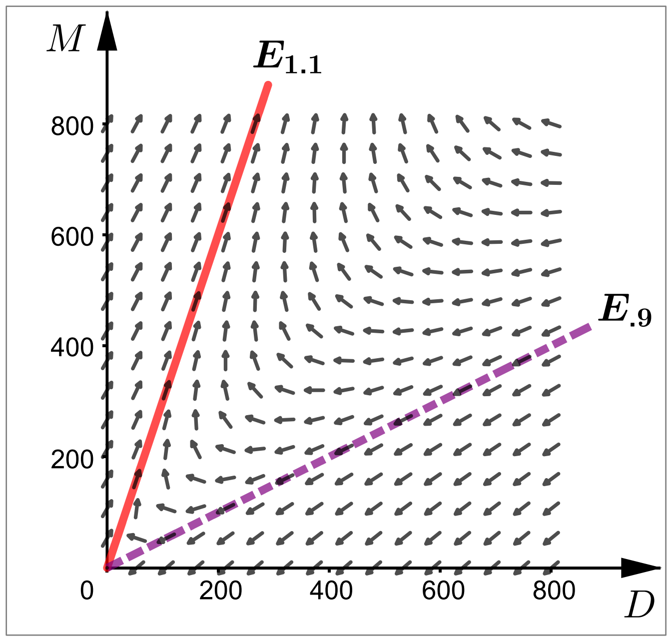
As for the model's second equation, it tells us that in the absence of dachshunds, the unchecked mailman population would grow by 14% each year.* However, for every hundred dachshunds, twelve mailmen will succumb to their ruthless predation. Incidentally, some have speculated that dachshund-on-mailman violence was precisely what inspired Alfred Lord Tennyson's phrase "Nature, red in tooth and claw".

Naturally, the model above can be translated into a matrix-vector equation:

$$\begin{pmatrix} D_{n+1} \\ M_{n+1} \end{pmatrix} = \begin{pmatrix} .86 & 0.08 \\ -.12 & 1.14 \end{pmatrix} \begin{pmatrix} D_n \\ M_n \end{pmatrix},$$

and an "eigenanalysis" like the one we employed in our Fibonacci problem will reveal much about this system's long run behavior. Carrying this out, we find that the matrix has two eigenvalues: 1.1 and 0.9. Their corresponding eigenspaces are shown in the figure below.

The figure reveals how the fate of this predator-prey system will depend on its initial conditions. For example, if we start with 800 dachshunds and 200 mailmen (i.e. if we drop a point into the "force field" at $(800, 200)$), the arrows point out a dire fate for the mailmen: They'll be hunted to extinction, driven down to the D-axis of doom. However, had those 800 dachshunds been paired initially with 600 mailmen, our moving point would be drawn to a happier fate: eigenspace $E_{1.1}$, where a stable ratio of 3 mailmen for every 1 dachshund eventually holds sway, with both populations then growing by 10% in each subsequent year. Of course, that growth rate can't be sustained forever in practice, which is one weakness of this crude predator-prey model.† More refined models exist, but this example was only meant to convey the basic idea.



At any rate, you'll have the chance to dig more deeply into these ideas in future classes. My intention here was just to gesture in their general direction, pointing towards the mathematical horizon – towards an educational eigenspace that you may find yourself drawn to bit by bit should your own initial conditions happen to predispose you to becoming sucked into the world of dynamical systems.

* Astute readers may now be wondering if these feral mailmen have evolved to reproduce asexually. Not so. Although I've used the gendered term "mailmen", it should be understood as including mail*women*, known sometimes as *femailmen*.

† A better model would account for the area's "carrying capacity" – a measure of how many dachshunds and mailmen the area can support before overcrowding leads to population decline.

Exercises.

19. Historically, the golden ratio arose from the following problem: Given any line segment, cut it into two pieces so that that the whole segment is to the longer piece as the longer piece is to the shorter piece. If we can do this, we define the numerical value of this ratio (whole/long, or equivalently, long/short) to be the **golden ratio**, ϕ.

Your problem: From the golden ratio's definition, deduce its exact numerical value.

[**Hint**: *Define the short piece of our divided line segment to be* 1 *unit. Call the long piece* x*, and use* ϕ*'s definition to deduce* x*'s length. Then by* ϕ*'s definition (again), it follows that* $\phi = x/1 = x$*, giving us* ϕ*'s numerical value.*]

20. The polynomial $x^2 - x - 1$ is intimately linked to the golden ratio: Its positive root is ϕ, while its negative root, which I'll call ψ, is ϕ's "conjugate". That is, the polynomial's two roots are:

$$\phi = \frac{1+\sqrt{5}}{2}, \qquad \psi = \frac{1-\sqrt{5}}{2}.$$

Demonstrate each of the following facts, some of which may prove useful to you in the next exercise:

a) $\phi^2 = 1 + \phi$ b) $\phi^{-1} = -\psi$ c) $\psi = 1 - \phi$ d) $\phi \approx 1.618$ e) $\psi \approx -0.618$

21. (The n^{th} Fibonacci number) Eigenstuff can lead us to an exact formula for F_n, the n^{th} Fibonacci number. In this exercise, I'll walk you through the argument, letting you fill in the details.

a) Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, the matrix we encountered in this section's Fibonacci example. Explain why

$$\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

b) Explain why it follows that

$$F_n = \begin{pmatrix} 0 & 1 \end{pmatrix} A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

c) Although the preceding equation gives us a recipe for F_n, it's not very satisfying, as it involves raising matrix A to a power, which is computationally expensive. But as you saw in Exercise 8D, we know a trick for raising a matrix to a power: *eigendecomposition*. Recall the eigendecomposition punchline (spelled out in Exercise 9B): If A admits an eigenbasis (as it does here), we can write $A = V\Lambda V^{-1}$, where Λ is a diagonal matrix whose diagonal entries are A's eigenvalues, while matrix V stores - in its corresponding columns - an eigenvector for each eigenvalue. Accordingly, eigendecomposition will turn Part B's formula into something of the form

$$F_n = \begin{pmatrix} 0 & 1 \end{pmatrix} V \Lambda^n V^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

This still contains a matrix raised to a power, but now it is a *diagonal* matrix, which is good, because raising a diagonal matrix to the n^{th} power is easy: we just raise each of its diagonal entries to the n^{th} power.

Your problems: Show that A's characteristic polynomial is $\lambda^2 - \lambda - 1$, and thus (by Exercise 20), A's eigenvalues are ϕ and ψ. Find a corresponding eigenvector for each eigenvalue. Having done this, conclude that the following expansion of our previous equation is valid:

$$F_n = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} \phi & \psi \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \phi & 0 \\ 0 & \psi \end{pmatrix}^n \begin{pmatrix} \phi & \psi \\ 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

d) Simplify the preceding until you obtain the following remarkable formula:

$$F_n = \frac{\phi^n - \psi^n}{\sqrt{5}}.^*$$

e) Use the preceding formula to compute F_{40} directly with a scientific calculator.

---

* Remarkably, a formula involving two irrational numbers always yields Fibonacci numbers, which are *integers*.

Chapter 8
Projections and Least Squares

Orthogonal Projection

JUSTICE SCALIA: What was that adjective? I liked that.
MR. FRIEDMAN: Orthogonal.
CHIEF JUSTICE ROBERTS: Orthogonal.
MR. FRIEDMAN: Right, right.
JUSTICE SCALIA: Orthogonal, ooh.
- Oral argument at the US Supreme Court in *Briscoe v. Virginia* (1/10/2010).

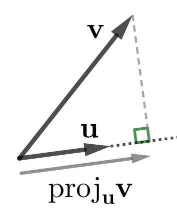
In earlier chapters, we've encountered orthogonal projections of vectors onto low-dimensional subspaces (lines and planes). In such low-dimensional contexts, we could rely exclusively on our geometric intuition. But in this chapter, we'll need to project vectors from $\mathbb{R}^n$ onto higher-dimensional subspaces, where our intuition is less reliable. What does it even mean, for example, to orthogonally project a vector in $\mathbb{R}^{10}$ onto a six-dimensional subspace? We obviously can't draw a faithful picture of such a situation, but as we'll see shortly, we can still figure out a sensible definition of such an orthogonal projection, use that definition to prove some theorems about such projections, and then use those theorems in applications.

We'll begin at the beginning – with an important formula for orthogonal projections onto a *line*.

Claim 1. The orthogonal projection of vector $\mathbf{v}$ onto the line spanned by a *unit* vector $\mathbf{u}$ is given by the following formula (and notation):

$$\text{proj}_{\mathbf{u}}\mathbf{v} = (\mathbf{v}\cdot\mathbf{u})\mathbf{u}.$$

Proof. Clearly, $\text{proj}_{\mathbf{u}}\mathbf{v} = \pm\|\text{proj}_{\mathbf{u}}\mathbf{v}\|\mathbf{u}$, where the sign depends on whether the angle between $\mathbf{v}$ and $\mathbf{u}$ is acute or obtuse. (If it's *right*, then the claim is obvious.) By the dot product's definition (Ch. 1), $\mathbf{v}\cdot\mathbf{u} = \pm\|\text{proj}_{\mathbf{u}}\mathbf{v}\|\|\mathbf{u}\| = \pm\|\text{proj}_{\mathbf{u}}\mathbf{v}\|$, where the sign depends once again on the type of angle between $\mathbf{u}$ and $\mathbf{v}$. Combining this with our first equation in this proof, we obtain



$$\text{proj}_{\mathbf{u}}\mathbf{v} = (\mathbf{v}\cdot\mathbf{u})\mathbf{u}$$

as claimed. ■

We'll use this projection formula soon, but before we do, I need to introduce a new definition and say a few words about it: An **orthonormal basis** is a basis consisting of *mutually perpendicular unit vectors*. Thus, for example, $\mathbb{R}^n$'s standard basis is orthonormal. We like orthonormal bases for many reasons, not least of which is that they facilitate orthogonal projections onto subspaces of any number of dimensions. Before we see how, we'll prove a useful theorem about representing a vector as a linear combination of an orthonormal basis's vectors.

Claim 2. Suppose a vector space has an *orthonormal* basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Then every vector $\mathbf{v}$ in that space is the sum of its orthogonal projections onto the orthonormal basis vectors:

$$\mathbf{v} = \sum_{i=1}^{n}\text{proj}_{\mathbf{u}_i}\mathbf{v} = (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}\cdot\mathbf{u}_n)\mathbf{u}_n.$$

Proof. We know, of course, that there exists a unique set of scalars $c_1, c_2, \dots, c_n$ such that

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n.$$

To determine what each scalar c_i must be, we will take the dot product of both sides with $\mathbf{u}_i$. Because the dot product distributes over vector addition, this yields

$$\mathbf{v}\cdot\mathbf{u}_i = (c_1\mathbf{u}_1)\cdot\mathbf{u}_i + (c_2\mathbf{u}_2)\cdot\mathbf{u}_i + \cdots + (c_i\mathbf{u}_i)\cdot\mathbf{u}_i + \cdots + (c_n\mathbf{u}_n)\cdot\mathbf{u}_i.$$

Pulling the scalars through each term's dot product, this becomes

$$\mathbf{v}\cdot\mathbf{u}_i = c_1(\mathbf{u}_1\cdot\mathbf{u}_i) + c_2(\mathbf{u}_2\cdot\mathbf{u}_i) + \cdots + c_i(\mathbf{u}_i\cdot\mathbf{u}_i) + \cdots + c_n(\mathbf{u}_n\cdot\mathbf{u}_i).$$

Since the basis here is orthonormal, the dot product of any two *distinct* basis vectors must be 0. Removing these zeros leaves us with

$$\mathbf{v}\cdot\mathbf{u}_i = c_i(\mathbf{u}_i\cdot\mathbf{u}_i).$$

Any vector's dot product with itself is the square of its length, so since all the basis vectors here are unit length (part of the definition of "orthonormal"), this dot product must be 1. Therefore,

$$\mathbf{v}\cdot\mathbf{u}_i = c_i$$

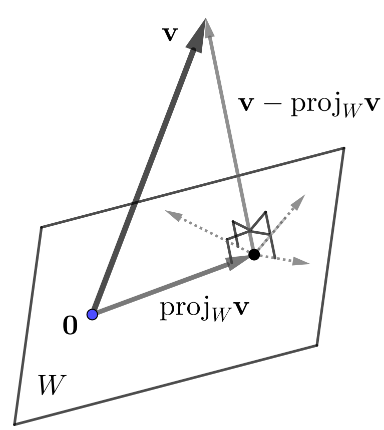
for all i. Substituting these back into our original expression for $\mathbf{v}$ as a linear combination of the orthonormal basis vectors, we obtain

$$\mathbf{v} = (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v}\cdot\mathbf{u}_n)\mathbf{u}_n,$$

which, by Claim 1, is the sum of $\mathbf{v}$'s orthogonal projections onto the basis vectors. ∎

The result we've just proved is important both pragmatically and theoretically. Pragmatically, it's a serious timesaver. Expressing $\mathbf{v}$ as a linear combination of basis vectors normally entails solving a linear system ($B\mathbf{x} = \mathbf{v}$, where B's columns are the basis vectors), but if the basis is orthonormal, we are freed from the dreary slog of Gaussian elimination. Instead, all that we - or our computer - must do is take some dot products. Nothing could be easier. As for the result's theoretical importance, we'll see an example of that in the discussion that follows, where it will help us meaningfully define the orthogonal projection of a vector onto arbitrary subspaces – not just onto lines and planes. To motivate that crucial definition, let's dwell a bit on something familiar: orthogonal projection onto a *plane*, where we can still visualize it.

What exactly does it mean to orthogonally project a vector $\mathbf{v}$ onto a *two*-dimensional subspace W? If $\mathbf{v}$ happens already to lie in the plane W, then of course $\mathbf{v}$'s projection onto W is vector $\mathbf{v}$ itself. But if $\mathbf{v}$ doesn't lie in W (as in the figure), we can precisely define $\mathbf{v}$'s orthogonal projection onto W as follows. Of all vectors in plane W, it's clear that one – and only one – possesses the following property: the vector from *its* tip to $\mathbf{v}$'s tip is perpendicular to *every* vector in plane W. We define this unique vector as $\mathbf{v}$'s orthogonal projection onto W, which we denote, unsurprisingly, as $\text{proj}_W\mathbf{v}$. The figure at right conveys the basic idea. Suitably amended, this definition can be generalized to define orthogonal projection of a vector onto *any* subspace, as we'll soon see.



It's tempting to leap triumphantly to our definition of an orthogonal projection onto *any* subspace W: $\text{proj}_W\mathbf{v}$ is the unique vector in W with the property that the vector from *its* tip to $\mathbf{v}$'s tip is perpendicular to all vectors in W, right? Well, this should work… but before we can endorse that definition, we must be sure that it still makes sense in higher dimensions! There are two questions to which we must attend. First, are we sure that a vector with that property even *exists*? When W is a line or a plane this is obvious, but when W is, say, a 17-dimensional hyperplane in $\mathbb{R}^{42}$, our geometric intuition fails, so we must prove the existence of such a vector analytically. Second, if a vector with this property does exist, is it *unique*? What if multiple vectors in W have that property? If so, we'd need to tighten up our proposed definition so that it specifies which one is "the" orthogonal projection. But if we can prove that only one vector has that property, then our proposed definition will work perfectly.

Claim 3. Given any vector $\mathbf{v}$ in a vector space V and any *sub*space W of the same vector space, there's a unique vector $\mathbf{p}$ in W with the property that $(\mathbf{v}-\mathbf{p})\cdot\mathbf{w}=0$ for *every* vector $\mathbf{w}$ in W.

Proof. Let $\mathbf{u}_1, \mathbf{u}_2, \dots \mathbf{u}_k$ be an orthonormal basis for W.[*] We know (from Claim 2) that every vector that lies *within* W is the sum of its orthogonal projections onto the orthonormal basis vectors $\mathbf{u}_i$. Vector $\mathbf{v}$, however, will generally lie *outside* of W. In that case, orthogonally projecting $\mathbf{v}$ onto each basis vector of W and summing those projections, all of which lie within W, obviously can't get us back to $\mathbf{v}$. But it will get us somewhere important! Namely, it will get us to a vector in W with the property we want. Call this vector sum $\mathbf{p}$, so that by definition,

$$\mathbf{p}=\sum_{i=1}^{k}\text{proj}_{\mathbf{u}_i}\mathbf{v}=(\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1+(\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2+\cdots+(\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k.$$

We must now show that $\mathbf{p}$ actually has the claimed property: namely, that $(\mathbf{v}-\mathbf{p})$, which extends from $\mathbf{p}$'s tip to $\mathbf{v}$'s tip, is perpendicular to all vectors in W. To see why this is so, we'll consider the dot product of $(\mathbf{v}-\mathbf{p})$ and an arbitrary vector $\mathbf{u}_i$ from W's orthonormal basis. Because the dot product distributes over vector addition, we have that

$$(\mathbf{v}-\mathbf{p})\cdot\mathbf{u}_i=(\mathbf{v}\cdot\mathbf{u}_i)-(\mathbf{p}\cdot\mathbf{u}_i)=(\mathbf{v}\cdot\mathbf{u}_i)-\left(\sum_{i=1}^{k}\text{proj}_{\mathbf{u}_i}\mathbf{v}\right)\cdot\mathbf{u}_i$$

$$=(\mathbf{v}\cdot\mathbf{u}_i)-\sum_{i=1}^{k}(\text{proj}_{\mathbf{u}_i}\mathbf{v})\cdot\mathbf{u}_i.$$

By Claim 1, this is equal to

$$(\mathbf{v}\cdot\mathbf{u}_i)-[(\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1\cdot\mathbf{u}_i+(\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2\cdot\mathbf{u}_i+\cdots+(\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k\cdot\mathbf{u}_i],$$

which by a basic property of dot products (scalars can be pulled out of dot products), is equal to

$$(\mathbf{v}\cdot\mathbf{u}_i)-[(\mathbf{v}\cdot\mathbf{u}_1)(\mathbf{u}_1\cdot\mathbf{u}_i)+(\mathbf{v}\cdot\mathbf{u}_2)(\mathbf{u}_2\cdot\mathbf{u}_i)+\cdots+(\mathbf{v}\cdot\mathbf{u}_k)(\mathbf{u}_k\cdot\mathbf{u}_i)].$$

[*] Clearly, every vector space admits an orthonormal basis. After all, if the space W is k-dimensional, we can imagine constructing an orthonormal basis for it as follows: Let *any* unit vector in W be $\mathbf{u}_1$. Now pick any vector in W that's perpendicular to $\mathbf{u}_1$ (there must be one, or W would be one-dimensional) and scale it down (or up) to unit length. Call it $\mathbf{u}_2$. Now pick any vector in W that's perpendicular to $\mathbf{u}_1$ *and* $\mathbf{u}_2$ (there must be one, or W would be two-dimensional) and scale it down to unit length. Call it $\mathbf{u}_3$. And so on until we have built up a collection of k mutually perpendicular unit vectors – i.e. an orthonormal basis for W.

