Accelerated Algebra 2018–19

Tom Leinster, University of Edinburgh

A short guide to university mathematics

This chapter and the next are to be read before the lecture of Monday, 24 September 2018

Welcome! This chapter contains guidance on how to read, write and do mathematics at university level, as well as a summary of some fundamental facts about sets and functions (some of which may be new to you). It will be useful for all your courses, not just this one.

School mathematics tends to emphasize calculation. At university level, the emphasis shifts more towards big, general concepts ('abstraction'), underpinned by watertight arguments ('proof'). You're here because you're extremely good at mathematics. But you're not just a calculating machine! This course will stretch your mind with abstract ideas and rigorous proofs.

The change in style from school to university mathematics takes some getting used to. Moreover, these accelerated courses go ***fast***. Be in no doubt: 'accelerated' is the right word! This chapter is full of tips and information to help you adapt to university mathematics quickly.

If you want to read more about the spirit of mathematics at an advanced level—this amazing, imaginary world where circles are perfect and numbers are exact, but which has repeatedly proved its value in solving problems in the imperfect and inexact world that we actually live in—try these two books. They both do a good job of evoking that spirit without going into technical detail.

- Timothy Gowers, *Mathematics: A Very Short Introduction.* Oxford University Press, 2002.
- Eugenia Cheng, *How to Bake* π. Profile Books, 2015.

How to read mathematics

The accelerated courses put a lot of responsibility on you. Aside from whatever you're asked to read and prepare for classes, you're expected to take the initiative and look things up in other sources whenever you find it necessary. There's a lot to do, and not much time to do it in.

However, mathematics needs to be read s l o w l y . I might read a 300-page novel in a day, but if you handed me a 300-page mathematics text, I'd expect it to take me months. So, you'll need to *organize yourself* and *make lots* *of time* for the reading that you're required to do. Here are some tips for using that time effectively.

Active reading I'm a big fan of swimming. I spend hours every day watching it on TV. I watch all the races, I know the names of all the star swimmers, and I know exactly how they perform the strokes in order to swim at record speeds. I've never actually got into the water myself, but after all that time watching, I know how to swim, right? So I confidently jump in the sea, and drown.

If you read mathematics without pen and paper at your side, you're making the same mistake. You won't learn it unless you *do* it.

Some concrete suggestions:

- *Print out the notes.* For this course, you're required to do that anyway. (And it's obviously in your interests: you can take a printout of the notes into the exam, but not an electronic copy.) However, I suggest you do it for all courses. As you're reading, annotate: highlight the parts that are important, circle the parts you don't understand, make notes in the margin, etc.

  Electronic annotation systems are getting better, but I still haven't seen one as good as pen and paper in terms of speed, flexibility, and handling mathematical symbols without fuss.

- *Constantly ask 'Is that really true?'* At its heart, a mathematics text is a sequence of statements, each one following logically from the previous ones. But you won't always be told exactly how it follows; the writer will assume that you've understood most of the text so far and can fill in the smaller gaps yourself. It's very easy to lose concentration and move your eyes over the text without really understanding how each step is justified. So keep asking yourself, every sentence: 'Is that true? Why?' If you can't answer those questions, make a note in the margin to come back to it later.

- *Do the exercises.* Apart from the homework and workshop activities, the text for this course (and many others) contains small exercises and questions for you to answer. Pick up that pen and do them! I promise that if you do, you'll understand the material better than if you don't.

- *Maintain a list of questions.* In a notebook, keep a running list of outstanding questions. Add to it every time you run into a part of the course that you don't understand or a homework problem you can't do, or a question occurs to you in the middle of the night. Get started right now. Find answers to those questions by asking me or your tutors or your classmates, or by reading, or by figuring it out yourself. Aim to have crossed every question off your list by the end of semester.

- *Before you read a proof, cover it up.* University mathematics texts contain a lot of theorems (and propositions, lemmas, etc.) followed by their proofs. Every time you get to a theorem, stop before you read the proof. Cover up the proof, and spend a few minutes trying to prove it yourself.

  If you succeed, you'll feel fantastic and you'll gain a sense of ownership. If you don't succeed, it might be because you realize that you don't really understand the *statement* of the theorem. (For instance, maybe you don't understand one of the words used.) That's a very important realization! In that case, you should go back to the earlier definitions until you're sure you do understand the statement. Or, it might be that you've failed to find the proof simply because it's difficult—in which case, you'll come away with an enhanced appreciation of where the difficulty lies. Whatever happens, attempting to prove the theorem yourself will help you.

How do I know what's important? These accelerated courses go at two or three times the speed of the non-accelerated versions, so we've had to remove almost all the material that isn't absolutely essential. In that sense, everything in them is important! But still, some things are more important than others.

In all mathematics courses,

The definitions are absolutely crucial.

It's almost impossible to exaggerate this point. The definitions are the foundation on which everything else is built.

For instance, suppose you're not quite sure of the definition of linear independence (a central concept in linear algebra, which we'll meet soon). Then it will be near-impossible for you to properly understand any theorem about linear independence, or follow any proof using linear independence, or solve any exercise on linear independence. You might be able to manage in a vague, half-understanding way, but as definition builds on definition and theorem builds on theorem, you'll soon find that even that vague understanding slips away, and before you know it you're thoroughly lost.

So, spend as much time as you need to understand the definitions, before you do anything else. Usually a definition is followed by some examples, and you can use those to help you understand what the definition means.

Mathematicians use different names for different types of result, and these names indicate which are the most important:

- *Theorems.* A theorem is a major, important result. These are the high points, the results to which you should pay special attention.

- *Propositions.* A proposition is a medium-sized result.

- *Lemmas.* A lemma is a small result, typically used as preparation for a proposition or theorem.

- *Corollaries.* A corollary is a result that follows easily from some other result, usually the last one stated. It might follow from the earlier result in such an obvious way that no proof is given at all; or if a proof is given, it is usually very short.

For instance, you might see two lemmas followed by a proposition, then a theorem, then a corollary. From this, you can deduce that the theorem and its corollary are the high point of the section—the dramatic climax—with the lemmas and proposition mainly acting as preparation.

Whether to call a result a theorem, proposition, lemma or corollary is a matter of judgement. There are no hard and fast rules.

A final tip for knowing what's important: listen for what your lecturers say about this, and write it down! If your lecturer spends five minutes in a class raving about how amazing and important the spectral theorem is, then at the very least pick up your pen and decorate your printout with stars etc. around the spectral theorem, and preferably jot down in the margin *why* it's so amazing and important.

How to write mathematics

There's not much point being a brilliant mathematician if you can't communicate your mathematics to other people. Part of what you'll learn in your degree is to *write* mathematics. We take this seriously, and reward it with marks in homework and exams.

Writing mathematics is different from *doing* mathematics. It is also different from writing ordinary English. It's a skill that no one is born with. You'll keep learning how to write mathematics for as long as you keep doing mathematics, and we'll help you in this throughout your degree.

The overarching principle is:

Have mercy on the reader.

Always remember: you're writing for someone else, not yourself.

Here are some specific tips.

- *Use words.* You don't need a lot of words to write mathematics clearly, and you shouldn't waffle. But you do need *some* words.

- *Write in sentences.* You're writing English: mathematical English, but English all the same. Use capital letters at the start of sentences, full stops at the end, and other punctuation in between. If it's not clear where one sentence ends and the next begins, your meaning may be ambiguous.

- *Use logical symbols (and use them correctly).* Every one of your homework answers should be a logically coherent argument. As well as small words like *so* and *if*, symbols like $\implies$ and $\iff$ are essential.

  Remember that these symbols have precise *meanings* (explained in the next section). Some students are in the habit of scattering the symbol $\to$ around their work as a kind of all-purpose connector. This symbol has at least two meanings in mathematics: convergence (as in '$x_n \to 1$ as $n \to \infty$') and function notation (as in 'a function $f\colon \mathbb{R} \to \mathbb{R}$'). If you're not using it with one of those two meanings, you're probably not using it right. If you mean $\implies$, write $\implies$!

- *Connect up your equations.* If someone asked you to solve the equation $2x + 7 = 15$ out loud, you wouldn't say '$2x + 7 = 15$ $2x = 8$ $x = 4$'. Instead, you'd say something like '$2x + 7 = 15$, so $2x = 8$, so $x = 4$'. That way, you're telling the listener how the equations are connected up logically: what implies what.

For exactly the same reason, if you were asked to solve the equation in writing, you shouldn't write

$$\begin{aligned} 2x + 7 &= 15 \\ 2x &= 8 \\ x &= 4. \end{aligned}$$

Instead, you should write something like

$$\begin{aligned} & 2x + 7 = 15 \\ \text{so} \qquad & 2x = 8 \\ \text{so} \qquad & x = 4. \end{aligned}$$

Alternatively, you might use $\Longrightarrow$ in place of 'so'; or maybe $\Longleftrightarrow$ would be more appropriate, depending on the context. The 'Logic' section below goes into detail about which symbol or word you should use.

Your answers should be logical arguments, not disconnected lists of equations. Small words and symbols like *so*, *i.e.* ('that is'), *iff*, $\therefore$, $\Longrightarrow$ and $\Longleftrightarrow$ can be used to turn a disjointed list of equations into a watertight chain of reasoning.

- *Work in rough first.* When you're doing homework, you'll find it really frees up your thinking if you begin by doing everything in rough. First, while you're solving the problem, write down your thoughts in whatever way comes naturally. Then, once it's solved, write out the argument properly, this time with your attention on clear communication. That way, you don't have to concentrate on two different tasks at once.

- *Write left-to-right, top-to-bottom.* This sounds obvious... but you'd be surprised how many people don't do it! Some people tend to write their thoughts all over the page: in the right-hand margin, in the middle, wherever. This creates at least two problems. First, the reader doesn't know what order to read it in. Second, if something's written way over on the right, is it actually intended to be read, or is it just rough work? (If it's rough work, best to write it only on rough paper, not on your final hand-in.)

- *Don't write the conclusion first.* Many mathematics questions ask you to prove something. If the question says 'Prove $X = Y$', don't begin your answer by writing '$X = Y$'. You don't *know* yet that $X = Y$. That's meant to be the end point of your argument, not the beginning!

  There's a natural tendency to write down the statement to be proved. Do it in rough if you want, but I see no reason to do in your final draft. Or if you must, always write 'Claim:' beforehand, to make clear that you're not making the mistake of assuming the statement that you've been asked to prove.

- *Define every letter you introduce.* This is probably the most common error I see. The golden rule is:

$\boxed{\textit{If you use a letter that isn't in the question, you have to define it.}}$

Just about the only exceptions are π, e and i.

For instance, if the question asks you about the angle between two vectors $\mathbf{u}$ and $\mathbf{v}$, then you might write down the formula $\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\|\cos\theta$, but you need to say that by 'θ' you mean the angle between $\mathbf{u}$ and $\mathbf{v}$. Never make the reader guess what you mean: tell them!

Logic

Logic is as essential to a mathematical argument as your skeleton is to your body. Without it, everything collapses into a mushy heap.

Common logical symbols In the explanations that follow, the letters P and Q stand for statements that are either true or false.

- $\therefore$ (therefore). Use '$P \therefore Q$' when you *know* that P is true and are deducing that Q is true. For instance, if you are given that x is a real number greater than 2, you could write:

$$\begin{aligned} x &> 2 \\ \therefore\ x^2 &> 4. \end{aligned}$$

- $\implies$ (implies). The expression '$P \implies Q$' means 'P implies Q' or 'if P then Q'. The only way for '$P \implies Q$' to be false is if P is true but Q is false.

  The implies sign can be used in situations where we don't know whether P is true. For instance, suppose you are given that x is a real number. Then it's correct to write 'If $x > 2$ then $x^2 > 4$' or '$x > 2 \implies x^2 > 4$'. But it's not correct to write '$x > 2$, therefore $x^2 > 4$' or '$x > 2 \therefore x^2 > 4$', because we don't know that x *is* greater than 2.

- $\iff$ (if and only if). The expression '$P \iff Q$' means 'P is true if and only if Q is true' or 'P is equivalent to Q'. It means that $P \implies Q$ and $Q \implies P$. So, $P \iff Q$ is true when P and Q are *either* both true *or* both false. For example, when x is a real number, $|x| > 2 \iff x^2 > 4$. Sometimes 'if and only if' is abbreviated to 'iff'.

- Occasionally people use the symbol $\because$ (because). *I strongly recommend that you don't use it.* Why? Because every time I've seen it used, it has been used badly, with chains of $\therefore$ and $\because$ signs whose logical meaning is highly ambiguous. Personally, I don't use either $\therefore$ or $\because$ at all.

On page 5, I encouraged you to *connect up your equations*, and explained why you shouldn't write things like

$$\begin{aligned} 2x + 7 &= 15 \\ 2x &= 8 \\ x &= 4. \end{aligned}$$

So, what logical symbols or words should you connect them up with?

It depends on the context. If you *know* that x satisfies $2x + 7 = 15$, you could use *so* or $\implies$ or $\therefore$. If you don't know that x satisfies $2x + 7 = 15$, but merely want to say that *if* x satisfies $2x + 7 = 15$ then $x = 4$, you should use $\implies$. If you want to say that the statements $2x+7 = 15$ and $x = 4$ are *equivalent* (in other words, $2x + 7 = 15 \implies x = 4$ and $x = 4 \implies 2x + 7 = 15$), then you should use $\iff$. There's no all-purpose solution: you always need to think about what you mean!

Converse and contrapositive The **converse** of the implication $P \implies Q$ is the implication $Q \implies P$. An implication and its converse are logically independent; that is, knowing whether one is true tells you nothing about whether the other is true.

For instance, suppose we are considering a real number x. Then the implication '$x > 2 \implies x^2 > 4$' is true, but its converse is '$x^2 > 4 \implies x > 2$', which is false (e.g. because $(-10)^2 > 4$ but $-10 \leq 2$). So in that case, the original implication is true but its converse is false. You should be able to think of other examples where an implication is false but its converse is true, or both are true, or both are false.

On the other hand, the **contrapositive** of the implication $P \implies Q$ is the implication (not Q) $\implies$ (not P). Any implication is logically equivalent to its contrapositive: either both implications are true or both are false.

For instance, the contrapositive of the implication '$11x - x^2 > 24 \implies x < 5$' is '$x \geq 5 \implies 11x - x^2 \leq 24$'. These two statements have exactly the same content. Either both implications are true, or both are false. (In fact, both are false.) The contrapositive of 'if it's Saturday, it's the weekend' is 'if it's not the weekend, it's not Saturday'. (Here, both implications are true.)

The word 'if' in definitions All human languages contain inconsistencies and exceptions to rules, and that's true for mathematical language too (although it's much more consistent than most). Here's one inconsistency: in definitions, the word 'if' is sometimes used to mean 'if and only if'.

This is best illustrated by an example. Here's the definition of prime number:

> An integer $n > 1$ is **prime** if it has no factors except 1 and n.

Strictly speaking, the word 'if' here should be 'if and only if'. An integer n is prime if it has no factors except 1 and n, *and only if* it it has no factors except 1 and n. That's what the definition means. Generally, any definition like 'a splodge X is **purple** if ...' really means that a splodge X is defined to be purple if *and only if* .... Although this is a slight inconsistency of language, you'll get used to it very quickly.

Quantifiers In mathematics, we often want to say that something is true *for all* x, or alternatively that *there exists* some x satisfying a certain condition. The terms in italics here are called 'quantifiers'.

- Suppose we want to say that the square of a real number is always at least zero. This is usually phrased as 'For all $x \in \mathbb{R}$, $x^2 \geq 0$'. We could also use 'for every' or 'for each' instead of 'for all'. Or we could use the symbol $\forall$, which means 'for all', so that the statement becomes '$\forall x \in \mathbb{R}$, $x^2 \geq 0$'.

To prove a statement beginning with the words ‘For all x’, you have to show that it is true for *every* x. It’s not enough to give just one or two values of x for which it’s true.

On the other hand, to *disprove* a statement beginning with ‘For all x’, it’s enough to find just a single value of x for which it fails. This is called a **counterexample** to the original (false) statement. For instance, to disprove the statement ‘for all $x \in \mathbb{R},\ \ x^2 \geq x$’, it is enough to point out that $x = 1/2$ is a counterexample (since $(1/2)^2 < 1/2$).

- Now suppose we want to say that there is some real number whose square is 2. This is usually phrased as ‘There exists $x \in \mathbb{R}$ such that $x^2 = 2$’, or in symbols, ‘$\exists x \in \mathbb{R} : x^2 = 2$’. An equivalent way of saying this is ‘There is at least one $x \in \mathbb{R}$ such that $x^2 = 2$’. (In fact, there are two such x: one positive, one negative.) Or, we could say ‘$x^2 = 2$ for some $x \in \mathbb{R}$’—but see the warning below about where to put the quantifiers.

  To prove a statement beginning with the words ‘There exists x such that’, you only need to find a single value of x satisfying the given condition. On the other hand, to disprove it, you have to show that no matter which value of x you pick, the condition is not satisfied.

Sometimes, you’ll have several quantifiers in a row. For instance, consider the statement

$$\text{for all } x \in \mathbb{R}, \text{ there exists } y \in \mathbb{R} \text{ such that } x + y > 0.$$

In cases like this, *the order of the quantifiers is crucial.* This statement is *not* the same as

$$\text{there exists } y \in \mathbb{R} \text{ such that for all } x \in \mathbb{R},\ x + y > 0.$$

In fact, the first statement is true and the second is false. (Have a think about why.)

Because the order of the quantifiers is crucial, it’s advisable to put them all at the start of the sentence. For instance, a statement like

$$\text{for all } x \in \mathbb{R},\ x + y > 0, \text{ for some } y \in \mathbb{R}$$

is ambiguous, because it’s not clear which of the previous two statements it’s supposed to mean. (In other words, is y allowed to depend on x or not?)

Sometimes it feels more natural to put a quantifier at the end of the sentence, and I’ll probably do that myself now and then, but make sure when you’re doing it that you don’t introduce any ambiguity.

Or In ordinary English, the word *or* sometimes means ‘one or the other but not both’. (If I hold out a tray of cakes to you and say ‘take a chocolate one or a lemon one’, and you take one of each, I’ll probably be annoyed.) But in mathematics, *or* always allows the possibility that both things are true. For instance, the set

$$\{n \in \mathbb{Z} : n \text{ is odd or } n > 10\}$$

contains the number 21, even though 21 is both odd *and* greater than 10. We could say ‘n is odd or $n > 10$ or both’—but in mathematics we never need to say ‘or both’, because it’s already contained in the mathematical definition of *or*.