Each term within the brackets can be simplified further, thanks to the orthonormality of the basis. Since the orthonormal basis vectors are mutually perpendicular, the dot product of $\mathbf{u}_i$ with any other $\mathbf{u}_j$ (where $j \neq i$) is zero. Thus, all but one of those terms are wiped out, leaving us with

$$(\mathbf{v} \cdot \mathbf{u}_i) - [(\mathbf{v} \cdot \mathbf{u}_i)(\mathbf{u}_i \cdot \mathbf{u}_i)].$$

Since the dot product of any vector with itself is the square of the vector's length, and $\mathbf{u}_i$ is unit length, it follows that $\mathbf{u}_i \cdot \mathbf{u}_i = 1$, reducing the expression above still further to

$$(\mathbf{v} \cdot \mathbf{u}_i) - (\mathbf{v} \cdot \mathbf{u}_i),$$

which is of course 0 for all i. Thus, we've shown that for all i,

$$(\mathbf{v} - \mathbf{p}) \cdot \mathbf{u}_i = 0.$$

That is, we've shown that $(\mathbf{v} - \mathbf{p})$, the vector extending from $\mathbf{p}$'s tip to $\mathbf{v}$'s tip, is perpendicular to all k of the vectors in our orthonormal basis for W. From this, it follows ("follows by linearity" as the pros would say) that $(\mathbf{v} - \mathbf{p})$ must in fact be perpendicular to *every* vector in W.*

We have now established the *existence* of a vector in W that satisfies the property we want. But is it unique? Is $\mathbf{p}$ the only such vector? To find out, let's suppose that

$$\mathbf{x} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k$$

represents *any* old vector in W that satisfies the property. To establish uniqueness, we must prove that any such $\mathbf{x}$ is in fact $\mathbf{p}$. To do that, we'll show that each coefficient c_i in $\mathbf{x}$'s expansion equals the corresponding coefficient in $\mathbf{p}$'s expansion; that is, we'll show that $c_i = \mathbf{v} \cdot \mathbf{u}_i$ for all i. (If this is unclear, review $\mathbf{p}$'s original definition.) To this end, recall the special property that $\mathbf{x}$ satisfies: $(\mathbf{v} - \mathbf{x}) \cdot \mathbf{w} = 0$ for all $\mathbf{w}$ in W. Written out more expansively, this says that for all $\mathbf{w}$ in W,

$$\big(\mathbf{v} - (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k)\big) \cdot \mathbf{w} = 0.$$

In particular, this must hold when $\mathbf{w} = \mathbf{u}_i$ for each i. That is, for each i, the following must hold:

$$\big(\mathbf{v} - (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k)\big) \cdot \mathbf{u}_i = 0.$$

After distributing that dot product (twice), we deduce that

$$(\mathbf{v} \cdot \mathbf{u}_i) - [c_1(\mathbf{u}_1 \cdot \mathbf{u}_i) + c_2(\mathbf{u}_2 \cdot \mathbf{u}_i) + \cdots + c_k(\mathbf{u}_k \cdot \mathbf{u}_i)] = 0.$$

For reasons explained earlier in this proof, the orthonormality of the basis guarantees that all but one of the bracketed terms will be zero, while the one exception, $c_i(\mathbf{u}_i \cdot \mathbf{u}_i)$, will be equal to c_i. The preceding equation thus reduces to $(\mathbf{v} \cdot \mathbf{u}_i) - c_i = 0$. Equivalently, it tells us that for each i, we must have $c_i = \mathbf{v} \cdot \mathbf{u}_i$. Substituting these expressions back into $\mathbf{x}$'s original expression yields

$$\mathbf{x} = (\mathbf{v} \cdot \mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v} \cdot \mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{v} \cdot \mathbf{u}_k)\mathbf{u}_k = \mathbf{p},$$

so $\mathbf{p}$ is in fact the unique vector in W with this property. ■

* Spelled out: Each vector $\boldsymbol{w}$ in W is some linear combination of the orthonormal basis vectors. We've seen above that $(\mathbf{v} - \mathbf{p})$ is perpendicular to each *component* (i.e. each term) of any such linear combination. Since the dot product distributes over vector addition, it follows that $(\mathbf{v} - \mathbf{p})$ must also be perpendicular to the components' *sum*, which is, of course, $\mathbf{w}$.

Thanks to the preceding result, we can now carefully define $\text{proj}_W\mathbf{v}$ for any vector $\mathbf{v}$ and subspace W. Moreover, we have even discovered how to compute it. Let's summarize our findings in a box.

Definition. The **orthogonal projection** of a vector $\mathbf{v}$ onto a subspace W (which we denote by $\text{proj}_W\mathbf{v}$) is the unique vector in W with the property that the vector extending from its tip to $\mathbf{v}$'s tip is perpendicular to every vector in the subspace W.

Projection Formula.

If $\mathbf{u}_1, \mathbf{u}_2, \ldots \mathbf{u}_k$ is any orthonormal basis for W, then $\text{proj}_W\mathbf{v}$ is the sum of $\mathbf{v}$'s orthogonal projections onto W's orthonormal basis vectors. That is,

$$\begin{aligned}\text{proj}_W\mathbf{v} &= \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i}\mathbf{v} \\ &= (\mathbf{v}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{v}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots (\mathbf{v}\cdot\mathbf{u}_k)\mathbf{u}_k.\end{aligned}$$

Hooray! Ah… but even as we raucously celebrate the theoretical achievement of our projection definition, we must also note an apparent pragmatic difficulty with our projection formula: The projection formula requires us to have an orthonormal basis for W already. As discussed in a footnote two pages back, we can confidently assert that every subspace *has* an orthonormal basis, but if we wish to use our projection formula, we can't rest content with knowing that an orthonormal basis exists in the mind of God. We must actually be able to construct an orthonormal basis for W with our bare hands (or bare computers) so that we have not merely a theoretical orthonormal basis, but a concrete one – one whose vectors' standard coordinates can be listed in black and white on a page or stored in a computer program.

In the next section, we'll learn the so-called Gram-Schmidt Orthonormalization Process, which starts with any basis of a subspace and "upgrades" it into an orthonormal basis. This will dispose of our pragmatic difficulty with the projection formula, rendering it moot.

Exercises.

1. In Claim 1, I proved that $\text{proj}_{\mathbf{u}}\mathbf{v} = (\mathbf{v} \cdot \mathbf{u})\mathbf{u}$ for any vector $\mathbf{v}$ and unit vector $\mathbf{u}$. To accompany the proof, I only drew a figure for the case in which the angle between $\mathbf{v}$ and $\mathbf{u}$ is acute. Draw a picture for the obtuse case and be sure that you understand why the proof holds in that case, too.

2. Find the orthogonal projection of vector $2\mathbf{i} + 3\mathbf{j} - \mathbf{k}$ onto the line through the origin and $(1, 5, 2)$. [*Hint*: First find a unit vector lying on the line.]

3. Find the orthogonal projections of $\mathbf{v}_1 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} -1 \\ -2 \\ 0 \end{pmatrix}$ onto $\mathbf{u} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$.

4. Find the orthogonal projection of $\mathbf{v} = \begin{pmatrix} -1 \\ 2 \\ -4 \end{pmatrix}$ onto the plane spanned by $\mathbf{u}_1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ and $\mathbf{u}_2 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

5. Consider the following two bases of $\mathbb{R}^4$, whose vectors are all given in standard coordinates:

$$\mathcal{A}: \quad \mathbf{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 2 \\ 3 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{a}_4 = \begin{pmatrix} 5 \\ 4 \\ 3 \\ 1 \end{pmatrix};$$

$$\mathcal{B}: \quad \mathbf{b}_1 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} -1/2 \\ -1/2 \\ 1/2 \\ 1/2 \end{pmatrix}, \quad \mathbf{b}_3 = \begin{pmatrix} -1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{pmatrix}, \quad \mathbf{b}_4 = \begin{pmatrix} -1/2 \\ 1/2 \\ 1/2 \\ -1/2 \end{pmatrix}.$$

a) Is either basis orthonormal?

b) Let $\mathbf{v} = 10\mathbf{e}_1 + 7\mathbf{e}_2 + 7\mathbf{e}_3 + 3\mathbf{e}_4$ be a vector in $\mathbb{R}^4$. Express $\mathbf{v}$ as a linear combination of $\mathcal{A}$'s vectors. Then express it as a linear combination of $\mathcal{B}$'s vectors. In the latter case, you should be able to find the linear combination *without* resorting to Gaussian elimination. Do so.

6. Find the orthogonal projection of $\mathbf{v} = 10\mathbf{e}_1 + 7\mathbf{e}_2 + 7\mathbf{e}_3 + 3\mathbf{e}_4$ onto the subspace of $\mathbb{R}^4$ spanned by

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}.$$

(Careful!)

7. If we want to let the "target" vector in an orthogonal projection be *any* vector $\mathbf{w}$ (not necessarily a unit vector), we can come up with a more complicated projection formula:

$$\text{proj}_{\mathbf{w}}\mathbf{v} = \left(\frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}}\right)\mathbf{w}.$$

Explain why this formula holds – but there's no need to memorize it. [*Hint*: Let $\mathbf{u} = \mathbf{w}/\|\mathbf{w}\|$.]

8. The matrix representations in this exercise should all be with respect to the standard basis.

a) Find the matrix A that orthogonally projects points in $\mathbb{R}^2$ onto the x-axis.
b) Find the matrix B that orthogonally projects points in $\mathbb{R}^2$ onto the line $y = 2x$.
c) What is the rank of the matrix AB? Explain your answer geometrically.
d) Find two 2×2 matrices A and B, both having positive rank, such that AB has rank 0.

9. In Chapter 1, we defined $\mathbf{v}$ and $\mathbf{w}$'s dot product as the product of their scalar projections onto $\mathbf{w}$. We then used this geometric definition to give coordinate-free proofs of six of the most familiar properties of the dot product: (1) It is commutative. (2) It distributes over vector addition. (3) We can pull scalars out of it. (4) It equals 0 if and only if the vectors are perpendicular. (5) Dotting any vector with itself yields the square of the vector's length. (6) the dot product of any two vectors is the product of their lengths and the cosine of the angle between them. All six of those properties are intrinsic to the dot product itself and have nothing to do with coordinates.

The one property whose proof required us to use coordinates was the "dot product formula", which states that the dot product of $\mathbf{v}$ and $\mathbf{w}$ can be obtained by summing the products of their corresponding coordinates.

Unlike the other six properties, that "dot product formula" is *not* intrinsic to the dot product itself. Rather, the formula arose from the marriage of properties of the dot product *and properties of the standard basis vectors*. Hence, if we switch to a different basis, that formula may no longer produce the dot product as we've defined it. In Chapter 6 (Exercise 5), you saw an example where that happens.

We'll now refine our understanding of this formula yet again. In this exercise, you'll prove that if we use not just any old nonstandard basis, but an *orthonormal* basis, then the "dot product formula" **does** in fact hold: We can count on it to find the dot product as we've defined it. (This is yet another reason to love orthonormal bases.)

a) Let $\mathbf{v}$ and $\mathbf{w}$ be vectors in a vector space with an orthonormal basis $\mathcal{B}$ consisting of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$. Suppose that

$$[\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \quad \text{and} \quad [\mathbf{w}]_{\mathcal{B}} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}.$$

Prove that $\mathbf{v} \cdot \mathbf{w} = v_1 w_1 + v_2 w_2 + \cdots v_n w_n$.

[**Hint**: Review our proof of this result (in the specific context of $\mathcal{E}$) from Chapter 1. With a few tiny changes, you can adapt it to the present circumstances.]

b) In Exercise 5 of Chapter 6, you also saw that the familiar formula for a vector's *length* (the square root of the sum of its squared components) can fail in a non-standard coordinate system. Happily, it works in a coordinate system based on an *orthonormal* basis. Prove that this is so.

[**Hint**: You can adapt the proof from Chapter 1, but there's a *much* easier way. It follows directly from Part A.]

The Gram-Schmidt Process

My Lord, he's going to his mother's closet.
Behind the arras I'll convey myself
To hear the process.
- Polonius (*Hamlet*, Act III, Scene 3)

Let us imagine Mr. Gram and Mr. Schmidt, who own a shop. You bring them a subspace's battered basis, its vectors all pointing in crazy directions and sticking out at different lengths, and for a fee, they'll twist, snip, and extend its vectors as needed until they're tidy again: unit length and mutually perpendicular. They do this by employing their patented Gram-Schmidt Process, whose details you'll learn in this section.

The goal of the Gram-Schmidt Process is to convert any basis of a subspace into an *orthonormal* basis. The process is easy to grasp geometrically, even if carrying out its details by hand can be quite a hassle. I'll introduce the idea in the context of a 2-dimensional subspace and then build up from there.

Example. The graph of $x - 2y + 3z = 0$ is a two-dimensional subspace of $\mathbb{R}^3$. Let us call it W. As you should verify, a basis for W is

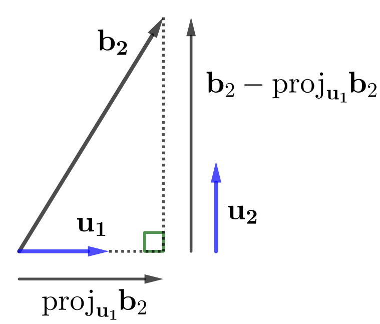
$$\mathbf{b}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{b}_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Now let's go to Gram and Schmidt's Olde Orthonormalization Shoppe and watch (from behind the arras) how they upgrade this basis to an orthonormal one, whose vectors they'll call $\mathbf{u}_1$ and $\mathbf{u}_2$.

Observations. First**,** Gram *normalizes* $\mathbf{b}_1$, dividing it by its own length to turn it into a unit vector. "Well, Schmidt," he says, "My work's done. See you later." After writing his result on a blackboard,

$$\mathbf{u}_1 = \tfrac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix},$$

Gram dons his hat and goes out for a cup of coffee, leaving Schmidt to find $\mathbf{u}_2$. Schmidt quickly draws a schematic picture of $\mathbf{u}_1$ and $\mathbf{b}_2$, then sketches the latter's projection onto the former, as at right. Finally, he draws a vector extending from the projection's tip to $\mathbf{b}_2$'s tip. He calculates that this vector, which clearly is orthogonal to $\mathbf{u}_1$, is



$$\mathbf{b}_2 - \mathrm{proj}_{\mathbf{u}_1}\mathbf{b}_2 \;=\; \mathbf{b}_2 - (\mathbf{b}_2 \cdot \mathbf{u}_1)\mathbf{u}_1 \;=\; \begin{pmatrix} -3/5 \\ 6/5 \\ 1 \end{pmatrix}.$$

This vector is properly oriented, but Schmidt still needs to normalize it. After calculating that its length is $\sqrt{70}/5$, as you should verify, he divides his vector by this length, and concludes that

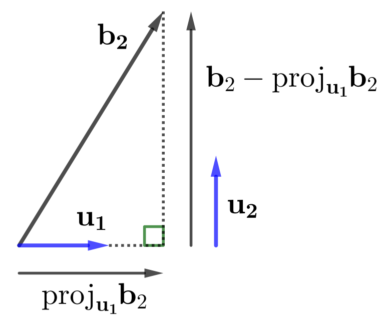
$$\mathbf{u}_2 = \tfrac{5}{\sqrt{70}}\begin{pmatrix} -3/5 \\ 6/5 \\ 1 \end{pmatrix}.$$

After recording this on the board, he grabs his cane and heads down to the coffee shop himself. It's a good life in the orthonormalization business. ♦

If you've understood that first example, you already understand most of the Gram-Schmidt Process. The process just requires one extra step for each extra dimension in the subspace. Suppose, for example, that we take a basis $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ for some 3-dimensional subspace to Gram and Schmidt's shop. To upgrade it to an orthonormal basis, they'll begin as before: Gram normalizes the first vector to obtain

$$\mathbf{u}_1 = \mathbf{b}_1/\|\mathbf{b}_1\|.$$

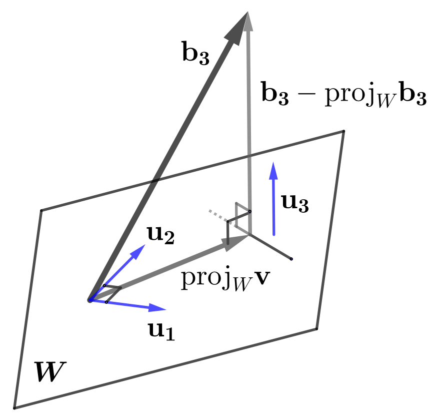
Now Schmidt redraws the same picture he drew on the previous page, and lets it guide him to the second vector:



$$\mathbf{u}_2 = \frac{\mathbf{b}_2 - \text{proj}_{\mathbf{u}_1}\mathbf{b}_2}{\|\mathbf{b}_2 - \text{proj}_{\mathbf{u}_1}\mathbf{b}_2\|} = \frac{\mathbf{b}_2 - (\mathbf{b}_2 \cdot \mathbf{u}_1)\mathbf{u}_1}{\|\mathbf{b}_2 - (\mathbf{b}_2 \cdot \mathbf{u}_1)\mathbf{u}_1\|}.$$

"But how do you remember that crazy formula?" an onlooker asks. Schmidt, squinting through his monocle, looks up incredulously and laughs. "I don't. No one, apart from a fool, would ever memorize that ugly thing. Why do you think I drew the picture? It shows me exactly what I must do and why it will work."