TFAE It's a fact that for real numbers x,

$$x \geq 0 \iff \text{there exists } y \in \mathbb{R} \text{ such that } x = y^2 \iff |x| = x.$$

When you have a list of logically equivalent conditions like this, sometimes you'll see them presented with the words 'The following are equivalent', as in:

> **Theorem** *Let $x \in \mathbb{R}$. The following are equivalent:*
>
> i. $x \geq 0$;
>
> ii. *there exists $y \in \mathbb{R}$ such that $x = y^2$;*
>
> iii. $|x| = x$.

This means that each of the three conditions implies the other two.

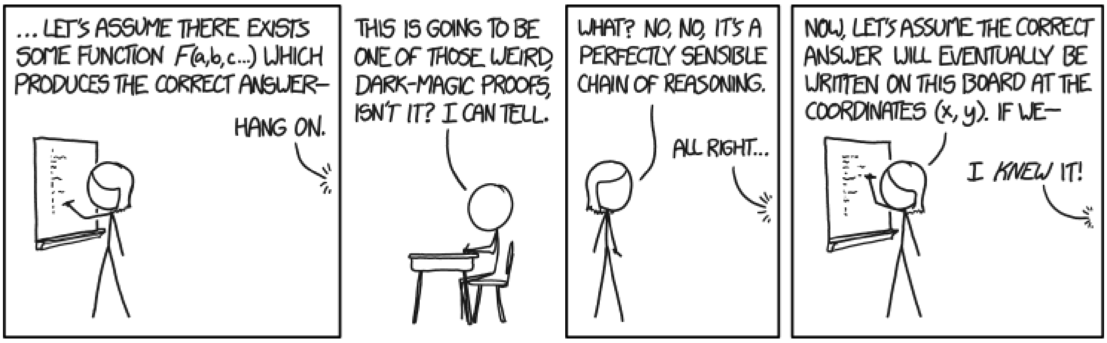
WLOG This stands for 'without loss of generality'. It's used when you're about to make an assumption that on the face of it is unjustified, but is actually harmless (usually for reasons of symmetry). Here's an example:

> **Lemma** *Let $m, n \in \mathbb{Z}$. If m is even or n is even then mn is even.*
>
> **Proof** Assume without loss of generality that m is even. Then $m = 2k$ for some integer k, so $mn = 2(kn)$, so mn is even. □

The point here is that although we're not given that m is even, we know that if it's not then n is even, and in that case we can carry out the same argument with the roles of m and n reversed. So in assuming that m is even, we haven't really lost any 'generality'—that is, we haven't made any genuinely unjustified assumptions.

The symbol □ marks the end of a proof.

Proofs


...LET'S ASSUME THERE EXISTS SOME FUNCTION F(a,b,c...) WHICH PRODUCES THE CORRECT ANSWER–
HANG ON.
THIS IS GOING TO BE ONE OF THOSE WEIRD DARK-MAGIC PROOFS, ISN'T IT? I CAN TELL.
WHAT? NO, NO, IT'S A PERFECTLY SENSIBLE CHAIN OF REASONING.
ALL RIGHT...
NOW, LET'S ASSUME THE CORRECT ANSWER WILL EVENTUALLY BE WRITTEN ON THIS BOARD AT THE COORDINATES (x, y). IF WE–
I KNEW IT!

xkcd.com/1724

A proof is a watertight argument. It is a complete chain of reasoning that leaves no room for doubt. At least, that's what 'proof' means in mathematics. There are other uses of the word 'proof' outside mathematics; for instance, your toothpaste may proclaim that it's 'clinically proven to reduce cavities', but that's not proof in the mathematical sense.

Much of university-level mathematics is proof-based. When you're studying big abstract concepts like infinity and 100-dimensional space, your intuition easily fails you, so you need rigorous, formal reasoning to stop you from making mistakes.

Watertight arguments are at the heart of mathematics, but other kinds of argument can be useful too (as long as you don't mistake them for proofs). For instance, imagine you're on a bus and you *think* you hear someone saying that the infinite sum

$$\frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots$$

is equal to $\pi^2/6$. You go home and try to prove it, but you can't. So you ask your computer to calculate

$$\frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{1000^2}$$

and discover that the result is within 0.001 of $\pi^2/6$. Is that a proof that what you overheard was correct? No. Is it reasonable evidence that it was *probably* correct? Yes. Will this kind of reasoning be used in this course? No. Is this kind of reasoning useful in other parts of mathematics? In some cases, yes.

Reading proofs In this course, you'll be reading a lot of proofs. But *why* should you read them? There are lots of reasons:

- Because a good proof explains why the result is true.
- Because it demonstrates that the result is correct. Authority counts for nothing in mathematics: just because your lecturer says something or it's printed in a book, doesn't mean it's true! You should always be sceptical.
- Because if you don't read the proof of a theorem, you're less likely to understand the statement of the theorem. It's too easy to read a theorem, say to yourself 'yeah yeah, I believe that', and move on, missing the point. But working through the proof really forces you to understand what the theorem says.
- Because it shows you which previous results are needed in order to prove the result at hand. In other words, it shows you what depends on what. Sometimes you'll read a lemma and think 'why did they bother stating that?' It might only be half a semester later, when that lemma plays a crucial part in the proof of some big theorem, that you'll see the point.
- Because it shows you how to reason about whatever subject you're learning. So, it will help you think up your own proofs, as you'll have to do in the homework, workshops and exams.
- Because each proof is a demonstration of how to write mathematics. It's something for you to imitate to improve your writing style.

That's *why* you should read proofs. But *how* should you read them?

- I already recommended on page 3 that before you read a proof, you should cover it up and try to prove the result yourself.

- For longer proofs—say, half a page or more—it's well worth trying to reduce the proof to two or three bullet points. The idea is that if you were stuck on a desert island and only able to remember those bullet points, you'd be able to reconstruct the whole proof. This is a really excellent way of deepening your understanding.

  (It's also a good habit to get into. Your exams this year are open-book, but in later years they'll be closed-book, and some of the questions will ask you to reproduce proofs from the course. The exam hall then becomes your desert island.)

Occasionally in this course (and others), a theorem will be stated with the words 'Proof omitted'. There are various reasons why we do this. Sometimes there's no time in the course, sometimes the proof's too hard, sometimes it's just tedious, and sometimes it would require too much of a digression.

Thinking up your own proofs In the step up from school to university mathematics, a big challenge is learning to think up your own proofs. Lots of people find this hard. Here are my tips:

- *Read your notes thoroughly.* For homework questions that require you to *calculate* something, you might be able to manage by just dipping into the notes, finding a similar example, and imitating it. But for questions that ask you to *prove* something, you often need to have read and understood your notes—including the proofs!—from the beginning. There really is no substitute for this.

- *Make sure you understand all the relevant definitions.* If you're asked to prove something about the nullity of a linear transformation, and you're not quite sure of the definition of nullity, your chances of being able to produce a correct proof are very low. So, take the time to go back and absorb those definitions. Once more: ***definitions are key!***

- To repeat myself again: *work in rough first.* This frees your mind.

- *Begin by writing down what you know.* Many questions are of the form 'assuming P, prove Q'. In rough, write down P at the top of the page—including all the relevant definitions. Write down Q at the bottom of the page—including all the relevant definitions. Then ask yourself how to fill in the gap. You might be surprised how often this last stage is really easy.

- *Draw a picture.* Not enough students do this! Pictures are helpful in almost all parts of mathematics, definitely including this course.

- *If stuck, try an example.* For instance, if you're asked to prove something about $n \times n$ matrices, try it for 2×2 or 3×3 (or even 1×1) matrices. If you succeed, it will usually give you a big clue as to how to do it in the general case.

- *If stuck, try to prove the opposite.* If a question asks you to prove that Q is true, and you're struggling, try to prove that Q is false. You won't succeed at that either (unless the question is wrong!), but the attempt might help you to see why Q must in fact be true.

Checking your proof Congratulations! You've finally found a proof. Or at least, you think you have... but how can you be sure?

The next step is to *write it out neatly.* You worked in rough first, right? Now's the time to go through your proof carefully again, writing it out properly with correct language and logic. Mathematical language is designed deliberately to make errors stand out, in the same way that lab scientists wear white coats in order that stains and spills are immediately visible. Make sure you're using notation and terminology correctly. If you don't, it's all too easy to fool yourself that you've got a correct argument, and as the physicist Richard Feynman said:

> The first principle is that you must not fool yourself—and you are the easiest person to fool.

Finally, ask yourself whether you're using all the hypotheses (conditions) in the question. For instance, suppose the question is this:

> Let A be a square matrix whose rows are linearly independent. Prove that the columns of A are linearly independent.

If your proof doesn't use the hypothesis that A is square, you should be suspicious! *Maybe* the person setting the question put that condition in there by accident, or to make the question easier, or to trick you. But it could also be a sign that you've made a mistake.

Sets and functions

Sets and functions play a part in almost every area of mathematics, and it's important that you're in control of both the fundamental concepts and the notation.

Some standard sets Here are some sets that you'll come across repeatedly.

- $\varnothing$, the **empty set**. (This symbol is different from ϕ, the Greek letter phi.)
- $\mathbb{N}$, the set of **natural numbers**. Some people count 0 as a natural number, so that $\mathbb{N} = \{0, 1, 2, \ldots\}$. Others don't, so that $\mathbb{N} = \{1, 2, 3, \ldots\}$. In this course, I plan to avoid the issue by not using the symbol $\mathbb{N}$ or the term 'natural number' at all. (But in case I slip up, I mean to include 0.)
- $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, the set of integers.
- $\mathbb{Q}$, the set of **rational numbers** (numbers that can be expressed as m/n for some integers m and n, with $n \neq 0$).
- $\mathbb{R}$, the set of real numbers.
- $\mathbb{C}$, the set of complex numbers.

Specifying a set To specify a *finite* set, you can simply list its elements: $\{1, 3, 5\}$. This doesn't work so well for infinite sets, but instead we can use notation such as

$$A = \{a \in \mathbb{Z} : a^2 - a \geq 6\} \qquad \text{or} \qquad B = \{t^2 : t \in \mathbb{R}\}.$$

Elements and subsets There is a difference between $\in$ ('is an element of') and $\subseteq$ ('is a subset of', which some people write as $\subset$). The statement '$x \in Y$' means that x is an element of Y, but the statement '$X \subseteq Y$' means that every element of X is an element of Y.

For instance, 2 is an element of the set $B = \{t^2 : t \in \mathbb{R}\}$, because there exists some $t \in \mathbb{R}$ such that $t^2 = 2$. But 2 is not a subset of B: that doesn't even make sense, since 2 is not a set at all. On the other hand, $\{2\}$ is a subset of B, because every element of $\{2\}$ is an element of B. Indeed, $\{2\}$ only *has* one element, 2 itself, so the statement that $\{2\} \subseteq B$ is equivalent to the true statement that $2 \in B$.

The word **contains** is used in two different ways. Sometimes people say 'Y contains x' to mean that $x \in Y$, and sometimes people say 'Y contains X' to mean that $X \subseteq Y$. It's not really ambiguous, since the context always makes it clear which is intended.

How do you prove that two subsets are equal? Specifically, suppose we have two subsets A and B of a set X, and we want to prove that $A = B$. A useful strategy is to first prove that $A \subseteq B$, then prove that $B \subseteq A$. This means that every element of A is an element of B and vice versa, and so A and B are equal. If you're going to use this strategy, your argument must have two parts. The first part looks like this:

> Let $a \in A$. ... [argument goes here] ... So $a \in B$. Hence $A \subseteq B$.

The second part looks like this:

> Let $b \in B$. ... [argument goes here] ... So $b \in A$. Hence $B \subseteq A$.

Then you conclude that $A = B$. We'll see some examples of this strategy later.

New sets from old Given subsets A and B of a set X, we can form several new subsets of X:

- The **intersection** $A \cap B = \{x \in X : x \in A \text{ and } x \in B\}$.
- The **union** $A \cup B = \{x \in X : x \in A \text{ or } x \in B\}$. Remember from page 9 that the word 'or' in mathematics always permits the possibility that both things are true, so $A \cap B \subseteq A \cup B$.
- The **complement** $A \setminus B = \{x \in X : x \in A \text{ but } x \notin B\}$. Note that this is defined whether or not B is a subset of A.

Functions, domains and codomains Here's the fundamental definition. A **function** (or **mapping**) consists of three things:

- a set A, called the **domain** of the function;
- another set B, called the **codomain** of the function; and
- a rule f that assigns to each element $a \in A$ an element $f(a) \in B$.

We write

$$f\colon A \to B$$

to mean that f is a function with domain A and codomain B.

For instance, there is a function $f\colon \mathbb{R} \to \mathbb{R}$ given by $f(a) = a^2$ $(a \in \mathbb{R})$. Here the domain is $\mathbb{R}$ and the codomain is also $\mathbb{R}$. Different elements of the domain can be mapped to the same element of the codomain (e.g. $f(-3) = f(3)$). Also, some elements of the codomain may not have anything mapping to them at all (e.g. there is no a in the domain such that $f(a) = -1$). However, f assigns to each element of the domain exactly *one* element of the codomain. There is no function $g\colon \mathbb{R} \to \mathbb{R}$ defined by '$g(x) = \pm x$', for instance.

The following point tends not to be emphasized in school mathematics, but is crucial at university level:

The domain and codomain are part of the function.

For example, there are functions

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad g\colon \mathbb{R} \to \mathbb{C}, \qquad h\colon \mathbb{C} \to \mathbb{C}$$

defined by

$$f(x) = x^2 \ (x \in \mathbb{R}), \qquad g(x) = x^2 \ (x \in \mathbb{R}), \qquad h(x) = x^2 \ (x \in \mathbb{C}).$$

Although f, g and h are all given by the same formula, ***they are different functions***. In order for two functions to be equal, they must have the same domain and the same codomain, as well as having the same effect on elements. For instance, $f \neq g$ because f and g have different codomains.

For similar reasons, if someone says to you 'define a function F by $F(x) = x^2$' then they have not, in fact, defined a function, because they have not specified a domain or a codomain.

Two types of arrow When we have a function $f\colon A \to B$, we sometimes write $f\colon a \mapsto b$ (or just $a \mapsto b$) to mean that $f(a) = b$. For instance, instead of writing 'define a function $f\colon \mathbb{Z} \to \mathbb{Q}$ by $f(n) = n^2/2$ $(n \in \mathbb{Z})$', we might write 'define

$$\begin{array}{cccc} f\colon & \mathbb{Z} & \to & \mathbb{Q} \\ & n & \mapsto & n^2/2 \end{array}$$

$(n \in \mathbb{Z})$'. But notice that $\to$ and $\mapsto$ are different symbols. Don't get them mixed up! The $\to$ arrow goes between sets; the $\mapsto$ arrow goes between elements. For example, '$n \to n^2/2$' is wrong.

Composition and identities Suppose we have functions $f\colon A \to B$ and $g\colon B \to C$. Then we can feed the output of f into the input of g to make a new function $g \circ f\colon A \to C$:

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[rr, bend right, "g \circ f"'] & B \arrow[r, "g"] & C \end{tikzcd}$$

This new function $g \circ f$ is given by

$$(g \circ f)(a) = g(f(a))$$

$(a \in A)$, and is called the **composite** of g and f.

For any set A, there is a function $1_A \colon A \to A$, called the **identity** on A (and sometimes written as id_A). It is given by

$$1_A(a) = a$$

($a \in A$). In other words, it does nothing at all! You might wonder what use this could possibly have. Thousands of years ago, people wondered what use the number zero could possibly have—and in fact, there was serious resistance to calling it a number at all. But both things turn out to be quite useful.

Injective, surjective and bijective functions Let A and B be sets. A function $f \colon A \to B$ is:

- **injective** (or **one-to-one**) if for each $b \in B$, there is *at most one* $a \in A$ such that $f(a) = b$;
- **surjective** (or **onto**) if for each $b \in B$, there is *at least one* $a \in A$ such that $f(a) = b$;
- **bijective** (or a **one-to-one correspondence**) if for each $b \in B$, there is *exactly one* $a \in A$ such that $f(a) = b$ (or equivalently if f is both injective and surjective).

For example, the function $f \colon \{2, 3\} \to \{1, \ldots, 10\}$ defined by $f(a) = a^2$ is injective, because for each $b \in \{1, \ldots, 10\}$, the equation $a^2 = b$ has at most one solution a belonging to the set $\{2, 3\}$. (There is exactly one solution if $b = 4$ or $b = 9$, and there are none otherwise.) However, f is not surjective, because there is no $a \in \{2, 3\}$ such that $a^2 = 1$, for instance.

The function $g \colon \{-2, 2\} \to \{4\}$ defined by $g(a) = a^2$ is surjective, because the only element of the codomain is 4, and $(-2)^2 = 4$. (Also $2^2 = 4$, but to prove surjectivity we only need to find *one* element a of $\{-2, 2\}$ such that $a^2 = 4$.) However, g is not injective, because $(-2)^2 = 2^2$ and $-2 \neq 2$.

The function $h \colon \mathbb{R} \to \mathbb{R}$ defined by $h(a) = a^3$ is bijective, because every real number has exactly one real cube root.

Inverse functions Let A and B be sets and let $f \colon A \to B$ be a function. If f is bijective then there is a unique function $f^{-1} \colon B \to A$ such that $f^{-1} \circ f = 1_A$ and $f \circ f^{-1} = 1_B$. This function f^{-1} is called the **inverse** of f. Conversely, if f has an inverse then f is bijective.

The words '**there is a unique** ...' mean 'there is one and only one ...'. So I am saying that if f is bijective then there is exactly one function $g \colon B \to A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$, and I am writing this function g as f^{-1}.

The Greek alphabet

Mathematicians like to use Greek letters as well as English letters, so you'll need to learn them sooner or later. The table below shows the Greek alphabet, along with the name of each letter and how the name is pronounced by (British) mathematicians.

To any native Greek speakers, apologies for what we have done to your language. How we say your letters is not how you say them. But if you want to make yourself understood, I'm afraid you'll have to do as we do.