And now Gram – who hasn't yet stepped out for coffee – sketches his own picture, which guides him to a third vector for the orthonormal basis that he's constructing. Gram's picture is much like Schmidt's, but now he's projecting $\mathbf{b}_3$ (instead of $\mathbf{b}_2$) onto the subspace that is spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$ (instead of the subspace spanned by $\mathbf{u}_1$ alone). I've labelled this subspace W in the figure. Thanks to our work in the previous section, it's very easy to project $\mathbf{b}_3$ orthogonally onto W: We just add up $\mathbf{b}_3$'s orthogonal projections onto $\mathbf{u}_1$ and $\mathbf{u}_2$. Thus,



$$\mathbf{u}_3 = \frac{\mathbf{b}_3 - \text{proj}_W\mathbf{b}_3}{\|\mathbf{b}_3 - \text{proj}_W\mathbf{b}_3\|} = \frac{\mathbf{b}_3 - (\mathbf{b}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{b}_3 \cdot \mathbf{u}_2)\mathbf{u}_2}{\|\mathbf{b}_3 - (\mathbf{b}_3 \cdot \mathbf{u}_1)\mathbf{u}_1 - (\mathbf{b}_3 \cdot \mathbf{u}_2)\mathbf{u}_2\|}.$$

(Again, this isn't something to memorize. Just think your way through the process, letting your thoughts and a figure guide you.) And with that, Gram and Schmidt shake hands and congratulate each other on a job well done.

But what if the subspace whose basis we want to upgrade has 4 (or even, say, 400,000) dimensions? The Gram-Schmidt Process still works. To upgrade a given basis $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \ldots, \mathbf{b}_n$ to an *orthonormal* basis, $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_n$, we still proceed one vector at a time.

In Step 1, we normalize $\mathbf{b}_1$ to obtain $\mathbf{u}_1$.

For every step i beyond that, we do the same thing: We orthogonally project $\mathbf{b}_i$ onto the subspace spanned by $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}$. We know from our work in the previous section that the vector extending from the orthogonal projection's tip to $\mathbf{b}_i$'s tip (namely, $\mathbf{b}_i - \text{proj}_{\text{span}(\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1})}\mathbf{b}_i$) points in a direction that suits our purposes (namely, perpendicular to all the $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{i-1}$ that we've already constructed). Now all we must do is normalize this vector to give it unit length, and we can declare the result to be $\mathbf{u}_i$.

Do this until we've obtained $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_n$, and that's that. The Gram-Schmidt Process is complete, and we have our orthonormal basis.

In the exercises that follow, you'll have the opportunity to try it out yourself.

Exercises.

10. Consider the following sets of vectors. The span of each set is, of course, a subspace, and since the vectors in each set are linearly independent, they constitute a basis for the space they span. For each of these spaces, come up with an *orthonormal* basis. Draw pictures like those on the previous page to guide you.

a) $\begin{pmatrix}2\\1\\2\end{pmatrix}, \begin{pmatrix}1\\0\\5\end{pmatrix}$ b) $\begin{pmatrix}5\\0\\0\end{pmatrix}, \begin{pmatrix}4\\3\\0\end{pmatrix}, \begin{pmatrix}5\\-7\\6\end{pmatrix}$ c) $\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}7\\0\\2\\7\end{pmatrix}, \begin{pmatrix}8\\6\\0\\0\end{pmatrix}$

11. Earlier in this chapter, we saw that if W, a k-dimensional subspace of $\mathbb{R}^n$, has an orthonormal basis $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k$, then the orthogonal projection of any vector $\mathbf{v}$ in $\mathbb{R}^n$ onto W is the sum of its projections onto W's orthonormal basis vectors. That is,

$$\text{proj}_W \mathbf{v} = \sum_{i=1}^{k} \text{proj}_{\mathbf{u}_i} \mathbf{v}.$$

Consider the vector $\mathbf{v} = 5\mathbf{i} + 3\mathbf{j} - 5\mathbf{k}$ in $\mathbb{R}^3$.

a) Find $\mathbf{v}$'s orthogonal projection onto the xy-plane. (No need to compute. Just think geometrically.)
b) Find $\mathbf{v}$'s orthogonal projection onto the yz-plane. (Same story.)
c) Let W_1 be the subspace (a plane) of $\mathbb{R}^3$ consisting of points satisfying the equation $-x + y - z = 0$. Verify that the following vectors constitute an *orthonormal* basis for W_1:

$$\mathbf{u}_1 = \tfrac{1}{\sqrt{2}}\begin{pmatrix}1\\1\\0\end{pmatrix} \quad \text{and} \quad \mathbf{u}_2 = \tfrac{1}{\sqrt{6}}\begin{pmatrix}1\\-1\\-2\end{pmatrix}.$$

d) Find $\mathbf{v}$'s orthogonal projection onto W_1.
e) Let W_2 be the subspace consisting of points satisfying the equation $2x + y - 3z = 0$. Find any old basis of W_2. You must, of course, verify that your prospective basis vectors do in fact lie in W_2, and that they constitute a basis for that plane.
f) Use the Gram-Schmidt Process to "upgrade" your basis for W_2 to an *orthonormal* basis.
g) Find $\mathbf{v}$'s orthogonal projection onto W_2.

12. (QR Decomposition) If a matrix A's columns are linearly independent, it turns out that we can always decompose it into two factors, $A = QR$, where Q's columns are unit length and orthogonal, and R is an upper triangular matrix. For example, the matrix

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

has the following "QR Decomposition":

$$\underbrace{\begin{pmatrix}1 & -1 & 4\\1 & 4 & -2\\1 & 4 & 2\\1 & -1 & 0\end{pmatrix}}_{A} = \underbrace{\begin{pmatrix}1/2 & -1/2 & 1/2\\1/2 & 1/2 & -1/2\\1/2 & 1/2 & 1/2\\1/2 & -1/2 & -1/2\end{pmatrix}}_{Q}\underbrace{\begin{pmatrix}2 & 3 & 2\\0 & 5 & -2\\0 & 0 & 4\end{pmatrix}}_{R}$$

I'll leave it to you to verify that the product of the matrices on the right really is the matrix on the left, and that the "Q" matrix's columns are indeed unit length and mutually orthogonal.

Like an eigendecomposition, a QR decomposition of a matrix can be useful in various contexts – especially when implementing efficient algorithms in a computer program. (Computer scientists are especially fond of it.) In this exercise, you'll see why the QR decomposition exists, and how to find it. I'll begin by defining the matrices Q and R. We'll then prove that Q and R have the properties that I've claimed for them above.

Let's get to work.

Since A's columns are linearly independent, these columns constitute a basis $\mathcal{A}$ for the space they span.
Carrying out Gram-Schmidt on A's columns yields an *orthonormal* basis $\mathcal{B}$ for the same space.
Definition: Q is the matrix whose columns are the vectors of this new orthonormal basis $\mathcal{B}$.
Definition: R is the $\mathcal{A}$-to-$\mathcal{B}$ change of basis matrix.

Given these definitions of Q and R, we must now prove that R is an upper triangular matrix, and that $A = QR$.

a) Explain why R, as defined above, is an upper triangular matrix.

b) To prove that $A = QR$, we'll show that feeding the same input to A and QR always yields the same output. This argument is a little trickier, so your job here is just to verify that the following argument holds:

First, we'll set the stage with some classical scenery: columns. Namely, let

$$A = \begin{pmatrix} | & | & & | \\ \mathbf{a_1} & \mathbf{a_2} & \cdots & \mathbf{a_n} \\ | & | & & | \end{pmatrix} \text{ and } Q = \begin{pmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \cdots & \mathbf{b_n} \\ | & | & & | \end{pmatrix}, \quad \text{and let} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

be any old list of n numbers. This last column will be the common input we'll feed to both A and QR.

A maps this last column, of course, to $c_1\mathbf{a_1} + c_2\mathbf{a_2} + \cdots + c_n\mathbf{a_n}$, which I'll call $\mathbf{w}$ for short.

We'll now show that QR maps the common input to $\mathbf{w}$, too. To this end, define $d_1, \ldots, d_n$ as the unique set of scalars that make $d_1\mathbf{b_1} + d_2\mathbf{b_2} + \cdots + d_n\mathbf{b_n} = \mathbf{w}$. (That is, the d_i are $\mathbf{w}$'s $\mathcal{B}$-coordinates). Then

$$QR\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = QR[\mathbf{w}]_{\mathcal{A}} = Q[\mathbf{w}]_{\mathcal{B}} = \begin{pmatrix} | & | & & | \\ \mathbf{b_1} & \mathbf{b_2} & \cdots & \mathbf{b_n} \\ | & | & & | \end{pmatrix}\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} = d_1\mathbf{b_1} + d_2\mathbf{b_2} + \cdots + d_n\mathbf{b_n} = \mathbf{w}.$$

Since A and QR have the same effect on any common input, they are equal. Thus, $A = QR$, as claimed.

c) Now we know that the QR decomposition exists, and we know how to find Q (via Gram-Schmidt). We could compute R by going through the slog of computing the $\mathcal{A}$-to-$\mathcal{B}$ change of basis matrix, but happily, there's a quicker way. To understand the quick way, you'll first need to prove a little lemma:

If Q is any matrix whose columns are unit length and mutually perpendicular, then $Q^TQ = I$.

d) With the lemma in hand, explain why the following quick formula for R (once we have Q) works: $R = Q^TA$.

e) In practice, no one does QR-decomposition by hand. Find a matrix calculator online that will do it for you and try it out on some matrices. (The matrices should have linearly independent columns, of course.)

f) (An example of how QR factorization can speed up a numerical algorithm.)
Solving $A\mathbf{x} = \mathbf{b}$ by Gaussian elimination can be computationally expensive if A is a large matrix. What to do? One strategy: Decompose A into QR. Then we have $QR\mathbf{x} = \mathbf{b}$, or equivalently, $R\mathbf{x} = Q^T\mathbf{b}$ (thanks to Part C). Because transposing a matrix is computationally cheap, as is matrix-vector multiplication, the vector $Q^T\mathbf{b}$ is inexpensive to compute. Call it $\mathbf{c}$. Now it remains to solve $R\mathbf{x} = \mathbf{c}$, where R is an upper-triangular matrix.

Your problem: Explain why $R\mathbf{x} = \mathbf{c}$ will be much less computationally expensive to solve than $A\mathbf{x} = \mathbf{b}$.

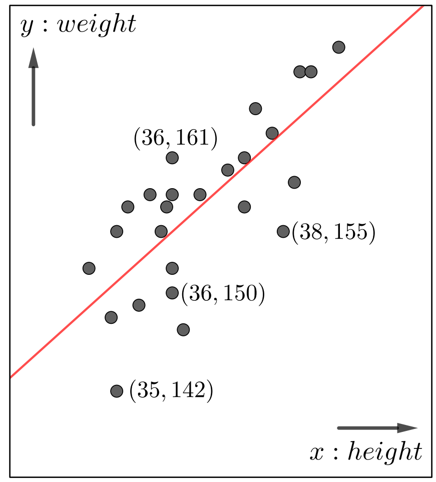
g) There are other matrix factorizations that are useful for creating efficient algorithms. One of the best-known is called LU decomposition. Look it up and see what it's all about.

Best Fit Lines, Least-Squares Solutions

I fight against thee? No! I will go seek
Some ditch wherein to die; the foul'st best fits
My latter part of life.
- Enobarbus (*Antony and Cleopatra*, Act IV, Scene 6)

Suppose we've collected some paired data: the heights and weights of 25 adult male Martians. Each pair yields a point in $\mathbb{R}^2$ whose first coordinate is a given Martian's height (in inches) and whose second coordinate is the same Martian's weight (in Martian pounds):

$$\begin{pmatrix}36\\150\end{pmatrix}, \begin{pmatrix}38\\155\end{pmatrix}, \begin{pmatrix}35\\142\end{pmatrix}, \dots, \begin{pmatrix}36\\161\end{pmatrix}.$$



We can then graph our data in a scatterplot, as I've done at right. The mysterious line running through the scatterplot is often called the "best-fit line" because it supposedly "fits" the data as best as a line possibly can. In statistics courses, the meaning of "best" tends to be explained hastily, with many questions left unanswered.* Linear algebra offers a crystal-clear conceptual explanation of what that line is and why it can lay claim to the title of "best fit". This section is devoted to that explanation.

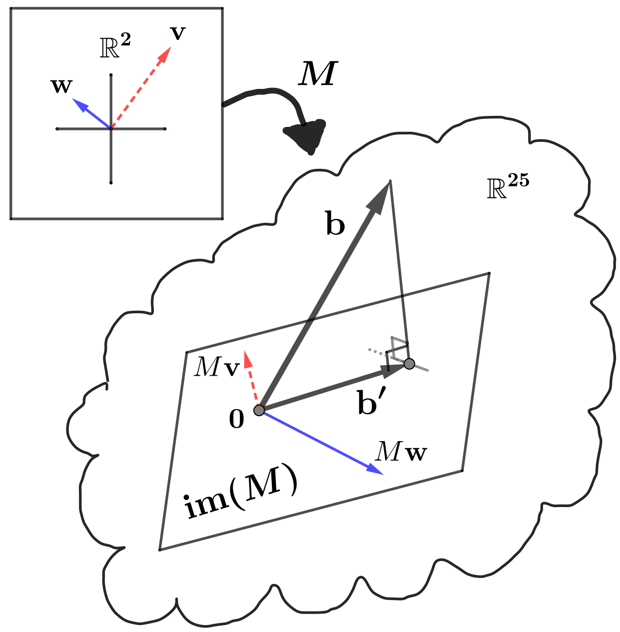
We'll begin with an act of imagination. We want a line that in some sense comes as close as possible to running through all 25 points in our scatterplot. The ideal, of course, would be a line that passes through all 25 points, which is clearly impossible. Still, playing with impossible ideas can be mathematically fruitful, so let's pretend that there *is* an ideal line, $y = c_1 x + c_2$, which, by some act of mathematical black magic, does pass through all 25 points. Popping our points' coordinates into its equation, we'd find that

$$\begin{aligned} 36c_1 + c_2 &= 150 \\ 38c_1 + c_2 &= 155 \\ 35c_1 + c_2 &= 142 \\ &\vdots \\ 36c_1 + c_2 &= 161 \end{aligned}$$

This is, I repeat, a hopelessly inconsistent system of 25 linear equations in 2 unknowns, but still, let's play. Recasting this system into $M\mathbf{x} = \mathbf{b}$ form, $\mathbf{c} = (c_1, c_2)$ is our "fantasy solution" to the inconsistent system

$$\underbrace{\begin{pmatrix} 36 & 1 \\ 38 & 1 \\ 35 & 1 \\ \vdots & \vdots \\ 36 & 1 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} 150 \\ 155 \\ 142 \\ \vdots \\ 161 \end{pmatrix}}_{\mathbf{b}}.$$

* The idea in a nutshell: Draw any old line through the scatterplot. Relative to that line, each point has a "residual": the (signed) vertical distance between it and the line. The line then gets a score: the sum of the *squared* residuals. Of all lines in the plane, the one with the lowest score is deemed the "best-fit line", or – as it's also called – the *least-squares line*. At the end of this chapter (Exercise 21), you'll see that our cleaner linear algebraic approach to best-fit lines still satisfies this criterion.



Rewriting our system as $M\mathbf{x} = \mathbf{b}$ allows us to see its geometry in a strikingly different way: We can now view M as a transformation from $\mathbb{R}^2$ to $\mathbb{R}^{25}$. We *wish* there was some ideal vector $\mathbf{c} = (c_1, c_2)$ in $\mathbb{R}^2$ that M would map to the vector $\mathbf{b}$ in $\mathbb{R}^{25}$. But of course, no such vector $\mathbf{c}$ actually exists. Matrix M can send inputs from $\mathbb{R}^2$ to any vector in $\text{im}(M)$, a 2-dimensional subspace of $\mathbb{R}^{25}$, but it can't reach vectors outside of it. In particular, it can't reach our "target vector" $\mathbf{b}$, which sticks out into the ambient space of $\mathbb{R}^{25}$.