Lower case	Upper case	Name	Pronunciation
α		alpha	AL-fa
β		beta	BEE-ta
γ	Γ	gamma	GAM-ma
δ	Δ	delta	DEL-ta
ε		epsilon	EP-si-lon
ζ		zeta	ZEE-ta
η		eta	EE-ta
θ	Θ	theta	THEE-ta (soft *th*, as in *think*)
ι		iota	eye-OH-ta
κ		kappa	KAP-pa
λ	Λ	lambda	LAM-da
μ		mu	myoo (rhymes with *few*)
ν		nu	nyoo (rhymes with *few*)
ξ	Ξ	xi	ksy (rhymes with *pie*)
o		omicron	never used in mathematics
π	Π	pi	pie
ρ		rho	roe (rhymes with *go*)
σ	Σ	sigma	SIG-ma
τ		tau	rhymes with *now*
υ		upsilon	almost never used in mathematics
ϕ	Φ	phi	fy (rhymes with *pie*)
χ		chi	ky (rhymes with *pie*)
ψ	Ψ	psi	psy (rhymes with *pie*); the *p* is pronounced
ω	Ω	omega	OH-me-ga

Many uppercase Greek letters look like uppercase English letters. For example, an uppercase α is A. These ones are not used in mathematics, or rather, are interpreted as English rather than Greek letters. For that reason, they are not shown in the table.

Chapter A

Background

This chapter and the previous one are to be read before the lecture of Monday, 24 September 2018

This is a course on linear algebra. It is about the interplay between *geometry*—the study of space and position—and *algebra*—the study of symbolic expressions. Geometry gives us intuition and lets us harness our visual sense. Algebra gives us certainty. We can often reduce hard-to-visualize geometric situations to simple algebraic calculations, and we can often understand algebraic constructions by viewing them geometrically. The two ways of thinking are complementary and mutually beneficial.

What is *linear* algebra? From the geometric viewpoint, 'linear' means that we're concerned with straight lines, planes, and so on: flat shapes, not curved ones. From the algebraic viewpoint, it means that we will mostly encounter expressions such as $2x + 3y$, not $2x^2 + 3y^3$ or $e^x \sin y$.

This chapter collects together various pieces of background material that we will need later in the course. Mostly it is about vectors and matrices (Sections A2–A5), including a section on the dot and cross products (Section A3) that you may find useful if you are taking Several Variable Calculus and Differential Equations. There is also some background on complex numbers and the fundamental theorem of algebra (Section A6). But first of all, we meet some very useful notation for adding things up.

A1 Summation notation

Suppose we have numbers $a_1, a_2, \dots, a_n$ and we want to consider their sum. We could of course write it as

$$a_1 + a_2 + \cdots + a_n,$$

but it is often convenient to write it instead as $\sum_{i=1}^n a_i$. Here the letter i is a 'dummy variable', which means that you could use any other letter $(j, q, Z, \dots)$ and it would make no difference. Thus,

$$\sum_{i=1}^n a_i = \sum_{j=1}^n a_j = \sum_{q=1}^n a_q = \sum_{Z=1}^n a_Z = a_1 + a_2 + \cdots + a_n.$$

This is very similar to the fact that $\int_a^b f(t)\,dt = \int_a^b f(x)\,dx = \int_a^b f(q)\,dq = \cdots$ (and indeed, $\int$ can be thought of as a continuous version of $\sum$).

I said that it makes no difference which letter you use, but that's not *quite* true: you need to be very careful about re-using letters that you've already used elsewhere. Certainly you shouldn't write $\sum_{n=1}^n a_n$, because then you're using n to mean two different things. But there are some more subtle cases.

For example, if you're given numbers $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$, is it safe to write the sum of all of them as

$$\sum_{i=1}^n a_i + \sum_{i=1}^n b_i\,?$$

Yes, and it means

$$a_1 + a_2 + \cdots + a_n + b_1 + b_2 + \cdots + b_n.$$

If you're not comfortable with this notation, you could clarify it by inserting some brackets:

$$\Bigl(\sum_{i=1}^n a_i\Bigr) + \Bigl(\sum_{i=1}^n b_i\Bigr).$$

Or you could change the second pair of is to js. But it's OK not to.

Now suppose that you're given numbers $a_1, a_2, \ldots, a_m$ and $b_1, b_2, \ldots, b_n$, where perhaps $m \neq n$. Is it safe to write

$$\sum_{i=1}^m a_i + \sum_{i=1}^n b_i\,?$$

It's not *wrong*, but I'd recommend against doing it. If i runs over $1, 2, \ldots, m$ in one place and $1, 2, \ldots, n$ in another, it's a recipe for confusion. It's much neater if you can stick to the convention that throughout an argument, i always runs over $1, 2, \ldots, m$ and j always runs over $1, 2, \ldots, n$, for instance. So in this case, it would be better to write

$$\sum_{i=1}^m a_i + \sum_{j=1}^n b_j$$

instead.

The notation $\sum_{i=1}^n a_i$ has some variants. Sometimes we write it as $\sum_{1\leq i\leq n} a_i$, or $\sum_i a_i$, or $\sum_1^n a_i$, or simply $\sum a_i$. Some of these forms omit information that should in principle be included, but they're safe as long as there's no possible misunderstanding in the context.

There's nothing to stop you putting sums inside other sums. For instance, if we have a grid of numbers

$$\begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{matrix} \tag{A:1}$$

then we can write down the expression

$$\sum_{i=1}^m \sum_{j=1}^n a_{ij}.$$

This expression means $\sum_{i=1}^{m} s_i$, where $s_i = \sum_{j=1}^{n} a_{ij}$ for $1 \leq i \leq m$. Now s_i is the sum of the ith row of the grid, so $\sum_{i=1}^{m} s_i$ is

$$\text{(sum of the 1st row)} + \text{(sum of the 2nd row)} + \cdots + \text{(sum of the } m\text{th row)}$$

or equivalently

$$(a_{11} + a_{12} + \cdots + a_{1n}) + (a_{21} + a_{22} + \cdots + a_{2n}) + \ \cdots \ + (a_{m1} + a_{m2} + \cdots + a_{mn}).$$

That's what $\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}$ means. But the summation notation is much shorter!

We've just seen that $\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}$ is the sum of all the numbers in the grid (A:1), taken row by row. But we'd get the same grand total if we added up the numbers column by column. So,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} = \sum_{j=1}^{n} \sum_{i=1}^{m} a_{ij}. \tag{A:2}$$

In words, *you can change the order of summation.* This is an important principle of working with sums.

Another important principle is that given numbers $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$,

$$\sum_{i=1}^{n} (a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i. \tag{A:3}$$

For example,

$$(a_1 + b_1) + (a_2 + b_2) + (a_3 + b_3) = (a_1 + a_2 + a_3) + (b_1 + b_2 + b_3).$$

And another one is that given numbers $a_1, \ldots, a_n$ and c,

$$\sum_{i=1}^{n} c a_i = c \sum_{i=1}^{n} a_i \tag{A:4}$$

(e.g. $ca_1 + ca_2 = c(a_1 + a_2)$). For instance, if $c = \sum_{k=1}^{p} b_k$ for some $b_1, b_2, \ldots, b_k$ then equation (A:4) gives

$$\sum_{i=1}^{n} \left(\sum_{k=1}^{p} b_k \right) a_i = \left(\sum_{k=1}^{p} b_k \right) \left(\sum_{i=1}^{n} a_i \right).$$

It's worth getting used to summation notation. It might not come naturally at first, but if you force yourself to use it then eventually it will pay off. So that you don't make mistakes while you're adapting, use this policy:

If in doubt, write it out.

That is, if you're not sure whether something you've done involving $\sum$s is valid, just write it out using $+$ and $\cdots$. That way, you'll be able to tell whether what you've done is correct.

A2 Vectors

From a geometric viewpoint, a vector is something with direction and length. This description might be good enough in 2- and 3-dimensional space, but what does it mean in higher dimensions? What *are* higher dimensions?

Let $n \geq 0$. For us, an n-dimensional **vector** is simply a list of n real numbers $x_1, x_2, \ldots, x_n$, which we write in a column:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

It saves space if we write

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \tag{A:5}$$

and so on. ***We will use this convention for the whole course.*** That is, whenever we have a vector $\mathbf{x}$, I will write its ith entry as x_i, and I will not usually explain that this is what I mean by x_i. Similarly, the ith entry of a vector $\mathbf{y}$ will be written as y_i, the kth entry of a vector $\mathbf{u}$ will be written as u_k, and so on.

Remark A2.1 It is standard to use bold typeface for vectors. But in handwriting, it's hard to write bold symbols like $\mathbf{x}$, so we indicate a vector by underlining instead: $\underline{x}$. The point of this convention is to make sure we don't get vectors confused with **scalars** (elements of $\mathbb{R}$).

So don't forget: ***in your work, always underline vectors.***

The set of all n-dimensional vectors is called $\mathbb{R}^n$. Often elements of $\mathbb{R}^n$ are written in the horizontal notation

$$(x_1, x_2, \ldots, x_n)$$

instead of as a column. It makes no real difference which notation one uses, but in this course we are going to use column notation. The elements of $\mathbb{R}^n$ are sometimes called 'vectors', sometimes 'points', and sometimes just 'elements'.

What does $\mathbb{R}^n$ *mean*? I probably can't visualize $\mathbb{R}^{10}$ much better than you. It's just the set whose elements are lists $x_1, x_2, \ldots, x_{10}$ of ten real numbers. For instance, suppose I'm interested in foot shapes, and I do a survey of residents of Edinburgh, recording for each person the length in millimetres of each of their toes. That means that for each person surveyed, I have an element of $\mathbb{R}^{10}$: a 10-dimensional vector. There's no deep meaning; it's just a list of numbers.

When you tell non-mathematicians that you're studying spaces of arbitrarily high dimension, they sometimes ask 'If time is the fourth dimension, what is the fifth?' If anyone asks you this, you should reply 'It's the length of your left big toe'.

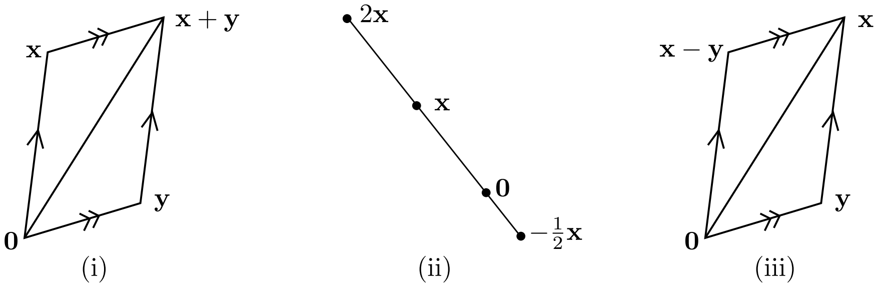


Figure A.1: Algebraic operations on vectors: (i) addition, (ii) scalar multiplication, and (iii) subtraction

When are two vectors equal? Again, there is no mystery or deep meaning. Simply, the rule is that for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$\mathbf{x} = \mathbf{y} \iff x_i = y_i \text{ for all } i \in \{1, \ldots, n\}.$$

Note that if $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$ then we can't possibly have $\mathbf{x} = \mathbf{y}$ unless $n = m$. When $n \neq m$, it doesn't even make sense to ask whether $\mathbf{x} = \mathbf{y}$.

New vectors from old

Any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ can be added together to get a third vector $\mathbf{x} + \mathbf{y} \in \mathbb{R}^n$, defined like this:

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

Any vector $\mathbf{x}$ in $\mathbb{R}^n$ can be multiplied by any scalar $a \in \mathbb{R}$ to get another vector $a\mathbf{x} \in \mathbb{R}^n$, defined by

$$a\mathbf{x} = \begin{pmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_n \end{pmatrix}.$$

One especially important vector in $\mathbb{R}^n$ is the **zero vector**, defined by

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The geometric interpretations of addition and scalar multiplication are indicated in Figure A.1.

Now here are some properties of addition and scalar multiplication of vectors.

Lemma A2.2 *Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$. Then:*

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

ii. $\mathbf{x}+\mathbf{y}=\mathbf{y}+\mathbf{x}$*;*

iii. $\mathbf{x}+\mathbf{0}=\mathbf{x}$*;*

iv. $a(b\mathbf{x})=(ab)\mathbf{x}$*;*

v. $1\mathbf{x}=\mathbf{x}$*;*

vi. $a(\mathbf{x}+\mathbf{y})=a\mathbf{x}+a\mathbf{y}$*;*

vii. $(a+b)\mathbf{x}=a\mathbf{x}+b\mathbf{x}$*;*

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

Proof This is a series of routine checks using the definitions. I will just do part (vi) and leave the rest to you. You should do a few of them, until you're confident that you could do them all.

For (vi), we have

$$\begin{aligned}
a(\mathbf{x}+\mathbf{y}) &= a\left[\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}+\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix}\right] && \text{by the convention introduced in (A:5)}\\
&= a\begin{pmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{pmatrix} && \text{by definition of vector addition}\\
&= \begin{pmatrix}a(x_1+y_1)\\a(x_2+y_2)\\\vdots\\a(x_n+y_n)\end{pmatrix} && \text{by definition of scalar multipliciation}\\
&= \begin{pmatrix}ax_1+ay_1\\ax_2+ay_2\\\vdots\\ax_n+ay_n\end{pmatrix} && \text{since } a(p+q)=ap+aq \text{ for all } p,q\in\mathbb{R}\\
&= \begin{pmatrix}ax_1\\ax_2\\\vdots\\ax_n\end{pmatrix}+\begin{pmatrix}ay_1\\ay_2\\\vdots\\ay_n\end{pmatrix} && \text{by definition of vector addition}\\
&= a\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}+a\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix} && \text{by definition of scalar multiplication}\\
&= a\mathbf{x}+a\mathbf{y},
\end{aligned}$$

as required. □

We write $(-1)\mathbf{x}$ as $-\mathbf{x}$; it is the vector whose ith entry is $-x_i$, and it satisfies $-\mathbf{x}+\mathbf{x}=\mathbf{0}$. As you'd guess, we write $\mathbf{x}+(-\mathbf{y})$ as $\mathbf{x}-\mathbf{y}$. Figure A.1(iii) shows an example of subtraction of vectors.

Distance

Pythagoras's theorem tells us that in the plane $\mathbb{R}^2$, the distance between the origin $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ and another point $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ is $\sqrt{x_1^2 + x_2^2}$. Put another way, this is the length of the vector $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Similarly, you may be familiar with the fact that in $\mathbb{R}^3$, the distance between $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ is $\sqrt{x_1^2 + x_2^2 + x_3^2}$. So when n is 2 or 3, the length of a vector $\mathbf{x} \in \mathbb{R}^n$ is

$$\sqrt{\sum_{i=1}^{n} x_i^2}. \tag{A:6}$$

That's fine for two and three dimensions, but what is the length of a 248-dimensional vector? Rather than try to develop our intuition about $\mathbb{R}^{248}$, we simply *define* it by the formula above. That is:

Definition A2.3 Let $n \geq 0$. The **length** of a vector $\mathbf{x} \in \mathbb{R}^n$ is

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}. \tag{A:7}$$

(Some people write $|\mathbf{x}|$ rather than $\|\mathbf{x}\|$. It doesn't matter which we use; it's just a matter of convention.)

Let's pause for a moment and reflect on what we just did. We took a geometric concept, length, that we knew about in low dimensions (two and three) but not in higher dimensions. We noted that in low dimensions, the geometric concept has an algebraic formulation (equation (A:6)). We noted that the algebraic formulation could be generalized in an obvious way to higher dimensions. And we then used that as our *definition* of the concept in higher dimensions. We'll see this same pattern over and over again in this course, where we generalize a geometric concept from low to high dimensions by turning it into a piece of algebra.

We've considered distance between the origin and any other point of $\mathbb{R}^n$. More generally, the **distance** between two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ is defined to be $\|\mathbf{y} - \mathbf{x}\|$. For instance, if $n = 10$ and $\mathbf{x}$ and $\mathbf{y}$ represent the toe-lengths of two people (as above), then the smaller $\|\mathbf{y} - \mathbf{x}\|$ is, the more similar those two people are in terms of the lengths of their toes.

Lemma A2.4 *Let $n \geq 0$, let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and let $a \in \mathbb{R}$. Then:*

i. $\|\mathbf{x}\| \geq 0$*, with equality if and only if* $\mathbf{x} = \mathbf{0}$*;*

ii. $\|\mathbf{x} - \mathbf{y}\| \geq 0$*, with equality if and only if* $\mathbf{x} = \mathbf{y}$*;*

iii. $\|a\mathbf{x}\| = |a|\,\|\mathbf{x}\|$*;*

iv. $\|\mathbf{y} - \mathbf{x}\| = \|\mathbf{x} - \mathbf{y}\|$.

Remark A2.5 Before I give you the proof, I need to explain the term **with equality if and only if**. Part (i) is stating two things: first, that $\|\mathbf{x}\| \geq 0$ for all $\mathbf{x}$, and second, that $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$. And similarly for part (ii).

Sometimes, when we're talking about an inequality $p \geq q$, we say '**equality holds**' to mean that $p = q$.

Proof of Lemma A2.4 For (i), $\|\mathbf{x}\|$ is defined as the square root of the nonnegative real number $\sum_i x_i^2$, and the square root of a nonnegative real always means the *nonnegative* square root, so $\|\mathbf{x}\| \geq 0$. Now

$$\begin{aligned}
\|\mathbf{x}\| = 0 &\iff \sqrt{\sum_{i=1}^{n} x_i^2} = 0 \\
&\iff \sum_{i=1}^{n} x_i^2 = 0 \\
&\iff x_i^2 = 0 \text{ for all } i \in \{1, \ldots, n\} \qquad (\text{since } x_i^2 \geq 0 \text{ for each } i) \\
&\iff x_i = 0 \text{ for all } i \in \{1, \ldots, n\} \\
&\iff \mathbf{x} = \mathbf{0}.
\end{aligned}$$

So $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$, as claimed.

For (ii), simply replace $\mathbf{x}$ by $\mathbf{x} - \mathbf{y}$ in (i).

For (iii), we have

$$\|a\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} (ax_i)^2} = \sqrt{\sum_{i=1}^{n} a^2 x_i^2} = |a| \sqrt{\sum_{i=1}^{n} x_i^2} = |a| \, \|\mathbf{x}\|.$$

(Here we've been careful not to fall into a trap: $\sqrt{a^2}$ is $|a|$, not a!)

Finally, (iv) follows from (iii) by taking $a = -1$ and replacing $\mathbf{x}$ by $\mathbf{x} - \mathbf{y}$.□

A3 The dot and cross products

The dot product

Any two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ have a **dot product** or **scalar product** $\mathbf{x} \cdot \mathbf{y} \in \mathbb{R}$, defined by

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

Note that $\mathbf{x} \cdot \mathbf{y}$ is a scalar, not a vector—hence the name 'scalar product'.

The dot product is very important in mathematics, but it's a little bit subtle. It doesn't directly correspond to a geometric concept such as length or angle, although it's closely related to both. But you can write the length of any vector in terms of the dot product, as part (v) of the following lemma says.