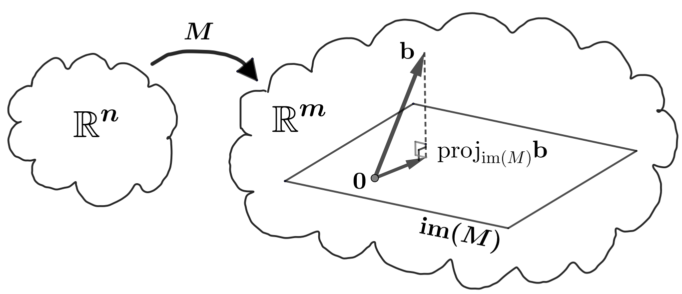
Matrix M can't "reach" $\mathbf{b}$ as an output, but… *how close* can it get? Clearly, of all vectors in $\text{im}(M)$, the one lying *closest* to $\mathbf{b}$ is $\text{proj}_{\text{im}(M)}\mathbf{b}$, which I'll call $\mathbf{b}'$ for short. (See the figure at right.) Since $\mathbf{b}'$ is in M's image, we know there *is* a vector $\mathbf{s}$ in $\mathbb{R}^2$ that gets mapped to it. This vector $\mathbf{s}$, which satisfies $M\mathbf{x} = \mathbf{b}'$, is thus the best *approximation* to a solution to our inconsistent system $M\mathbf{x} = \mathbf{b}$. While $\mathbf{c} = (c_1, c_2)$ was a "fantasy solution" giving rise to a fictional line $y = c_1x + c_2$ passing through all the scatterplot's points, $\mathbf{s} = (s_1, s_2)$ gives rise to $y = s_1x + s_2$, which is a very real line: the best-fit line.

We call $\mathbf{s}$ the **least-squares solution** for the inconsistent system $M\mathbf{x} = \mathbf{b}$. The least-squares solution is the best *approximation* to a solution that we can find.

Although I motivated the idea of a least-squares solution to an inconsistent system by focusing on a specific application - finding a best-fit line for a scatterplot - we can find a least-squares solution for *any* inconsistent linear system whatsoever.

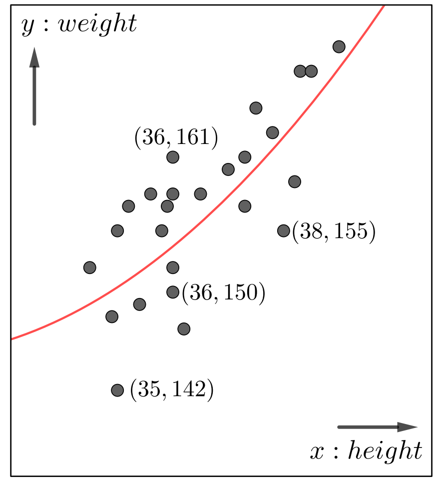
> If $M\mathbf{x} = \mathbf{b}$ is an inconsistent system, we obtain the best *approximation* to a solution (called a **least-squares solution** for $M\mathbf{x} = \mathbf{b}$) by solving $M\mathbf{x} = \text{proj}_{\text{im}(M)}\mathbf{b}$ instead.*

The schematic picture for the general case is essentially the same as the specific one I've drawn above. The only difference is that if M is an $m \times n$ matrix, then the "domain space" and "target space" will be $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively (to replace the specific cases of $\mathbb{R}^2$ and $\mathbb{R}^{25}$ above).



* The equation $M\mathbf{x} = \mathbf{b}'$ typically has a unique solution in best-fit applications, but it can, in principle, have multiple solutions, which will happen if M's columns are linearly dependent. In that case, we'd have a "tie for first place" in the "best fit" contest.

Returning to the specific theme of scatterplots, let's consider some natural variations on the theme of best-fit lines. For example, one might wish to fit a *second*-degree polynomial (rather than a linear one) to a paired data set. To do this, we need only make a few tiny adjustments in the preceding argument. For example, suppose that we want to determine the best-fit *quadratic* function for our paired Martian height/weight data. Well, if there existed a "fantasy quadratic" $y = c_1x^2 + c_2x + c_3$ that somehow passed through all 25 data points, the following equations would hold:



$$\begin{aligned} 36^2c_1 + 36c_2 + c_3 &= 150 \\ 38^2c_1 + 36c_2 + c_3 &= 155 \\ 35^2c_1 + 36c_2 + c_3 &= 142 \\ &\vdots \\ 36^2c_1 + 36c_2 + c_3 &= 161. \end{aligned}$$

Recasting this into $M\mathbf{x} = \mathbf{b}$ form, $\mathbf{c} = (c_1, c_2, c_3)$ is our "fantasy solution" to the inconsistent system

$$\underbrace{\begin{pmatrix} 36^2 & 36 & 1 \\ 38^2 & 38 & 1 \\ 35^2 & 35 & 1 \\ \vdots & \vdots & \vdots \\ 36^2 & 36 & 1 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} 150 \\ 155 \\ 142 \\ \vdots \\ 161 \end{pmatrix}}_{\mathbf{b}}.$$

Alas, $\mathbf{b}$ isn't in M's image, so a solution $\mathbf{c}$ to $M\mathbf{x} = \mathbf{b}$ really is just fantasy. Still, we can get a least-squares solution for our inconsistent system by solving $M\mathbf{x} = \mathbf{b}'$ instead, where $\mathbf{b}' = \text{proj}_{\text{im}(M)}\mathbf{b}$. Doing that will yield a vector $\mathbf{s} = (s_1, s_2, s_3)$ whose components will be the coefficients of the best-fit parabola we seek, $y = s_1x^2 + s_2x + s_3$.

I've not worked out the numerical details of the two examples above (or the one example below) for two main reasons. First: The details are gory. (We'd need to compute $\mathbf{b}' = \text{proj}_{\text{im}(M)}\mathbf{b}$, which involves producing a basis for $\text{im}(M)$, then upgrading it to an orthonormal basis, and then finally solving $M\mathbf{x} = \mathbf{b}'$ via Gaussian elimination.) Second: In the next section, I'll introduce a more efficient way that builds directly on the concepts you're learning here, but which *won't* require us to explicitly compute $\mathbf{b}'$.

One last example: Fitting various functions to data is part of what statisticians call *regression analysis*. The examples above are considered "*single* regression" because they produce an equation that "predicts" something (Martian weight) from a *single* variable (Martian height). In contrast, *multiple regression* produces an equation that predicts something from *multiple* variables. For example, suppose that Martian weights depend both on height and number of fingers. The data we collect from our sample of 25 Martians should thus consist of ordered triples: height, number of fingers, and weight. Something like this:

$$\begin{pmatrix} 36 \\ 12 \\ 150 \end{pmatrix}, \begin{pmatrix} 38 \\ 12 \\ 155 \end{pmatrix}, \begin{pmatrix} 35 \\ 10 \\ 142 \end{pmatrix}, \dots, \begin{pmatrix} 36 \\ 14 \\ 161 \end{pmatrix}.$$

Although we could make a scatterplot in $\mathbb{R}^3$, such things are hard to read and not very useful in practice. Still, you might imagine a cloud of points in $\mathbb{R}^3$ and the best-fit *plane* $z = c_1x + c_2y + c_3$ for that data. The analysis we've run through twice already will work here, too. An idealized but impossible plane that ran through all 25 points would lead to this system of equations:

$$\begin{gathered} 36c_1 + 12c_2 + c_3 = 150 \\ 38c_1 + 12c_2 + c_3 = 155 \\ 35c_1 + 10c_2 + c_3 = 142 \\ \vdots \\ 36c_1 + 14c_2 + c_3 = 161. \end{gathered}$$

The vector (c_1, c_2, c_3) can be understood as a "fantasy solution" to an inconsistent linear system, which we can rewrite in $M\mathbf{x} = \mathbf{b}$ form:

$$\underbrace{\begin{pmatrix} 36 & 12 & 1 \\ 38 & 12 & 1 \\ 35 & 10 & 1 \\ \vdots & \vdots & \vdots \\ 36 & 14 & 1 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} 150 \\ 155 \\ 142 \\ \vdots \\ 161 \end{pmatrix}}_{\mathbf{b}}.$$

Of course, this equation lacks a solution, but we know that the closely related equation $M\mathbf{x} = \text{proj}_{\text{im}(M)}\mathbf{b}$ *has* some solution $\mathbf{s} = (s_1, s_2, s_3)$ (the least-squares solution of our original system), whose components will be the coefficients of our best-fit plane's equation.

In the next section, you will see how we can dramatically simplify the computational side of finding a least-squares solution. But first, a few exercises.

Exercises.

13. The following paired data shows 30 students' scores on their midterm and final exams in a linear algebra class:

$$\begin{pmatrix} 90 \\ 96 \end{pmatrix}, \begin{pmatrix} 72 \\ 80 \end{pmatrix}, \begin{pmatrix} 70 \\ 35 \end{pmatrix}, \cdots, \begin{pmatrix} 84 \\ 84 \end{pmatrix}.$$

Suppose we want the best-fit line that predicts final exam scores (the second component) from midterm scores (the first component). Explain how to do this and why this method works.

14. Same story as the previous problem, but now with a best-fit *parabola*, $y = c_1x^2 + c_2x + c_3$.

15. Given the data above on Martian heights, finger counts, and weights, suppose we wanted to find the *paraboloid* whose equation has the form $z = c_1x^2 + c_2y^2 + c_3xy + c_4x + c_5y + c_6$ that best fits the data. Describe a matrix-vector equation whose solution would give us the c_i coefficients.

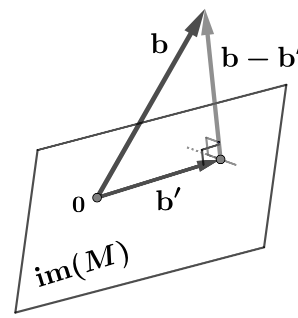
16. In a footnote a few pages back, I implied that an inconsistent system $M\mathbf{x} = \mathbf{b}$ has a *unique* least-squares solution, provided that M's columns are linearly independent. Explain why this is so.

17. Suppose that $M\mathbf{x} = \mathbf{b}$ is a *consistent* linear system, but we decide, just for kicks, to seek a least-squares solution instead, solving $\boldsymbol{M}\mathbf{x} = \mathbf{b}'$, where $\mathbf{b}' = \text{proj}_{\text{im}(M)}\mathbf{b}$. What will we find?

Transpose to the Rescue

In the last section, you learned that when we have an inconsistent linear system $M\mathbf{x} = \mathbf{b}$, we can still find a "least-squares solution" $\mathbf{s}$, the system's closest possible *approximate* solution, by solving $M\mathbf{x} = \mathbf{b}'$, where $\mathbf{b}'$ is the orthogonal projection of $\mathbf{b}$ onto $\mathrm{im}(M)$. Though conceptually clean, finding $\mathbf{s}$ that way can be computationally dirty, since it involves computing the projection $\mathbf{b}'$. Fortunately, some clever algebra will show us how to find $\mathbf{s}$ in a sneaky way – without having to compute $\mathbf{b}'$.

Begin by recalling that $(\mathbf{b} - \mathbf{b}')$ is perpendicular to the subspace $\mathrm{im}(M)$.[*] That is, $(\mathbf{b} - \mathbf{b}')$ is perpendicular to all vectors in $\mathrm{im}(M)$. In particular, it is perpendicular to each column of M. Hence, if M's i^{th} column is $\mathbf{m}_i$, we have



$$\begin{aligned} \mathbf{m}_1 \cdot (\mathbf{b} - \mathbf{b}') &= 0 \\ \mathbf{m}_2 \cdot (\mathbf{b} - \mathbf{b}') &= 0 \\ &\vdots \\ \mathbf{m}_k \cdot (\mathbf{b} - \mathbf{b}') &= 0 \end{aligned}$$

(where $\mathbf{m}_k$ is M's last column). We can wrap all these equations up in a single linear-algebraic package: a matrix-vector product involving M's *transpose*, wherein the $\mathbf{m}_i$ vectors appear as rows instead of columns:

$$\begin{pmatrix} - & \mathbf{m}_1 & - \\ - & \mathbf{m}_2 & - \\ & \vdots & \\ - & \mathbf{m}_k & - \end{pmatrix} \begin{pmatrix} | \\ (\mathbf{b} - \mathbf{b}') \\ | \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.^{\dagger}$$

That is,

$$M^T(\mathbf{b} - \mathbf{b}') = \mathbf{0}.$$

Distributing the M^T, we see that this is equivalent to

$$M^T\mathbf{b}' = M^T\mathbf{b}.$$

Because $\mathbf{b}' = M\mathbf{s}$ (recall that, by definition, $\mathbf{s}$ satisfies $M\mathbf{x} = \mathbf{b}'$), we can rewrite the previous line as

$$M^T M\mathbf{s} = M^T\mathbf{b}.$$

This last equation provides a new characterization of $\mathbf{s}$, our least-squares solution for $M\mathbf{x} = \mathbf{b}$. Namely, $\mathbf{s}$ is the solution to $M^T M\mathbf{x} = M^T\mathbf{b}$. And happily, this characterization of $\mathbf{s}$ doesn't refer explicitly to $\mathbf{b}'$.

Thus, faced with an inconsistent system $M\mathbf{x} = \mathbf{b}$, we can find a least-squares solution $\mathbf{s}$ as follows: Left-multiply both sides of the system by M^T and solve the resulting system, $M^T M\mathbf{x} = M^T\mathbf{b}$. That's it. Algebraic magic! And here it is again, safely packaged in a box for posterity:

> **Theorem.** To find a least-squares solution to an inconsistent system $M\mathbf{x} = \mathbf{b}$, multiply both sides by M^T and then solve the resulting system.

---

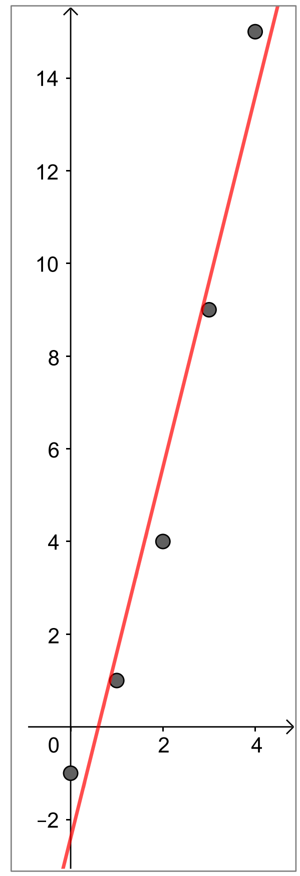
[*] Indeed, this is what it means for $\mathbf{b}'$ to *be* the orthogonal projection of $\mathbf{b}$ onto that subspace.

[†] This follows from the "i^{th}-entry formula for Matrix-Vector multiplication" (See the section in Chapter 3 called "Another Look at the Matrix-Vector Product".)

Example. Find the equation of the best-fit line for the following paired data:

$$(0,-1), (1,1), (2,4), (3,9), \text{ and } (4,15).$$

Solution. By the reasoning we employed in the previous section, we know that the coefficients of an imaginary line $y = c_1 x + c_2$ that somehow passed through all five of these points would be the components of the vector solution of the equation



$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} -1 \\ 1 \\ 4 \\ 9 \\ 15 \end{pmatrix}.$$

But of course, no such line exists. Still, we can find the system's best-fit solution, and as discussed in the previous section, its components will be the coefficients of the best-fit line for the five given points. By this section's theorem, our inconsistent system's best-fit solution must be the solution to the related system

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 4 \\ 9 \\ 15 \end{pmatrix}.$$

After doing the matrix multiplication, this becomes

$$\begin{pmatrix} 30 & 10 \\ 10 & 5 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 96 \\ 28 \end{pmatrix}.$$

We could solve this through Gaussian elimination, but since 2×2 matrices are so easy to invert, we'll solve it by multiplying both sides by the 2×2 matrix's inverse, obtaining

$$\mathbf{x} = \begin{pmatrix} 30 & 10 \\ 10 & 5 \end{pmatrix}^{-1} \begin{pmatrix} 96 \\ 28 \end{pmatrix} = \cdots = \begin{pmatrix} 4 \\ -2.4 \end{pmatrix}.$$

Thus, the best-fit line for the five given points will be $y = 4x - 2.4$. ♦

In Chapter 3's Exercise 27, you saw (and proved) that if M is any matrix, $M^T M$ is a *symmetric* matrix. It follows that when we use this section's theorem to find a least-squares solution to an inconsistent linear system, we'll encounter symmetric matrices along the way. Symmetric matrices have very nice properties that you can learn about in a second linear algebra course. They're all "orthogonally diagonalizable", for example, meaning that for any real symmetric matrix, there's always an *orthonomal* basis relative to which the underlying map has a diagonal representation. Essentially, every real symmetric matrix *is* diagonal… if you tilt your head at the appropriate angle when looking at it.

This book, and our time together, now draws to an end, but others will happily guide you should you care to venture deeper. Apart from one last exercise set, you have now finished this introductory textbook (it's a good feeling, is it not?), and I hereby formally declare you initiated into the dark art of linear algebra.