Lemma A3.1 *Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ and $a \in \mathbb{R}$. Then:*

i. $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$*;*

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

iii. $\mathbf{x} \cdot \mathbf{0} = 0$;

iv. $\mathbf{x} \cdot (a\mathbf{y}) = a(\mathbf{x} \cdot \mathbf{y})$;

v. $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$.

Proof Again, these are routine checks using the definitions. I will do part (ii) and leave the rest to you.

For (ii), first note that both sides of the equation are scalars, so the equation does make sense. (Reminder: it is inconceivable that a vector could be equal to a scalar, or to a vector of different dimension. You should constantly keep your eye on this kind of thing.) Write $\mathbf{w} = \mathbf{y} + \mathbf{z}$. Then

$$\begin{aligned}
\mathbf{x} \cdot (\mathbf{y} + \mathbf{z}) &= \mathbf{x} \cdot \mathbf{w} && \text{by definition of } \mathbf{w} \\
&= \sum_{i=1}^{n} x_i w_i && \text{by definition of } \cdot \\
&= \sum_{i=1}^{n} x_i (y_i + z_i) && \text{by definition of vector addition} \\
&= \sum_{i=1}^{n} (x_i y_i + x_i z_i) && \text{since } p(q+r) = pq + pr \text{ for all } p, q, r \in \mathbb{R} \\
&= \sum_{i=1}^{n} x_i y_i + \sum_{i=1}^{n} x_i z_i && \text{by equation (A:3)} \\
&= \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \cdot \mathbf{z} && \text{by definition of } \cdot,
\end{aligned}$$

as required. □

This lemma shows that in many ways, the dot product behaves like ordinary multiplication of real numbers. *But* the big difference is that $\mathbf{x} \cdot \mathbf{y}$ is a different type of thing than $\mathbf{x}$ and $\mathbf{y}$: it's a scalar, whereas $\mathbf{x}$ and $\mathbf{y}$ are vectors. And multiple dot products such as '$\mathbf{x} \cdot \mathbf{y} \cdot \mathbf{z}$' simply do not make sense.

You may have encountered the fact that whenever $\mathbf{x}$ and $\mathbf{y}$ are vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$,

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta$$

where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$. Since $|\cos\theta| \leq 1$ for all θ, it follows that

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\| \tag{A:8}$$

for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^2$ or $\mathbb{R}^3$. (The argument above only applies to *nonzero* vectors, but it's also obviously true if $\mathbf{x}$ or $\mathbf{y}$ is $\mathbf{0}$.)

The inequality (A:8) turns out to be very important, so let's think about it further. When does equality hold? (This is always a good question to ask, whenever you meet an inequality.) That is, when is $|\mathbf{x} \cdot \mathbf{y}|$ *equal* to $\|\mathbf{x}\| \, \|\mathbf{y}\|$? Certainly equality holds if $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$. Assuming now that neither $\mathbf{x}$ nor $\mathbf{y}$ is $\mathbf{0}$, equality holds in (A:8) if and only if $|\cos\theta| = 1$, or equivalently $\cos\theta = \pm 1$. This means that the angle between $\mathbf{x}$ and $\mathbf{y}$ is 0 or π. So in summary: equality holds if and only if the points $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{y}$ are **collinear** (all lie on some straight line).

We now show that the inequality (A:8), and the condition for when equality holds, generalize without change from dimensions 2 and 3 to all higher dimensions:

Lemma A3.2 (Cauchy–Schwarz inequality) *For all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$$

with equality if and only if the points $\mathbf{0}$, $\mathbf{x}$ *and* $\mathbf{y}$ *are collinear.*

Proof If $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$ then both sides of the inequality are 0 and the points $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{y}$ are collinear. Now assume that $\mathbf{x} \neq \mathbf{0} \neq \mathbf{y}$.

The points $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{y}$ are collinear if and only if $\mathbf{x} = a\mathbf{y}$ for some $a \in \mathbb{R}$. If $\mathbf{x} = a\mathbf{y}$ then

$$\mathbf{x} \cdot \mathbf{y} = a\mathbf{y} \cdot \mathbf{y} = a\|\mathbf{y}\|^2,$$

and so $a = \mathbf{x} \cdot \mathbf{y}/\|\mathbf{y}\|^2$. Hence $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{y}$ are collinear if and only if

$$\mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y}. \tag{A:9}$$

We don't know whether $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{y}$ *are* collinear, but in any case, we can consider the distance between the left- and right-hand sides of (A:9). We have

$$\begin{aligned}
0 &\leq \left\| \mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y} \right\|^2 \\
&= \left(\mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y} \right) \cdot \left(\mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y} \right) \\
&= \|\mathbf{x}\|^2 - 2\frac{(\mathbf{x} \cdot \mathbf{y})^2}{\|\mathbf{y}\|^2} + \frac{(\mathbf{x} \cdot \mathbf{y})^2}{\|\mathbf{y}\|^2} \\
&= \frac{1}{\|\mathbf{y}\|^2}\Big(\|\mathbf{x}\|^2 \, \|\mathbf{y}\|^2 - (\mathbf{x} \cdot \mathbf{y})^2 \Big),
\end{aligned}$$

using Lemmas A2.4 and A3.1. So, rearranging,

$$(\mathbf{x} \cdot \mathbf{y})^2 \leq \|\mathbf{x}\|^2 \, \|\mathbf{y}\|^2.$$

Taking square roots on both sides gives

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|,$$

as required. Equality holds if and only if $\mathbf{x} - \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{y}\|^2}\mathbf{y} = \mathbf{0}$; but we have already shown that this is equivalent to the condition that $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{y}$ are collinear. □

This is a good example of how we can use our knowledge of $\mathbb{R}^2$ and $\mathbb{R}^3$ to guess (and then prove!) a fact about $\mathbb{R}^n$ for general n.

You may be familiar with the 'triangle inequality' in $\mathbb{R}^2$. This says that for any triangle, the length of each side is less than or equal to the sum of the lengths of the other two sides. (For instance, the distance from Edinburgh to Glasgow can't be more than the distance from Edinburgh to Stirling plus the distance from Stirling to Glasgow.) Now let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, and look back at Figure A.1(i), thinking about the triangle with vertices $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{x} + \mathbf{y}$.

- The distance from $\mathbf{0}$ to $\mathbf{x}$ is $\|\mathbf{x} - \mathbf{0}\| = \|\mathbf{x}\|$;

- the distance from $\mathbf{x}$ to $\mathbf{x}+\mathbf{y}$ is $\|(\mathbf{x}+\mathbf{y})-\mathbf{x}\| = \|\mathbf{y}\|$;
- the distance from $\mathbf{0}$ to $\mathbf{x}+\mathbf{y}$ is $\|(\mathbf{x}+\mathbf{y})-\mathbf{0}\| = \|\mathbf{x}+\mathbf{y}\|$.

So in this case, the triangle inequality states that $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\|+\|\mathbf{y}\|$. We have a geometrically plausible (although not quite rigorous) argument for why this should be true in $\mathbb{R}^2$, and maybe it's also clear in $\mathbb{R}^3$. But can you really claim that it's clearly true in $\mathbb{R}^{14382}$? Maybe not... but it *is* true, by the following algebraic argument.

Lemma A3.3 (Triangle inequality) *Let $n \geq 0$. For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,*

$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|.$$

Proof Using the Cauchy–Schwarz inequality,

$$\begin{aligned} \|\mathbf{x}+\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y})\cdot(\mathbf{x}+\mathbf{y}) \\ &= \|\mathbf{x}\|^2 + 2\mathbf{x}\cdot\mathbf{y} + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2|\mathbf{x}\cdot\mathbf{y}| + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2 \\ &= \big(\|\mathbf{x}\| + \|\mathbf{y}\|\big)^2. \end{aligned}$$

Taking square roots on both sides gives the result. □

Both this proof and the proof of the Cauchy–Schwarz inequality demonstrate an important lesson: it's often easier to work with *squares* of distances than with distances themselves, exploiting the fact that you can expand a squared distance $\|\mathbf{v}\|^2$ as a dot product $\mathbf{v}\cdot\mathbf{v}$.

Angles What is the angle between two nonzero vectors in $\mathbb{R}^n$? We know what 'angle' means when n is 2 or 3, but what does it mean for vectors in $\mathbb{R}^{14382}$? We don't even have a definition.

To answer this question, we will again take our inspiration from the 2- and 3-dimensional cases (just like when we defined *length* in $\mathbb{R}^n$). Recall that when $\mathbf{x}$ and $\mathbf{y}$ are nonzero vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, with angle θ between them, we have

$$\mathbf{x}\cdot\mathbf{y} = \|\mathbf{x}\|\,\|\mathbf{y}\|\cos\theta.$$

For nonzero $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, where n is arbitrary, we're going to *define* the angle between them to make this equation true. In other words:

Definition A3.4 Let $n \geq 0$ and let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ be nonzero vectors. The **angle** θ between $\mathbf{x}$ and $\mathbf{y}$ is defined to be

$$\theta = \cos^{-1}\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \in [0,\pi]. \qquad \text{(A:10)}$$

Here I am using the standard notation $[a,b] = \{x \in \mathbb{R} : a \leq x \leq b\}$.