Exercises.

18. Consider the system $\begin{pmatrix} 6 & 9 \\ 3 & 8 \\ 2 & 10 \end{pmatrix}\mathbf{x} = \begin{pmatrix} 0 \\ 49 \\ 0 \end{pmatrix}$.

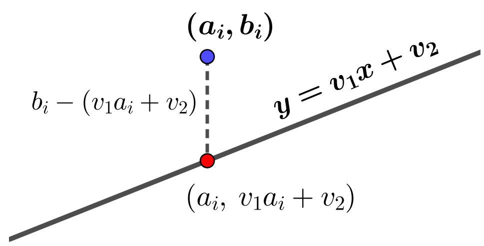
a) Is it consistent? If so, what are its solutions? If not, how do you know?
b) Is there a least-squares solution for the system? If so, find it. If not, why not?

19. Consider the following paired data: $(1,1), (2,3), (3,2), (4,5)$. Make a scatterplot of these points.

a) By hand, find the equation of the best-fit line.
b) Using a computer for the matrix operations, find the best-fit *parabola* (of the form $y = c_1x^2 + c_2x + c_3$).

20. Find the equation of the best-fit *plane* for these points in $\mathbb{R}^3$: $(0,0,0), (0,1,0), (1,0,1), (1,1,3), (1,-1,2)$.

21. In elementary statistics textbooks (and elsewhere), the best-fit line for a set of paired data $(a_1, b_1), \ldots, (a_n, b_n)$ is usually described as the line that minimizes the *sum of the points' squared residuals*. The idea is that relative to any line in the plane, a point's "residual" is, by definition, its signed vertical distance to the line. That is, relative to a line whose equation is $y = v_1x + v_2$, point (a_i, b_i)'s residual is $b_i - (v_1a_i + v_2)$, as indicated in the figure at right, and the sum of the points' squared residuals relative to the line is



$$\sum_{i=1}^{n} \big(b_i - (v_1a_i + v_2)\big)^2.$$

As v_1 and v_2 range over the real numbers (i.e. as the line $y = v_1x + v_2$ covers all possible positions in the plane), this quantity varies. The line that *minimizes* it is, in the statistics approach, defined as the data's best-fit line. One can then – although one doesn't do this in an elementary statistics class – find the best-fit line's coefficients, v_1 and v_2, by solving a multivariable calculus optimization problem.

In this exercise, you'll confirm that our linear algebraic approach to the best-fit line does indeed minimize the sum of squared residuals - and we won't need recourse to multivariable calculus. To set the stage, suppose we have a paired data set $(a_1, b_1), \ldots, (a_n, b_n)$, and we let

$$M = \begin{pmatrix} a_1 & 1 \\ a_2 & 1 \\ \vdots & \vdots \\ a_n & 1 \end{pmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

a) Remind yourself – geometrically - why a least-squares solution $\mathbf{s}$ to an inconsistent system $M\mathbf{x} = \mathbf{b}$ minimizes the quantity $||\mathbf{b} - M\mathbf{x}||$.
b) A vector that minimizes the quantity $||\mathbf{b} - M\mathbf{x}||$ also minimizes $||\mathbf{b} - M\mathbf{x}||^2$. Right?
c) Explain why $||\mathbf{b} - M\mathbf{x}||^2$, a quantity that we know our least-squares solution $\mathbf{s}$ minimizes, is equal to

$$\sum_{i=1}^{n} (b_i - (M\mathbf{x})_i)^2,$$

where $(M\mathbf{x})_i$ is the i^{th} component of $M\mathbf{x}$.
d) If we let $\mathbf{x} = \binom{v_1}{v_2}$ be a variable vector in $\mathbb{R}^2$, explain why $(M\mathbf{x})_i = a_iv_1 + v_2$.
e) Conclude that the least-squares solution $\mathbf{s} = (s_1, s_2)$ minimizes $\sum_{i=1}^{n}\big(b_i - (v_1a_i + v_2)\big)^2$. Hence, s_1 and s_2 are the values of v_1 and v_2 that minimize the sum of squared residuals. It follows that our linear algebra approach agrees, reassuringly, with your statistics teacher about which line is "best".

Selected Answers

To Exercises

Chapter 1

1. False.

3. a) Try division or exponentiation.

b) Try the *average* of two numbers. It helps to invent a symbol such as $\odot$ for the "averaging operation". (Thus, $5 \odot 11 = 8$.) Another fun one to play with: The "rock-paper-scissors operation", which follows the rules of the familiar children's game. (For example, $R \odot S = R$. You'll also need to declare that in the case of "ties", the operation's result is to return the element that was presented twice. For example, $P \odot P = P$.)

5. a) Associativity, commutativity, associativity.

b) Distributive property, scaling by 1 has no effect, ordinary arithmetic.

c) $\mathbf{v} + (4\mathbf{w} + 2\mathbf{v}) = \mathbf{v} + (2\mathbf{v} + 4\mathbf{w}) = (\mathbf{v} + 2\mathbf{v}) + 4\mathbf{w} = (1\mathbf{v} + 2\mathbf{v}) + 4\mathbf{w} = (1+2)\mathbf{v} + 4\mathbf{w} = 3\mathbf{v} + 4\mathbf{w}.$

6. a) No b) Yes c) $2\sqrt{3}$, $\sqrt{3}$

7. a) False b) True c) True d) False. (The vector is in fact 10 units long.)

9. a) $-4\mathbf{i} + 4\mathbf{j} + \mathbf{k}$ b) $3\mathbf{i} + 17\mathbf{j} + 8\mathbf{k}$ c) $\mathbf{i} + \mathbf{j} + \mathbf{k}$ d) $\sqrt{11}$ e) $\sqrt{33}$ f) $\frac{1}{\sqrt{30}}(-\mathbf{i} + 5\mathbf{j} + 2\mathbf{k})$ g) 1 h) 1

10. 8

11. a) Extend a cube's edges indefinitely. You'll find some skew lines among them.

b) One way: Let l be the first randomly chosen line. Set up your axes so that l is the x-axis. Unless the second randomly chosen line, m, happens to be parallel to the xy-plane (which is unlikely, given m's randomness), it will pierce the xy-plane at one - and only one - point. Unless that point of intersection happens to be on the x-axis (another unlikely event), lines l and m will never meet. Thus, for two random lines in space, parallelism and intersection are both exceptions to the rule. Skew lines are the norm.

12. a) (iii) Vertical lines *cannot* be put in the form $y = n + mx$. They can, however, still be expressed in the form $ax + bx = c$. Namely, the vertical line crossing the horizontal axis at k will have the equation $x = k$, which, spelled out in gory detail to show that it conforms to the usual pattern, is $1x + 0y = k$.

c) (ii) $w = 6 + 2x + y + 7z$.

d) (ii) an ordinary 2-dimensional plane.

13. $\mathbf{v} \cdot \mathbf{w} = 16$, $\mathbf{a} \cdot \mathbf{b} = -12$

14. If the dot product is positive, the angle between the vectors is acute. If negative, the angle is obtuse.

15. a) $\approx 77.3°$ b) $\approx 119.7°$ c) $\approx 61.2°$

16. Yes.

17. Check your candidate by confirming that its dot product with $\mathbf{v}$ is zero.

20. Thou shalt not cancel under these circumstances! (Simple counterexample: $\mathbf{i} \cdot \mathbf{j} = \mathbf{i} \cdot \mathbf{k}$, but $\mathbf{j} \neq \mathbf{k}$.)

21. Neither $\mathbf{a} \cdot (\mathbf{b} \cdot \mathbf{c})$ nor $(\mathbf{a} \cdot \mathbf{b}) \cdot \mathbf{c}$ is even *defined*, since you can't take the dot product of a vector and a scalar.

Chapter 2

1. Parts e, f, i are false. The rest are true. **2.** Yes to part B. No to the others.

3. No, since the zero vector would be a linear combination of all the other vectors in the set: just let their scalar coefficients all be zero.

4. a) The line $y = 2x$ b) $\mathbb{R}^2$ c) $\mathbb{R}^2$ d) The line $y = x/3$ e) The origin.

6. a) $\mathbf{u} = \frac{31}{8}\mathbf{v} + \frac{3}{8}\mathbf{w}$ b) $\mathbf{d} = \mathbf{a} + \mathbf{b} = 2\mathbf{a} + \frac{1}{2}\mathbf{c} = 3\mathbf{a} - \mathbf{b} + \mathbf{c} = 2\mathbf{b} - \frac{1}{2}\mathbf{c}$ (among other possibilities)

7. Parts c, d, g, j, l, m are true. The rest are false.

8. a) the origin, lines through the origin, and all of $\mathbb{R}^2$.

b) the origin, lines through the origin, planes through the origin, and all of $\mathbb{R}^3$.

c) the origin; lines, planes, and 3-dimensional hyperplanes through the origin; all of $\mathbb{R}^4$.

d) the origin; lines, planes, 3-dimensional and 4-dimensional hyperplanes through the origin; all of $\mathbb{R}^4$.

9. Yes to a,b,and c. No to d.

11. Both conditions are necessary. In $\mathbb{R}^2$, the first quadrant is closed under vector addition, but not under scalar multiplication. The axes (considered as a pair) are closed under scalar multiplication, but not vector addition. These examples (others could be given) demonstrate that neither closure condition implies the other.

12. Closed under addition, multiplication, and differentiation, but *not* under integration.

13. Only options a, b, d, e are bases for $\mathbb{R}^2$.

14. The graphs of equations a, c, f, h are subspaces of $\mathbb{R}^3$.

15. Point A corresponds to $\mathbf{v} + 2\mathbf{w_1} - \mathbf{w_2}$.
Point B corresponds to $\mathbf{v} - \mathbf{w_1} + 2\mathbf{w_2}$.
Point C corresponds to $\mathbf{v} + 3\mathbf{w_1} + 3\mathbf{w_2}$.

16. a) Line in $\mathbb{R}^2$. Affine space. b) Line in $\mathbb{R}^3$. Affine space. c) Line in $\mathbb{R}^4$. Affine space. d) Plane in $\mathbb{R}^3$. Subspace. e) Plane in $\mathbb{R}^3$. Affine space. f) Three-dimensional hyperplane in $\mathbb{R}^4$. Subspace. g) Plane in $\mathbb{R}^5$. Affine space. h) Line in $\mathbb{R}^3$. Subspace. (Did you forget to check for linear independence?) i) Plane in $\mathbb{R}^3$. Affine space.

17. When we represent a line by an expression of the form $\mathbf{v} + t\mathbf{w}$, we have infinitely many different choices for $\mathbf{v}$ (our "anchor point") and $\mathbf{w}$ (a parallel vector). Different choices will yield different parametric representations. The same sort of thing applies to parametric representations of planes and hyperplanes.

18. Answers could vary, since parametric representations are not unique, but the most obvious possibilities are:

a) $x = 3 + 4t, \quad y = 1 - t$ b) $x = 3 + 4t, \quad y = 1 - t, \quad z = 2 + t$
c) $x = 3 + t - 5s, \quad y = 1 - 2t + 2s, \quad z = 2 - t - s.$

19. a) The vectors are linearly independent, so their span is a plane in $\mathbb{R}^3$ through the origin, given by
$x = 2t + 6s, \quad y = t + s, \quad z = t.$
b) The vectors are linearly **de**pendent, so their span is a mere line in $\mathbb{R}^3$ through the origin, given by
$x = 2t, \quad y = t, \quad z = t.$

20. a) 4 b) $m + 1$
c) There are exceptions. [Ex: *Infinitely many* planes pass through $(0,0,0)$, $(1,0,0)$, and $(2,0,0)$.]

Chapter 3

1. Part A is true, B is false.

2. a) $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. It sends (2,3) to (-2,3). b) $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. It sends (2,3) to (2,-3).

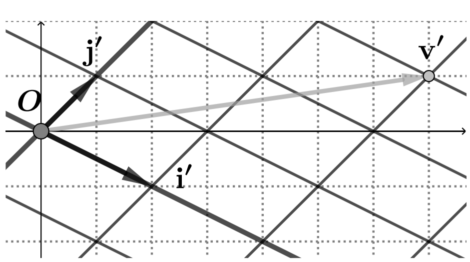
c) This is not a *linear* map (cf. Exercise 1a), so it has no matrix. [It sends (2,3) to (2,-1).]

d) $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. It sends (2,3) to $(2\cos\theta - 3\sin\theta,\ 2\sin\theta + 3\cos\theta)$.

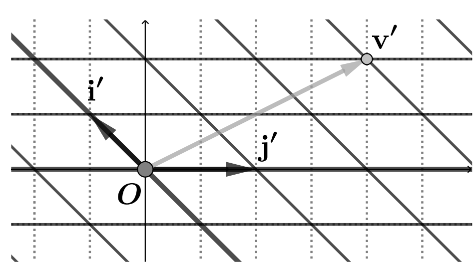
e) $\begin{pmatrix} \sqrt{3}/2 & -1/2 \\ 1/2 & \sqrt{3}/2 \end{pmatrix}$. It sends (2,3) to $\left(\sqrt{3} - \frac{3}{2},\ 1 + \frac{3}{2}\sqrt{3}\right)$. f) $\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$. It sends (2,3) to $\left(\frac{5}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

g) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. It doesn't move (2,3).

3. b)



c)



4. a) $\begin{pmatrix}1 & .5\\0 & 1\end{pmatrix}$ b) $(21, 22)$ c) $(-1, 22)$ d) It preserves their areas since it preserves their bases and heights.

e) The sheep – or any plane region – is made up of infinitesimal squares. Since squares' areas are preserved by the shear, the sheared sheep's area remains what it was before the shear.

6. The zero matrix crushes all of $\mathbb{R}^n$ into the origin.

7. a) $\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & -1\end{pmatrix}$ b) $\begin{pmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{pmatrix}$ c) $\begin{pmatrix}\cos\theta & -\sin\theta & 0\\\sin\theta & \cos\theta & 0\\0 & 0 & 1\end{pmatrix}$ d) $\begin{pmatrix}1 & 0 & 0.5\\0 & 1 & 0.5\\0 & 0 & 1\end{pmatrix}$ e) $\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}$

8. a) $\begin{pmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 0\\0 & 0 & 0 & \mathbf{-1}\end{pmatrix}$ b) $\begin{pmatrix}1 & 0 & \cdots & 0 & 0\\0 & 1 & \cdots & 0 & 0\\\vdots & \vdots & \ddots & \vdots & \vdots\\0 & 0 & \cdots & 1 & 0\\0 & 0 & \cdots & 0 & \mathbf{-1}\end{pmatrix}$ c) $\begin{pmatrix}1 & 0 & \cdots & 0 & 0\\0 & 1 & \cdots & 0 & 0\\\vdots & \vdots & \ddots & \vdots & \vdots\\0 & 0 & \cdots & 1 & 0\\0 & 0 & \cdots & 0 & 1\end{pmatrix}$

9. a) $\begin{pmatrix}23\\34\end{pmatrix}$ b) $\begin{pmatrix}12\\-9\\5\end{pmatrix}$ c) $\begin{pmatrix}-3\\-1\\4\\-4\end{pmatrix}$

10. The vector must have n entries. **11.** An $m \times n$ matrix determines a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$.

12. a) It maps $\mathbb{R}^2$ onto a plane in $\mathbb{R}^3$. b) It maps $\mathbb{R}^3$ onto $\mathbb{R}^2$. c) It maps $\mathbb{R}^3$ onto a single *line* in $\mathbb{R}^2$. d) It maps $\mathbb{R}^2$ onto $\mathbb{R}^3$'s origin. e) It maps $\mathbb{R}^3$ onto a 3-dimensional hyperplane in $\mathbb{R}^4$. f) It maps $\mathbb{R}^4$ onto $\mathbb{R}^2$.

14. a) The map represented by $2A$ does what A does, while also dilating the plane by a factor of 2.

15. a) $\begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}$ b) $\begin{pmatrix}0\\1\\1\\0\\0\end{pmatrix}$ c) $\begin{pmatrix}0\\0\\0\\3\\0\end{pmatrix}$ d) $\begin{pmatrix}-5\\0\\0\\3\\0\end{pmatrix}$ **16.** a) $\begin{pmatrix}-3 & 5\\-1 & 5\end{pmatrix}$ b) $\begin{pmatrix}4 & 2\\1 & -2\end{pmatrix}$ c) $\begin{pmatrix}0 & 6 & 0\\1 & 0 & 3\\0 & 6 & 0\end{pmatrix}$

18. Think geometrically.

19. b) AB is a 5×3 matrix, which represents a linear map from $\mathbb{R}^3$ to $\mathbb{R}^5$.
c) $BCAB$ is a 2×3 matrix, which represents a linear map $\mathbb{R}^3$ to $\mathbb{R}^2$.

20. $MN = \begin{pmatrix}4 & -1 & 1\\6 & -9 & 3\\-3 & 7 & -2\end{pmatrix}$, $NM = \begin{pmatrix}-9 & 5\\4 & 2\end{pmatrix}$, M^2 and N^2 are undefined; $NMN = \begin{pmatrix}-13 & 32 & -9\\10 & -10 & 4\end{pmatrix}$.

21. a) $-\frac{1}{2}\begin{pmatrix}1 & \sqrt{3}\\\sqrt{3} & -1\end{pmatrix}$ b) $-\frac{1}{2}\begin{pmatrix}3 + 8\sqrt{3}\\-8 + 3\sqrt{3}\end{pmatrix}$ c) $M^2 = I$, since M^2 fixes both standard basis vectors.

23. a) A^{-1} is a 90° *clockwise* rotation about the origin, so $A^{-1} = \begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}$. You should find that $A^{-1}A = I$.

b) Suppose A is a nonsquare $m \times n$ matrix, and C is some theoretical "candidate" auditioning to be A's inverse. Clearly, matrix C must have n rows (so we can form the product AC) and m columns (so that we can form CA). Hence, C would have to be an $n \times m$ matrix. But in that case, AC would be an $m \times m$ matrix, while CA would be an $n \times n$ matrix. Thus, it is impossible to find a matrix C such that $AC = CA$. It follows that nonsquare matrices can't have inverses. Only *square* matrices can have inverses.

c) Suppose B and C are inverses of A. Since B is an inverse of A, we know that $AB = I$. Left-multiplying both sides by C yields $CAB = C$, or, using the associativity of matrix multiplication, $(CA)B = C$. But since C is an inverse of A, this becomes $IB = C$, or equivalently, $B = C$, as claimed.

d) To be invertible, a matrix, and hence the map it represents, must be "one-to-one" (i.e. must always take distinct points in the domain to distinct points in the range). The zero map fails this test spectacularly.

e) For our formal proof that AB's inverse is $B^{-1}A^{-1}$, we must multiply these two matrices (in both orders) and confirm that we get I each time. The associativity of matrix multiplication does all the work for us:

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I.$$

I'll leave it to you to confirm that $(AB)(B^{-1}A^{-1})$ is also I.

f) In this case, you should find that the matrix's inverse is equal to the matrix itself:

$$-\frac{1}{2}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}.$$

g) Right-multiplying cA by $c^{-1}A^{-1}$ yields $(cA)(c^{-1}A^{-1}) = (cc^{-1})AA^{-1} = 1I = \boldsymbol{I}$. (The first equals sign is justified by the result you proved in Exercise 22.) Left-multiplying cA by $c^{-1}A^{-1}$ also yields I, as you can verify. Hence, $(cA)^{-1} = c^{-1}A^{-1}$, as claimed.

24. a) Lighter b) Darker

c) Reflected over the line extending from the top left corner (i.e. the pixel in row 1, column 1) through all the other pixels whose row and column numbers are equal.

d) Photo negative e) Reflected over the vertical line that divides the image in half.

f) N/A: Those matrices can't be multiplied in that order.

g) Reflected over the horizontal line that divides the image in half.

27. a) No. c) MM^T is not only square, but symmetric. Same story for M^TM.

Chapter 4

1. a) $\begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ b) $\begin{pmatrix} 0 & 2 & 4 \\ 1 & 3 & 5 \\ 3 & 7 & 7 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -5 \\ -2 \\ 6 \end{pmatrix}$ c) $\begin{pmatrix} 1 & -2 & -1 \\ 3 & -6 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ d) $\begin{pmatrix} 1 & 1 \\ -1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$

2. a) 1 b) 1 c) infinitely many d) 0

3. a) $\begin{aligned} x + 2y &= 5 \\ 3x + 4y &= 6 \end{aligned}$ b) $\begin{aligned} 3r + s + 4t &= 2 \\ r + 5s + 9t &= 6 \end{aligned}$ c) $\begin{aligned} 2x_1 + 7x_2 &= 4 \\ x_1 + 8x_2 &= 5 \\ 2x_1 + 8x_2 &= 9 \end{aligned}$

8. a) $\left(\begin{array}{cc|c} 1 & 3 & 5 \\ 2 & -1 & 2 \end{array}\right)$ b) $\left(\begin{array}{ccc|c} 0 & 2 & 4 & -5 \\ 1 & 3 & 5 & -2 \\ 3 & 7 & 7 & 6 \end{array}\right)$ c) $\left(\begin{array}{ccc|c} 1 & -2 & -1 & 3 \\ 3 & -6 & -2 & 2 \end{array}\right)$ d) $\left(\begin{array}{cc|c} 1 & 1 & 2 \\ -1 & 1 & 2 \\ 2 & 1 & 1 \end{array}\right)$

12. a) $\left(\begin{array}{cc|c} 1 & 2 & 3 \\ 2 & 3 & 5 \end{array}\right)$ b) $\left(\begin{array}{ccc|c} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{array}\right)$ c) $\left(\begin{array}{cccc|c} 2 & 3 & 25 & 1 & 2 \\ 0 & 1 & 8 & -3 & 7 \\ 0 & 0 & 6 & 0 & 15 \end{array}\right)$ d) $\left(\begin{array}{cccc|c} 1 & -2 & 3 & -4 & 5 \\ 0 & 14 & -7 & 14 & -32 \end{array}\right)$

13. a) $\left(\begin{array}{cc|c} 1 & -3 & 1 \\ 0 & -7 & 6 \\ 0 & 11 & 2 \\ 0 & -1 & 1 \end{array}\right)$ b) $\left(\begin{array}{ccc|c} 1 & -4 & 5 & 0 \\ 0 & \mathbf{1} & -2 & 6 \\ 0 & 1/2 & -3 & 4 \end{array}\right) \begin{array}{c} +4R_2 \\ \\ -(1/2)R_2 \end{array} \left(\begin{array}{ccc|c} 1 & 0 & -3 & 24 \\ 0 & 1 & -2 & 6 \\ 0 & 0 & -2 & 1 \end{array}\right)$

14. a) $\left(\begin{array}{cc|c} 2 & 3 & 0 \\ 4 & 5 & 0 \end{array}\right)$ is equivalent to $\left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right)$, so the original system's unique solution is $(0,0)$.

b) $\left(\begin{array}{cc|c} 4 & 3 & 2 \\ 7 & 5 & 3 \end{array}\right)$ is equivalent to $\left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 2 \end{array}\right)$, so the original system's unique solution is $(-1, 2)$.

c) $\left(\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 2 & 4 & 7 & 2 \\ 3 & 7 & 11 & 8 \end{array}\right)$ is equivalent to $\left(\begin{array}{ccc|c} 1 & 0 & 0 & -9 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 0 \end{array}\right)$, so the original system's unique solution is $(-9, 5, 0)$.

d) $\left(\begin{array}{ccc|c} 1 & 2 & 3 & 8 \\ 1 & 3 & 3 & 9 \\ 1 & 2 & 4 & 10 \end{array}\right)$ is equivalent to $\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right)$, so the original system's unique solution is $(1, 2, 1)$.

15. The row $(0 \quad 0 \quad \cdots \quad 0 \mid c)$ corresponds to the equation $0 = c$, a condition that *cannot* be satisfied, since $c \neq 0$. Since the a solution to a system is one that satisfies all its equations, any system containing an unsatisfiable condition like this one has **no solutions**.

16. a) neither b) rref c) neither d) row echelon form

17. a) $\big((-t-3),(2t+2),t\big)$ for all real t; line in $\mathbb{R}^3$. b) $\big((5-3t+9s),t,(1+5s),s\big)$ for all real t,s; plane in $\mathbb{R}^4$.

c) No solution. d) $\left(\frac{7}{4}t,\ \left(\frac{1}{2}-\frac{1}{4}t\right),\ t\right)$ for real t; a line in $\mathbb{R}^3$. e) $(3,\ 4,-2)$; a point in $\mathbb{R}^3$

f) $\big((-2t-s+u),\ t,\ (1+s-u),\ (2-2s+u),\ s,\ u\big)$, for all real s,t,u; a 3-dimensional hyperplane in $\mathbb{R}^6$.

18. a) $\begin{pmatrix}x\\y\\z\\w\end{pmatrix}=\begin{pmatrix}0\\2-s-2t\\s\\t\end{pmatrix}=\begin{pmatrix}0\\2\\1\\1\end{pmatrix}+s\begin{pmatrix}0\\-1\\1\\0\end{pmatrix}+t\begin{pmatrix}0\\-2\\0\\1\end{pmatrix}$. A plane in $\mathbb{R}^4$. b) no solutions

c) $\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1+2s-3t\\s\\t\end{pmatrix}=\begin{pmatrix}1\\0\\0\end{pmatrix}+s\begin{pmatrix}2\\1\\0\end{pmatrix}+t\begin{pmatrix}-3\\0\\1\end{pmatrix}$. A plane in $\mathbb{R}^3$. d) $\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}$. The origin of $\mathbb{R}^2$.

e) $\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix}=\begin{pmatrix}s\\7-3t-2u\\t\\8-5u\\u\end{pmatrix}=\begin{pmatrix}0\\7\\0\\8\\0\end{pmatrix}+s\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}+t\begin{pmatrix}0\\-3\\1\\0\\0\end{pmatrix}+u\begin{pmatrix}0\\-2\\0\\-5\\1\end{pmatrix}$. A 3-dimensional hyperplane in $\mathbb{R}^5$.

f) $\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\\x_6\\x_7\end{pmatrix}=\begin{pmatrix}8+2t_1+4t_3-6t_5\\t_1\\9-3t_3-5t_4-7t_5\\t_2\\t_3\\t_4\\t_5\end{pmatrix}=\begin{pmatrix}8\\0\\9\\0\\0\\0\\0\end{pmatrix}+t_1\begin{pmatrix}2\\1\\0\\0\\0\\0\\0\end{pmatrix}+t_2\begin{pmatrix}0\\0\\0\\1\\0\\0\\0\end{pmatrix}+t_3\begin{pmatrix}4\\0\\-3\\0\\1\\0\\0\end{pmatrix}+t_4\begin{pmatrix}0\\0\\-5\\0\\0\\1\\0\end{pmatrix}+t_5\begin{pmatrix}-6\\0\\-7\\0\\0\\0\\1\end{pmatrix}$.

(A 5-dimensional hyperplane in $\mathbb{R}^7$.)

19. No. The system $A\mathbf{x}=\mathbf{v}$ has no solution as you can show with Gaussian elimination. **20.** No.

21b. *i*) Linearly independent.

ii) Linearly **de**pendent. There are infinitely many "dependencies". One is $-2\mathbf{v_1}-3\mathbf{v_2}+\mathbf{v_3}=\mathbf{0}$, where $\mathbf{v_1},\mathbf{v_2},\mathbf{v_3}$ are the three given vectors (in the same order given in the problem).

iii) Linearly dependent. (As any four vectors in $\mathbb{R}^3$ must be!)

22 b) Let A be the matrix whose columns are the vectors in Exercise 21b, *ii*. To see if $\mathbf{v}$ lies in those vectors' span, we solve $A\mathbf{x}=\mathbf{v}$. A quick computer check shows that this equation has no solution, so $\mathbf{v}$ does *not* lie in their span. In contrast, $\mathbf{w}$ *does* lie in their span and can be expressed as a linear combination of A's columns in infinitely many different ways. In particular, if we call the columns $\mathbf{a_1},\mathbf{a_2},\mathbf{a_3}$, then for any real t, we have

$$\mathbf{w}=(22-2t)\mathbf{a_1}+(25-3t)\mathbf{a_2}+t\mathbf{a_3}.$$

24. a) $y=\frac{7}{6}x^2-\frac{1}{2}x-\frac{2}{3}$ b) $y=-2x^3-x^2+2x+1$ c) No quadratics. Infinitely many quartics.

25. If we know how many cars pass along any one block, then the number of cars on the other blocks are determined. If we let t be the number of cars on Teerts Street, for example, then the number of cars on Yaw, Daor, and Enal will be, respectively, $y=150-t$, $d=250-t$, and $e=270-t$. After playing around with these expressions, (or similar ones we'd obtain by using the number of cars on a *different* street – not Teerts – as our parameter), we see that at any given moment, $0\le t\le 150$, $0\le y\le 150$, $100\le d\le 250$, and $120\le e\le 270$.

26. If c is the number of chicks, then the number of roosters will be $r=(4/3)c-100$, and the number of hens will be $h=200-(7/3)c$. Four values of c will make all of these three numbers positive integers, so the problem has four different solutions: $(75+3t)$ chicks, $(4t)$ roosters, $(25-7t)$ hens, where t can be $0,1,2$, or 3.

28. $A^{-1}=\begin{pmatrix}7&-3&-3\\-1&1&0\\-1&0&1\end{pmatrix}$, $B^{-1}=\frac{1}{6}\begin{pmatrix}2&6&-2\\-3&3&0\\1&-3&2\end{pmatrix}$, $C^{-1}=\frac{1}{30}\begin{pmatrix}16&-6&-2\\20&0&-10\\7&3&1\end{pmatrix}$.

29. c) $A^{-1}=A$, $B^{-1}=\begin{pmatrix}2&-3\\-1&2\end{pmatrix}$, $C^{-1}=-\frac{1}{5}\begin{pmatrix}2&-3\\-7&8\end{pmatrix}$, D is not invertible.

30. a) The matrix would have an entire column of zeros, meaning that the map kills a standard basis vector, collapsing a dimension. Dimensional collapse means the map can't be one-to-one, so the matrix can't be invertible.

b) The map corresponding to this matrix just stretches the various standard basis vectors, while maintaining (or reversing) their directions. Stretches – by any nonzero factor – are obviously invertible, and the inverse matrix would just "undo" all these stretches.

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

d) It will be another diagonal matrix, whose entries are the reciprocals of those in the original matrix.

e) Hint: If D is a diagonal matrix, think about what it does to each standard basis vector. From there, think about what D^2, D^3, etc. would do to each standard basis vector.

31. The preceding section's footnote shows that A ⇔ B ⇔ C. It's geometrically clear that any n linearly independent vectors in $\mathbb{R}^n$ span $\mathbb{R}^n$ and that conversely, for any n vectors to span $\mathbb{R}^n$, they must be linearly independent. Hence, D ⇔ E. As discussed in Chapter 2, a set of n vectors in $\mathbb{R}^n$ is a basis if and only if the vectors are linearly independent and span the whole space. Thus, statements D and E (which we've just noted stand or fall together) are also equivalent to statement F. We've now established two separate "islands" of equivalent statements: A ⇔ B⇔ C and D ⇔ E ⇔ F. It remains only to unite the islands. First, a one-way bridge from the second island to the first (extending from F to A): If a matrix's columns are a basis for $\mathbb{R}^n$, then the matrix transforms the standard grid into a new "clean grid" (to use Chapter 2's language), and thus it will clearly map separate input points to separate output points. In other words, the map/matrix is one-to-one, and is therefore invertible. Thus, F ⇒ A. Now we'll build our one-way bridge from the first island back to the second, which will do from A to D: If the map/matrix is invertible, then it can't collapse any dimensions. (Otherwise many input points would get mapped to the same output point and the map wouldn't be one-to-one.) But to preserve all dimensions, its columns must be linearly independent. Thus, A ⇒ D. Our two islands are now fully united. We can get from any of the five statements to any other. They all stand or fall together.

32. No. **33.** a) It is 3 times $\text{rref}(A)$'s 6th column plus twice $\text{rref}(A)$'s 7th column. b) No. (By the IMT: Exercise 31.)

34. a) Image: $\mathbb{R}^2$, Kernel: $\mathbf{0}$. (Rank = 2, Nullity = 0.) b) Image: $\mathbf{0}$, Kernel: $\mathbb{R}^3$. (Rank = 0, Nullity = 3.)
c) Image: $y = x$, Kernel: $y = -x$. (Rank = Nullity = 1) d) Image: $\mathbb{R}^4$, Kernel: $\mathbf{0}$. (Rank = 4, Nullity = 0)
e) Image: $\mathbb{R}^n$, Kernel: $\mathbf{0}$. (Rank = n, Nullity = 0) f) Image: $\mathbb{R}$, Kernel: the yz-plane. (Rank = 1, Nullity = 2)

35. a) To use the suggested example, a rotation matrix is invertible, so by the Invertible Matrix Theorem, its rref is I, which obviously no longer rotates points. Hence, row operations have destroyed the original matrix's geometric effect.

b) Try $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. Its image is the line $y = 3x$. Scaling row 1 by 3 yields $\begin{pmatrix} 3 & 6 \\ 3 & 6 \end{pmatrix}$, whose image is the line $y = x$.

c) In any matrix M, if we weed out columns that depend linearly on their predecessors, we get a basis for $\text{im}(M)$. Suppose we do row operations on matrix A, obtaining matrix B, and then we weed both matrices' columns. Since B is a "descendant" of A via row operations, which preserve linear dependencies among the columns, we'll weed the columns in the same positions of the two matrices (e.g. if we weed columns 1, 3, and 7 in A, we'll weed columns 1, 3, and 7 in B.) Consequently, the bases of $\text{im}(A)$ and $\text{im}(B)$ must have the same number of vectors. But, by definition, the number of vectors in a subspace's basis is the *dimension* of that subspace. Thus, $\text{im}(A)$ and $\text{im}(B)$ have the same dimension. In other words, A and B have the same rank.

d) The kernel – unlike the image – is a set of solutions to an equation ($A\boldsymbol{x} = \mathbf{0}$), and solutions are preserved by row operations.

36. a) Image: $\mathbb{R}^2$, Kernel $\mathbf{0}$. (Rank = 2, Nullity = 0) b) Image: $y = 2x$, Kernel: $t(-2\mathbf{i} + \mathbf{j})$. (Rank = 1, Nullity = 1)
c) Im: $(\mathbf{i} + 4\mathbf{j} + 7\mathbf{k})t + (2\mathbf{i} + 5\mathbf{j} + 8\mathbf{k})s$ (plane in $\mathbb{R}^3$), Ker: $(\mathbf{i} - 2\mathbf{j} + \mathbf{k})t$, a line in $\mathbb{R}^3$. (Rank = 2, Nullity = 1)
d) Im: $\mathbb{R}^2$, Ker: $(\mathbf{i} - 2\mathbf{j} + \mathbf{k})t$, a line in $\mathbb{R}^3$. (Rank = 2, Nullity = 1)
e) Im: span of the columns, a plane in $\mathbb{R}^4$. Ker: $\mathbf{0}$, the origin in $\mathbb{R}^2$. (Rank = 2, Nullity = 0)
f) Im: $\mathbb{R}^4$, Ker: $\mathbf{0}$. (Rank = 4, Nullity = 0)

37. a) True, b) False, c) True, d) False, e) True, f) False (they need not be independent of one another!)

39. In that case, the matrix's nullity is 2, so by the rank-nullity theorem, its rank must be 1. Its image must therefore be a line. That can only happen if *the matrix's columns are all scalar multiples of one another*.

40. The equivalence of A-F was established in Exercise 31. Next, H and I are equivalent by the definition of *rank*. These, in turn, are equivalent to G by the rank-nullity theorem. Thus, we have a new "island" of three equivalent statements: G,H,I. To establish a bridge to the Invertible Matrix Theorem's "mainland" of statements A-F, we note that a mainland statement (E: A's columns span $\mathbb{R}^n$) is clearly equivalent to an island one (H: $\mathrm{im}(A) = \mathbb{R}^n$).

41. a) $\mathrm{im}\, A$: The line $y = x/2$. $\ker A$: The line $y = -x/3$. (i.e. $\ker A = t\begin{pmatrix} 3 \\ -1 \end{pmatrix}$ for all real t.)

b) $\begin{pmatrix} 2 \\ 0 \end{pmatrix} + t\begin{pmatrix} 3 \\ -1 \end{pmatrix}$ for all real t. c) $\begin{pmatrix} 0 \\ -1 \end{pmatrix} + t\begin{pmatrix} 3 \\ -1 \end{pmatrix}$ for all real t. d) All points on the line are mapped to $(14, 7)$.

Chapter 5

2. False: B,D,E. The others are true.

3. a) To move the standard basis vectors into this position requires "crossing the axes" that they determine.

4. a) A square matrix whose entries *above* the main diagonal are all zeros.

5. a) 16 vertices, 32 edges, 24 faces. b) 8 three-dimensional cells.

6. Since $A^{-1}A = I$, we have $\det(A^{-1}A) = \det I$. By property 2, this is $\det(A^{-1})\det(A) = \det I$. By Exercise 2F, the right-hand side is 1. Thus, $\det(A^{-1}) = 1/\det A$, as claimed. The result follows.

7. a) $\det A = 1$, $\det B = -5$, $\det C = 0$, $\det D = -1$, $\det E = 60$. b) All but C

c) A determinant of zero indicates dimensional collapse. When that occurs, multiple "input" points get crushed into the same "output". (i.e. the map is not one-to-one.) As a result, given a particular output, there's no way of saying which input it came from. Hence, the map/matrix can't be inverted.

d) A and D e) B and D

f) $\det(A^{-1}) = 1$, $\det(B^{-1}) = -1/5$, $\det(D^{-1}) = -1$, $\det(E^{-1}) = 1/60$.

g) $B^2 = \begin{pmatrix} 85 & 30 \\ 70 & 25 \end{pmatrix}$, $\det(B^2) = 25$. h) $\det(AB) = -5$

8. Exercise 7B shows that $\det A = 0 \Rightarrow$ not invertible. Moreover, if $\det A \neq 0$, there's no dimensional collapse, so $\mathrm{im}(A) = \mathbb{R}^n$, which – by the existing Invertible Matrix Theorem – implies that A is invertible. Thus, having a nonzero determinant and being invertible are logically equivalent, so we may adjoin statement I to the IMT's list.

10. If $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, then $A^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$

11. $\det(D^{1000}) = \det(DD\cdots D) = \det(D)\det(D)\cdots\det(D) = (\det(D))^{1000} = 1$.

13. a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ c) $\begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ **14.** a) -1 b) 4 c) 1 **15.** $\det A = 5/3$.

16. a) 18 b) 1 c) -96 d) 24 e) -8 f) -15 g) 9 h) 0 **17.** Hint: Look at the first two columns.

19. First, the three new statements (K, L, M) are clearly equivalent *to one another*, as they concern n vectors in $\mathbb{R}^n$. We'll now join this new "island" to the IVT mainland by showing that statements A and K are equivalent:

Statement A: A is invertible $\Leftrightarrow \det(A) \neq 0$ (by the existing IVT)
$\Leftrightarrow \det(A^T) \neq 0$ (since $\det(A) = \det(A^T)$)
$\Leftrightarrow$ the columns of A^T span $\mathbb{R}^n$ (by the existing IVT)
$\Leftrightarrow$ the rows of A span $\mathbb{R}^n$: **Statement L** (by definition of the transpose).

20. Yes. It's easy to see that the matrix's columns are linearly independent, so by the IMT, the matrix is invertible. As such, its rows span $\mathbb{R}^3$ by the IMT.

21. A zero column would indicate that a dimension is crushed, so the matrix's determinant would be zero. A matrix with a zero *row* would have a determinant of zero as well, since every square matrix has the same determinant as its transpose, and the transpose of a matrix with a zero row is a matrix with a zero column.

Chapter 6

1. a) $\begin{pmatrix} 3 & -1 \\ -2 & 2 \end{pmatrix}$ b) $\frac{1}{4}\begin{pmatrix} 2 & 1 \\ 2 & 3 \end{pmatrix}$ c) $\begin{pmatrix} 7 \\ 13 \end{pmatrix}$ d) $\begin{pmatrix} 12 \\ 8 \end{pmatrix}$

e) A vector has this property if and only if its $\mathbf{a}_2$-coordinate is 2/3 of its $\mathbf{a}_1$-coordinate. (Or equivalently, if its $\mathbf{b}_2$-coordinate is 3/2 of its $\mathbf{b}_1$-coordinate.) The set of all such vectors lie on a line through the origin.

2. b) $[\mathbf{b}_1]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $[\mathbf{b}_2]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. c) $B = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}$ d) $B^{-1} = \frac{1}{4}\begin{pmatrix} 2 & -1 \\ -2 & 3 \end{pmatrix}$ e) $\begin{pmatrix} 9/4 \\ -11/4 \end{pmatrix}$ f) $\begin{pmatrix} 11 \\ 6 \end{pmatrix}$

3. a) Yes. b) $\frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ -2 & 2 & 0 \\ 1 & -1 & 1 \end{pmatrix}$ c) $\begin{pmatrix} -3/2 \\ -4 \\ 9/2 \end{pmatrix}$ **4.** b) It will be a 3×3 matrix. c) $\begin{pmatrix} 2 & 2 & -1 \\ 0 & -2 & 0 \\ 1 & 1 & 0 \end{pmatrix}$

5. a) In the first case (basis $\mathcal{E}$), the square root of the squared coordinates is $\sqrt{5}$.
In the second case (basis $\mathcal{A}$), the square root of the squared coordinates is $\sqrt{2}$.
In the third case (basis $\mathcal{B}$), the square root of the squared coordinates is $\sqrt{13}$.
Only in the first case did the square root of the squared coordinates yield $\|\mathbf{v}\|$.

b) $[\mathbf{i}]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $[\mathbf{j}]_{\mathcal{B}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Hence, summing the products of the corresponding $\mathcal{B}$-coordinates yields -1.

6. a) $[R]_{\mathcal{B}} = C^{-1}[R]_{\mathcal{E}}C$ b) $[R]_{\mathcal{B}} = \frac{1}{4}\begin{pmatrix} 7 & 5 \\ -13 & -7 \end{pmatrix}$.

c) $\mathcal{B}$-coordinates: $(-17/4, 27/4)$. Standard coordinates: $(-6, 5)$. I'll leave the sketch to you. d) $(-3, 5)$.

8. $\frac{1}{5}\begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix}$ **9.** $\frac{1}{7}\begin{pmatrix} 6 & 3 \\ 2 & 1 \end{pmatrix}$

10. a) If A is similar to B, then $CAC^{-1} = B$ for some C. Right-multiply both sides by C^{-1}, then left-multiply both sides by C. The result: $A = C^{-1}BC$. If we let $D = C^{-1}$, this becomes $A = DBD^{-1}$. Thus, B is similar to A.

b) Hint: This is similar in spirit to the previous part.

d) Similar matrices represent the same linear map. The volume-expanding factor of this map is, of course, independent of its matrix representations.

e) Same idea as Part D. Think about the underlying map, and the geometric meaning of rank.

f) If A were an $m \times n$ matrix and similar to B, we'd have $CAC^{-1} = B$ for some invertible matrix C. Since C is invertible, it must be square, and of course, C and C^{-1} must have the same dimensions. Let's say that C and C^{-1} are $d \times d$ matrices. What is d? Well, for the product CA to be defined, d must equal m. But for AC^{-1} to be defined, d must equal n. Hence, $m = n$, so A is square as claimed.

11. It merely *scales* each basis vector. The factor by which T scales $\mathcal{B}$'s i^{th} vector is the i^{th} diagonal entry of $[T]_{\mathcal{B}}$.

12. a) $\begin{pmatrix} 1/4 & 1 \\ -5/4 & 1 \end{pmatrix}$ b) $\frac{1}{6}\begin{pmatrix} 4 & -4 \\ 5 & 1 \end{pmatrix}$ c) $[T]_{\mathcal{B}} = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}$, $[T]_{\mathcal{A}} = \frac{1}{24}\begin{pmatrix} 92 & -16 \\ -5 & 76 \end{pmatrix}$, $[T]_{\mathcal{E}} = \frac{1}{4}\begin{pmatrix} 13 & -3 \\ -1 & 15 \end{pmatrix}$.

13. a) $\det A = \pm 1$, since isometries preserve area, volume, etc.

b) No. The shear $S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ isn't an isometry (e.g. it sends **j** to a vector of length $\sqrt{2}$), but $\det S = 1$.

c) Matrix C, being a rotation matrix, represents an isometry. Hence, $C^T = C^{-1}$. Thus, $CC^T = I$.

e) If A is orthogonal, it represents a linear isometry. Since any linear isometry's *inverse* is another linear isometry (think geometrically and you'll see why), A^{-1} also represents a linear isometry, which means that A^{-1} is an orthogonal matrix. But we know that $A^{-1} = A^T$ (since A is orthogonal), so A^T is an orthogonal matrix, too. Hence, the columns of A^T – which are the *rows* of A! – are mutually perpendicular unit vectors, as claimed.

g) The composition of two linear isometries is obviously a linear isometry itself.

h) A reflection is an isometry, so the Devil's matrix can be inverted simply by writing down its transpose. You'll still be busy for a while: If you never stop to sleep, eat, or use the bathroom (this is taking place in Hell, after all), you'll be inverting for 5 days, 3 hours, 12 minutes, and 36 seconds.

14. a) $[T]_{\boldsymbol{\mathcal{B}}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

b) T fixes the "$\mathbf{b}_1$-axis" while stretching space by factors of 2 and 3 along the "$\mathbf{b}_2$-axis" and "$\mathbf{b}_3$-axis" respectively. By linearity, this extends to all vectors in $\mathbb{R}^3$: The effect of T is to stretch the $\mathbf{b}_2$ and $\mathbf{b}_2$ *components* of any vector in $\mathbb{R}^3$ (viewed as a linear combination of the $\mathcal{B}$ basis vectors) by factors of 2 and 3 respectively. Simple.

15. For all that follows, let M_{ij} represent the ij^{th} entry of a matrix M.

a) This one's fairly obvious, but to prove it formally, we'll show that $((cA)^T)_{ij} = (cA^T)_{ij}$ for all i and j. I'll leave it to you to justify each equals signs in the following chain: $((\boldsymbol{cA})^{\boldsymbol{T}})_{\boldsymbol{ij}} = (cA)_{ji} = c(A)_{ji} = c(A^T)_{ij} = (\boldsymbol{cA^T})_{\boldsymbol{ij}}$.

b) See the answer to exercise 23E in Chapter 3.

c) Begin with a preliminary proposition: *Scaling one column of a matrix by c multiplies its determinant by c.* (Proof: If $c \geq 0$, this is fairly obvious; one edge of the frame of the 'box' determined by the columns is stretched by c, thus multiplying the box's 'volume' by c. The box's orientation is preserved. The total effect? The determinant is multiplied by c, as claimed. On the other hand, if $c < 0$, one edge of the box has its length multiplied by $|c|$, thus multiplying the box's volume by $|c|$. Moreover, the box's orientation reverses as it 'flips' across the origin. Total effect? The determinant is multiplied by $|c|(-1) = (-c)(-1) = c$ as claimed.) With this preliminary result in hand, we observe that cA is the result of scaling all n columns by c. Each scaled column multiplies the matrix's determinant by c, so the total effect is that $\det(cA) = c^n \det(A)$ as claimed.

(Sketch of a different proof: $\det(cA) = \det\big((cI)A\big) = \det(cI)\det(A) = c^n \det(A)$. Justify those equals signs!)

Chapter 7

1. a) Yes. 0.

2. a) E_0: the whole plane b) E_{-1}: the whole plane c) E_1: the line $y = 2x$, E_0: the line $y = -x/2$

d) E_1: the y-axis, E_{-1}: the x-axis.

3. In all cases, $\mathbf{v}$ is an eigenvector; the eigenvalues are: a) 2λ b) $n\lambda$ c) λ^2 d) λ^n e) λ^{-1} f) 1 g) $2\lambda + 3$.
(Sketch of proof for part E: Given that $A\mathbf{v} = \lambda\mathbf{v}$, left-multiply both sides by A^{-1}, then solve for $A^{-1}\mathbf{v}$.)

4. All "eigenstuff" is defined in terms of eigen*vectors*, so it will suffice to show that only square matrices can have eigenvectors. If A is any matrix, $\mathbf{v}$ is an eigenvector if and only if $\mathbf{v}$ and $\boldsymbol{A}\mathbf{v}$ are scalar multiples of one another. But to be scalar multiples of one another, these vectors must obviously have the same number of components. For its input and output vectors to have the same number of components, the matrix obviously must be square. Hence eigenvectors (and so, *a fortiori*, all eigenstuff) can exist only for square matrices as claimed.

5. Every orthogonal matrix represents an isometry, and thus preserves distances. Accordingly, the only possible eigenvalues of an orthogonal matrix are ± 1.

6. (Statement N holds) $\Leftrightarrow$ A doesn't map any nonzero vectors to $\mathbf{0}$ $\Leftrightarrow$ (Statement G holds).

7. a) $(12, -16, -10)$. To see why, observe that $\mathbf{v} = 4\mathbf{b}_1 - 2\mathbf{b}_2 + 5\mathbf{b}_3$, then follow your nose.
b) $(3x_1, 8x_2, -2x_3)$ c) $(\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n)$

8. Suppose eigenbasis $\mathcal{B}$ consists of $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$. Since these are eigenvectors, there exist scalars $\lambda_1, \lambda_2, \dots, \lambda_n$ such that $T(\mathbf{b}_i) = \lambda_i \mathbf{b}_i$ for all i. Or to spell this out a bit more tediously, $T(\mathbf{b}_i) = 0\mathbf{b}_1 + 0\mathbf{b}_2 + \cdots + \lambda_i \mathbf{b}_i + \cdots + 0\mathbf{b}_n$. This means that $T(\mathbf{b}_i)$'s coordinates relative to the eigenbasis $\mathcal{B}$ are all 0, except for its i^{th} coordinate, which is λ_i. It follows that in the ith column of $[T]_{\mathcal{B}}$, the only nonzero entry is the i^{th} entry, which is λ_i. That is, the *first* column is all zeros, except for its first entry, which is the eigenvalue λ_1 corresponding to the first eigenvector $\mathbf{b}_1$. The *second* column is all zeros, except for its second entry, which is the eigenvalue λ_2 corresponding to the second eigenvector $\mathbf{b}_2$. And so on and so forth for all the columns. The net result is that $[T]_{\mathcal{B}}$ is a diagonal matrix whose diagonal entries are the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ corresponding to the eigenbasis's eigenvectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$.

9. a) This follows immediately from the Invertible Matrix Theorem.

b) If we view A as the standard matrix of a linear map T, then matrix V is the $\mathcal{B}$-to-$\mathcal{E}$ change of basis matrix. Hence, $A = [T]_{\mathcal{E}} = V[T]_{\mathcal{B}}V^{-1}$. From Exercise 6, we know that $[T]_{\mathcal{B}}$ is a diagonal matrix whose diagonal entries are the eigenvalues corresponding to the eigenvectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$. If we name this diagonal matrix Λ, then we have the decomposition $A = V\Lambda V^{-1}$, as claimed.

c) $V = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$, $\Lambda = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, $V^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$, so our eigendecomposition of A is

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

d) This requires only 2 "costly" matrix multiplications (no matter how large n is!) as opposed to $(n-1)$ of them.

e) $A^{10} = \begin{pmatrix} 59049 & 58025 \\ 0 & 1024 \end{pmatrix}$.

10. The vector whose entries are all 1 will be an eigenvector of the matrix, with eigenvalue s. (This follows from the definition of matrix-vector multiplication.)

11. Read the section again.

12. a) Eigenvalues: 2 and 9. Eigenspaces: E_2 is the line $y = -x/2$, and E_9 is the line $y = 3x$. Taking one eigenvector from each eigenspace gives us two linearly independent eigenvectors, so an example of an eigenbasis would be, using standard coordinates,

$\begin{pmatrix} 2 \\ -1 \end{pmatrix}$, with eigenvalue 2 and $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$, with eigenvalue 9.

The corresponding eigendecomposition would be

$$\begin{pmatrix} 3 & 2 \\ 3 & 8 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 9 \end{pmatrix}\begin{pmatrix} 3/7 & -1/7 \\ 1/7 & 2/7 \end{pmatrix}.$$

b) Eigenvalues: 2 and 1. I'll leave the rest to you, as this is so similar to Part A.

c) No real eigenvalues, hence no eigenspaces, and thus obviously no eigenbasis or eigendecomposition.

d) Eigenvalue: -2. Its eigenspace is the line $y = -2x$. We don't have enough linearly independent eigenvectors to make an eigenbasis. Hence, no eigendecomposition is possible.

e) Eigenvalues: 5 and 1. Eigenspaces: E_5 is the x-axis in $\mathbb{R}^3$; E_1 is another line, the span of $\mathbf{i} + 4\mathbf{j} - 4\mathbf{k}$. Hence, we can have at most two linearly independent eigenvectors, which isn't enough to create an eigenbasis of $\mathbb{R}^3$ relative to the matrix. Consequently, the matrix lacks an eigendecomposition.

f) Eigenvalues: 2 and 1. Eigenspaces: E_2 is a line in $\mathbb{R}^3$, the span of $2\mathbf{i} + \mathbf{j}$; E_1 is another line, the span of $\mathbf{i} + \mathbf{j}$. As in the previous part, we don't have enough linearly independent eigenvectors to form an eigenbasis.

g) Eigenvalues: 3 and 0. Eigenspaces: E_3 is a line, the span of $\mathbf{i} + \mathbf{j} + \mathbf{k}$; E_0 is the plane $x + y + z = 0$. Since the dimensions of these eigenspaces add up to 3, we can find 3 linearly independent eigenvectors: any eigenvector from E_3 along with any two linear independent vectors from E_0. Thus, one such eigenbasis is:

$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ with eigenvalue 3, $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ with eigenvalue 0, $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ with eigenvalue 0.

The corresponding eigendecomposition is:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & -2/3 & 1/3 \\ 1/3 & 1/3 & -2/3 \end{pmatrix}.$$

13. Suppose M is an $n \times n$ triangular matrix. Since M and λI are triangular, their difference $M - \lambda I$ is triangular. Recall from Chapter 5 that a triangular matrix's determinant is the product of its diagonal entries. What are these entries? Well, if M's ii^{th} entry is m_{ii}, then $(M - \lambda I)$'s ii^{th} entry must be $(m_{ii} - \lambda)$. M's characteristic polynomial must therefore be $(m_{11} - \lambda)(m_{22} - \lambda)\cdots(m_{nn} - \lambda)$. The values of λ that make this polynomial equal to zero are obviously $m_{11},\ m_{22},\ \dots,\ m_{nn}$, so these numbers are M's eigenvalues, as claimed.

14. a) First equals sign: Every matrix has the same determinant as its transpose, as proved in Chapter 5, at the end of the section "Computing Determinants (by Row Reduction)". Second equals sign: Linearity of transposition (Chapter 3, Exercise 26B-C.) Third equals sign: If this isn't obvious, write out the matrix λI and transpose it.

b) Yes to eigenvalues, no to eigenvectors.

c) 1 must be an eigenvalue of A (since it will be an eigenvalue of A^T), but that's all we can say for certain.

15. A is diagonalizable, but not invertible. B is both. C is neither. D is invertible but not diagonalizable.

16. Let A be a 3×3 matrix. Its characteristic polynomial is a cubic polynomial. Since every cubic's graph crosses the horizontal axis at least once, this one has at least one real root. Hence, A has at least one real eigenvalue.

17. a) 3 (with multiplicity 2); 1/2 and -2 (each with multiplicity 1);
0, i, $-i$ (all with multiplicity 1); 1, -3, 2 (with respective multiplicities 1, 1, and 2)

b) 2 (alg mult: 2, geom mult: 1); 1 (alg mult: 1, geom mult: 1)

c) If we include any complex eigenvalues, then the sum of the algebraic multiplicities will be always n by the FTA. Thus, if we consider only real eigenvalues, the sum is either n (if all the eigenvalues are real), or less than n (if some eigenvalues are not real). In short, the sum of real algebraic multiplicities is *at most* n as claimed.

d) Each eigenvalue's GM is less than or equal to its AM, so the sum of the GMs is less than or equal to the sum of the GMs, which, by the previous part, is less than or equal to n.

e) $\Longrightarrow$) If the matrix is diagonalizable, it has n linearly independent eigenvectors. By their independence, each contributes a dimension to an eigenspace. Thus, the sum of the GMs (the sum of the eigenspaces' dimensions) is *at least* n. But by Part D, the sum of the GMs is also *at most* n. Together, these imply that sum of GMs is n.
$\Longleftarrow$) Suppose the sum of the GMs is n. Then the eigenspaces contain n linearly independent eigenvectors. These constitute an eigenbasis, so the matrix is diagonalizable.

f) See matrices A and D in Exercise 15.

g) $\Longleftarrow$) Suppose conditions (1) and (2) are satisfied. By (2), we have (sum of the GMs) = (sum of AMs), so (1) then implies that (sum of GMs) = n. Thus, by Part E, the matrix is diagonalizable.
$\Longrightarrow$) Suppose the matrix is diagonalizable. Then $n =$ (sum of GMs) $\leq$ (sum of AMs) $\leq n$. (The $=$ is justified by Part E; the two $\leq$s are justified - in order – by the theorem we're taking for granted and by Part C). Since the cumulative reading of that chain is that $n \leq n$, which should obviously be an *equality*, we deduce that the two $\leq$ symbols can in fact be replaced by $=$ signs. Replacing the second one gives us condition (1). Replacing the first implies that no individual GM can be less than its corresponding AM. Hence condition (2) holds, too.

h) In this case, (sum of GMs) < (sum of AMs) $\leq n$ (the inequality is from Part C), so he cumulative reading is that (sum of GMs) $< n$. Hence, by Part E, the matrix is *not* diagonalizable.

i) Every GM is a positive integer. (A "zero-dimensional eigenspace" would consist solely of $\mathbf{0}$, and thus would contain no eigenvectors, and wouldn't be associated with any particular eigenvalue.) Combining this with the theorem that GM $\leq$ AM, we know hat for this particular eigenvalue, $0 <$ GM ≤ 1 . Since GM is an integer, it must be 1. Hence, its corresponding eigenspace is a line.

18. a) Let $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Compute the characteristic polynomial in the usual way, and do a little algebra.

c) Use the old algebraic trick of expressing something (here, A's characteristic polynomial) in two different ways. By Part A, it is $\lambda^2 - (\mathrm{tr}(A))\lambda + \det(A)$. By the FTA, it is $k(\lambda - \lambda_1)(\lambda - \lambda_2)$ for some constant k. Equate these expressions and compare their coefficients. (You'll first see that $k = 1$, and then the result will follow.)

d) Being a 2×2 matrix, A has two eigenvalues (possibly complex, and counting repeated roots). By Part C, the eigenvalues' product is 7 and their sum is 8. The only numbers that will satisfy these conditions are 7 and 1, so these are A's eigenvalues. (I'll leave B to you.)

e) Matrix A: the columns are linearly dependent (the last column is the sum of the first two), so it's not invertible. Hence by the IMT, A has **0** as an eigenvalue. The rows all sum to **8**, so that's another eigenvalue (Exercise 10). Finally, A's eigenvalues must add up to its trace, which is 7, so its last remaining eigenvalue must be **-1**.

Matrix B: The eigenvalues of a triangular matrix are its diagonal entries (Exercise 13), so B's are **2** and **3**.

Matrix C: Similar to A, but here, it's more obvious that the *rows* are linearly dependent.

Chapter 8

2. $\begin{pmatrix} 1/2 \\ 5/2 \\ 1 \end{pmatrix}$ **3.** $\text{proj}_{\mathbf{u}}\mathbf{v_1} = \begin{pmatrix} 10/3 \\ 20/3 \\ 10/3 \end{pmatrix}$, $\text{proj}_{\mathbf{u}}\mathbf{v_2} = \begin{pmatrix} -5/6 \\ -5/3 \\ -5/6 \end{pmatrix}$ **4.** $\begin{pmatrix} -5/2 \\ 2 \\ -5/2 \end{pmatrix}$

5. a) Basis $\mathcal{B}$ is orthonormal. b) $\mathbf{v} = -2\mathbf{a_1} + \mathbf{a_2} - 2\mathbf{a_3} + 3\mathbf{a_4} = \frac{27}{2}\mathbf{b_1} - \frac{7}{2}\mathbf{b_2} - \frac{7}{2}\mathbf{b_3} + \frac{1}{2}\mathbf{b_4}$.

6. The basis vectors are mutually orthogonal, but not unit length. Normalizing them yields an orthonormal basis for the subspace. Vector $\mathbf{v}$'s orthogonal projection onto that subspace is $\frac{1}{4}(33\mathbf{e_1} + 21\mathbf{e_2} + 35\mathbf{e_3} + 19\mathbf{e_4})$.

7. Let $\mathbf{u} = \mathbf{w}/\|\mathbf{w}\|$. Then the following chain of equalities holds. I'll leave it to you to explain why each one is valid:

$$\text{proj}_{\mathbf{w}}\mathbf{v} = \text{proj}_{\mathbf{u}}\mathbf{v} = (\mathbf{v}\cdot\mathbf{u})\mathbf{u} = \left(\frac{\mathbf{v}\cdot\mathbf{w}}{\|\mathbf{w}\|}\right)\frac{\mathbf{w}}{\|\mathbf{w}\|} = \frac{(\mathbf{v}\cdot\mathbf{w})}{\|\mathbf{w}\|^2}\mathbf{w} = \left(\frac{\mathbf{v}\cdot\mathbf{w}}{\mathbf{w}\cdot\mathbf{w}}\right)\mathbf{w}.$$

8. a) $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ b) $\frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ c) $\text{rank}(AB) = 1$ d) $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

9. b) Use the fact that $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$.

10. a) $\frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$, $\frac{1}{3\sqrt{10}}\begin{pmatrix} -5 \\ -4 \\ 7 \end{pmatrix}$. b) $\mathbf{i}, \mathbf{j}, \mathbf{k}$ (Don't compute. Just think.) c) $\frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $\frac{1}{\sqrt{38}}\begin{pmatrix} 3 \\ -4 \\ -2 \\ 3 \end{pmatrix}$, $\frac{1}{2\sqrt{51}}\begin{pmatrix} 9 \\ 5 \\ -7 \\ -7 \end{pmatrix}$.

11. a) $5\mathbf{i} + 3\mathbf{j}$ b) $3\mathbf{j} - 5\mathbf{k}$ d) $\begin{pmatrix} 6 \\ 2 \\ -4 \end{pmatrix}$ e) Many possibilities. For example, $\mathbf{v_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbf{v_2} = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}$.

f) Depends on Part E. The $\mathbf{v_1}$ and $\mathbf{v_2}$ above, for example, yield $\mathbf{u_1} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $\mathbf{u_2} = \frac{1}{\sqrt{42}}\begin{pmatrix} -4 \\ 5 \\ -1 \end{pmatrix}$.

g) $\text{proj}_{W_2}\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

12. a) Let bases $\mathcal{A}$ and $\mathcal{B}$ consist of $\mathbf{a}_1, \dots, \mathbf{a}_k$ and $\mathbf{b}_1, \dots, \mathbf{b}_k$ respectively. By definition, the first column of R, the $\mathcal{A}$-to-$\mathcal{B}$ change of basis matrix, will be $\mathbf{a}_1$ expressed in $\mathcal{B}$-coordinates. By Gram-Schmidt, $\mathbf{b}_1 = \mathbf{a}_1/\|\mathbf{a}_1\|$, or equivalently, $\mathbf{a}_1 = \|\mathbf{a}_1\|\mathbf{b}_1$, so we see that R's first column has $\|\mathbf{a}_1\|$ for its first entry and zeros below that. The second step of Gram-Schmidt gives us

$$\mathbf{b}_2 = \frac{\mathbf{a}_2 - (\mathbf{a}_2\cdot\mathbf{b}_1)\mathbf{b}_1}{\|\mathbf{a}_2 - (\mathbf{a}_2\cdot\mathbf{b}_1)\mathbf{b}_1\|},$$

or equivalently,

$$\mathbf{a}_2 = (\mathbf{a}_2\cdot\mathbf{b}_1)\mathbf{b}_1 + \|\mathbf{a}_2 - (\mathbf{a}_2\cdot\mathbf{b}_1)\mathbf{b}_1\|\mathbf{b}_2.$$

From this, we see that R's second column has $(\mathbf{a}_2\cdot\mathbf{b}_1)$ for its first entry, $\|\mathbf{a}_2 - (\mathbf{a}_2\cdot\mathbf{b}_1)\mathbf{b}_1\|$ for its second, and nothing but zeros in the entries below that. A little thought about the subsequent steps of Gram-Schmidt will convince you that this pattern holds all the way down: each $\mathbf{a}_i$ is a linear combination of only the first i of $\mathcal{A}$'s basis vectors. Consequently, R is an upper triangular matrix as claimed.

c) In the case when Q is a *square* matrix, Q is an orthogonal matrix, and we've already proved (in Chapter 6) that any orthogonal matrix's inverse is its transpose. Thus $Q^TQ = Q^{-1}Q = I$, as claimed. [*Hint for the case when Q isn't square: Revisit that old proof, and adapt as necessary. Almost nothing needs to be changed.*]

d) If $M = QR$, then $Q^TM = Q^TQR$, which (by part C of this exercise) gives us $R = Q^TM$, as claimed.

f) There's no need for Gaussian elimination. The new system can be solved quickly through "back substitution": Start at equation in the bottom row (which is easily solved since it has only one variable), substitute its solution back into the next row up (which has only two variables – one of which we've already solved for), and so on until we've reached the top row.

13. We should solve the system

$$M\mathbf{x} = \mathbf{b}', \text{ where } M = \begin{pmatrix} 90 & 1 \\ 72 & 1 \\ \vdots & \vdots \\ 84 & 1 \end{pmatrix} \text{ and } \mathbf{b}' = \text{proj}_{\text{im}(M)} \begin{pmatrix} 96 \\ 85 \\ \vdots \\ 84 \end{pmatrix}.$$

If the solution $\mathbf{s}$ has components s_1 and s_2, the best-fit line will be $y = s_1 x + s_2$.
I'll leave the explanation of why this works for you.

14. We should solve the system

$$M\mathbf{x} = \mathbf{b}', \text{ where } M = \begin{pmatrix} 90^2 & 90 & 1 \\ 72^2 & 72 & 1 \\ \vdots & \vdots & \vdots \\ 84^2 & 84 & 1 \end{pmatrix} \text{ and } \mathbf{b}' = \text{proj}_{\text{im}(M)} \begin{pmatrix} 96 \\ 85 \\ \vdots \\ 84 \end{pmatrix}.$$

If the solution $\mathbf{s}$ has components s_1 and s_2, the best-fit line will be $y = s_1 x^2 + s_2 x + s_3$.
I'll leave the explanation of why this works to you.

15. $M\mathbf{x} = \mathbf{b}'$, where $M = \begin{pmatrix} 36^2 & 12^2 & 36 \cdot 12 & 36 & 12 & 1 \\ 38^2 & 12^2 & 38 \cdot 12 & 38 & 12 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 36^2 & 14^2 & 36 \cdot 14 & 36 & 14 & 1 \end{pmatrix}$ and $\mathbf{b}' = \text{proj}_{\text{im}(M)} \begin{pmatrix} 150 \\ 155 \\ \vdots \\ 161 \end{pmatrix}$.

16. If M's columns are linearly independent, they constitute a basis for $\text{im}(M)$. It follows that since $\mathbf{b}'$ is in $\text{im}(M)$, $\mathbf{b}'$ is a *unique* linear combination of M's columns. Equivalently, $M\mathbf{x} = \mathbf{b}'$ has a unique solution. That is, $M\mathbf{x} = \mathbf{b}$ has a unique least-squares solution.

17. In that case, $\mathbf{b}' = \mathbf{b}$, so the least-squares solution(s) will be the actual solution(s) to the system $M\mathbf{x} = \mathbf{b}$.

18. a) The system is *not* consistent, as Gaussian elimination shows. b) A least-squares solution is $\begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

19. a) $y = \frac{11}{10}x$ b) $y = \frac{1}{4}x^2 - \frac{3}{20}x + \frac{5}{4}$ **20**. $z = \frac{11}{5}x + \frac{2}{5}y - \frac{1}{5}$

21. a) Hint: Recall our earlier discussion about replacing $M\mathbf{x} = \mathbf{b}$ with $M\mathbf{x} = \mathbf{b}'$, where $\mathbf{b}' = \text{proj}_{\text{im}(M)}\mathbf{b}$.
c) Hint: Recall that for any vector $\mathbf{v}$, we have $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$.
d) This follows from the i^{th}-entry formula for matrix-vector multiplication.

Index