Let's check that this 'definition' of angle really does make sense. The Cauchy–Schwarz inequality implies that $\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\,\|\mathbf{y}\|} \in [-1,1]$. Every number in

$[-1, 1]$ has infinitely many inverse cosines, but only one that belongs to $[0, \pi]$. So, the definition does indeed make sense. And it immediately implies that

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta$$

for all nonzero $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, extending the pattern we've already seen for $n = 2$ and $n = 3$.

Example A3.5 Let $\mathbf{x} = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \\ \sqrt{12} \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} -\sqrt{2} \\ -\sqrt{2} \\ \sqrt{12} \end{pmatrix}$. Then $\mathbf{x} \cdot \mathbf{y} = 8$ and $\|\mathbf{x}\| = \|\mathbf{y}\| = 4$. Hence the angle θ between $\mathbf{x}$ and $\mathbf{y}$ is given by

$$\theta = \cos^{-1} \frac{8}{4 \cdot 4} = \cos^{-1} \frac{1}{2} = \pi/3.$$

Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are said to be **orthogonal** if $\mathbf{x} \cdot \mathbf{y} = 0$. (Other ways of saying orthogonal are 'perpendicular' and 'at right angles'.) This happens exactly when $\mathbf{x} = \mathbf{0}$ or $\mathbf{y} = \mathbf{0}$ or the angle between $\mathbf{x}$ and $\mathbf{y}$ is $\pi/2$.

Lemma A3.6 (Pythagoras) *Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ be orthogonal vectors. Then*

$$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2.$$

(Draw a picture to see why I've attributed this result to Pythagoras!)

Proof Using the hypothesis that $\mathbf{x} \cdot \mathbf{y} = 0$, we have

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) \\ &= \mathbf{x} \cdot \mathbf{x} + 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y} \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2, \end{aligned}$$

as required. □

The cross product

Given two vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^3$, there is usually a unique direction orthogonal to both of them. (I say 'usually' because it's not unique if $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{y}$ happen to be collinear.) This is geometrically clear, but if I gave you two *specific* vectors, in terms of their coordinates, you might not find it so easy to write down another vector orthogonal to them both. However, the cross product provides a very easy way of doing exactly that.

Definition A3.7 Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$. The **cross product** or **vector product** $\mathbf{x} \times \mathbf{y} \in \mathbb{R}^3$ is defined by

$$\mathbf{x} \times \mathbf{y} = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}.$$

As the alternative name suggests, $\mathbf{x} \times \mathbf{y}$ is a vector, not a scalar (unlike the dot product). Note also that the cross product is only defined for three-dimensional vectors, not in $\mathbb{R}^n$ for arbitrary n (unlike the dot product). In this course, three-dimensional space plays no special role, so we'll hardly use the cross product at all. However, you'll need it for other courses.

The last part of the following lemma states that $\mathbf{x} \times \mathbf{y}$ is indeed orthogonal to $\mathbf{x}$ and $\mathbf{y}$. The other parts state other basic facts about the cross product.

Lemma A3.8 *Let* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^3$ *and* $a \in \mathbb{R}$. *Then:*

i. $\mathbf{x} \times \mathbf{y} = -(\mathbf{y} \times \mathbf{x})$*;*

ii. $\mathbf{x} \times \mathbf{x} = \mathbf{0}$*;*

iii. $\mathbf{x} \times (\mathbf{y} + \mathbf{z}) = (\mathbf{x} \times \mathbf{y}) + (\mathbf{x} \times \mathbf{z})$*;*

iv. $\mathbf{x} \times (a\mathbf{y}) = a(\mathbf{x} \times \mathbf{y})$*;*

v. $(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x} = 0 = (\mathbf{x} \times \mathbf{y}) \cdot \mathbf{y}$.

Proof Again, these are routine algebraic checks that you should try for yourself. I'll just prove the first part of (v): $(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x} = 0$. We have

$$\begin{aligned}
(\mathbf{x} \times \mathbf{y}) \cdot \mathbf{x} &= \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\
&= (x_2y_3 - x_3y_2)x_1 + (x_3y_1 - x_1y_3)x_2 + (x_1y_2 - x_2y_1)x_3 \\
&= x_1x_2y_3 - x_1x_3y_2 + x_2x_3y_1 - x_1x_2y_3 + x_1x_3y_2 - x_2x_3y_1 \\
&= 0,
\end{aligned}$$

as required. □

Example A3.9 As in Example A3.5, let $\mathbf{x} = \begin{pmatrix} \sqrt{2} \\ \sqrt{2} \\ \sqrt{12} \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} -\sqrt{2} \\ -\sqrt{2} \\ \sqrt{12} \end{pmatrix}$. Then

$$\mathbf{x} \times \mathbf{y} = \begin{pmatrix} \sqrt{2}\sqrt{12} + \sqrt{2}\sqrt{12} \\ -\sqrt{12}\sqrt{2} - \sqrt{2}\sqrt{12} \\ -\sqrt{2}\sqrt{2} + \sqrt{2}\sqrt{2} \end{pmatrix} = \begin{pmatrix} 2\sqrt{24} \\ -2\sqrt{24} \\ 0 \end{pmatrix}.$$

And this vector is indeed orthogonal to $\mathbf{x}$ and $\mathbf{y}$; that is, its dot product with $\mathbf{x}$ and its dot product with $\mathbf{y}$ are both zero, as you can check.

So we know about the *direction* of $\mathbf{x} \times \mathbf{y}$. What about its length? The following result gives the answer.

Lemma A3.10 *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be nonzero vectors in* $\mathbb{R}^3$. *Then*

$$\|\mathbf{x} \times \mathbf{y}\| = \|\mathbf{x}\| \, \|\mathbf{y}\| \sin\theta$$

where $\theta \in [0, \pi]$ *is the angle between* $\mathbf{x}$ *and* $\mathbf{y}$.

Proof We have

$$\begin{aligned}
\|\mathbf{x} \times \mathbf{y}\|^2 + (\mathbf{x} \cdot \mathbf{y})^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\
&\quad + (x_1y_1 + x_2y_2 + x_3y_3)^2 \\
&= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) \\
&= \|\mathbf{x}\|^2 \|\mathbf{y}\|^2
\end{aligned}$$

where the second equality is a routine calculation. Hence

$$\begin{aligned}\|\mathbf{x}\times\mathbf{y}\|^2 &= \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - (\mathbf{x}\cdot\mathbf{y})^2\\ &= \|\mathbf{x}\|^2\|\mathbf{y}\|^2 - (\|\mathbf{x}\|\|\mathbf{y}\|\cos\theta)^2\\ &= \big(\|\mathbf{x}\|\|\mathbf{y}\|\sin\theta\big)^2.\end{aligned}$$

Finally, $\sin\theta \geq 0$ for all $\theta \in [0, \pi]$, so taking square roots on both sides gives the result. □

This is a 'rabbit-out-of-a-hat' proof. When a magician pulls a rabbit from a hat, the audience's reaction is 'where did that come from?'. That might also be your reaction when you read a proof like this. In this case, it's just a long algebraic calculation that's been packaged up in a neat way. Once again, it exploits the fact that *squared* distances are easier to work with than actual distances.

Example A3.11 Let's verify Lemma A3.10 for the vectors in Example A3.9. On the one hand,

$$\|\mathbf{x}\times\mathbf{y}\| = \left\|\begin{pmatrix}2\sqrt{24}\\-2\sqrt{24}\\0\end{pmatrix}\right\| = \sqrt{(2\sqrt{24})^2+(2\sqrt{24})^2} = \sqrt{2\times 2^2\times 24} = \sqrt{192}.$$

On the other,

$$\|\mathbf{x}\|\,\|\mathbf{y}\|\sin\theta = 4\cdot 4\cdot\sin(\pi/3) = 16\cdot\sqrt{3}/2 = 8\sqrt{3} = \sqrt{8^2\times 3} = \sqrt{192}$$

(using the values of $\|\mathbf{x}\|$, $\|\mathbf{y}\|$ and θ that we found in Example A3.5). So in this case, Lemma A3.10 states that $\sqrt{192} = \sqrt{192}$…which is true!

A4 Matrices

Take integers $m, n \geq 0$. An $m \times n$ real **matrix** consists of a real number a_{ij} for each $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. We visualize them arranged in a grid with m rows and n columns:

$$\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}.$$

Let us call this matrix A. We refer to a_{ij} as the **(i, j)-entry** of A. Sometimes we write

$$A = (a_{ij})_{1\leq i\leq m,\, 1\leq j\leq n}$$

to mean that A is an $m\times n$ matrix with (i, j)-entry called a_{ij}. But more often, we just write $A = (a_{ij})$. This is a little bit ambiguous; brackets in mathematics are used for multiple purposes! But you'll see this notation a lot, so it's important to get used to it.

Often it's useful to adopt a different convention, where the (i, j)-entry of a matrix A is written as A_{ij}, the (i, j)-entry of B is written as B_{ij}, and so on.

For instance, this would mean that the (i,j)-entry of the matrix $(A+B)C$ is written as $((A+B)C)_{ij}$. (The meaning of $(A+B)C$ will be explained soon.) Whenever I'm about to use this convention, I'll say so.

An m-dimensional vector is just an $m \times 1$ matrix. For us, vectors are by default *column* vectors, but you can also consider *row* vectors. By definition, an n-dimensional **row vector** is a $1 \times n$ matrix.

A 1×1 matrix is just a scalar (a real number). You might argue that a 1×1 matrix is really a scalar *with a pair of brackets around it*, and you'd be right, but we won't worry about the difference!

Since a vector is a special kind of matrix, it's a little inconsistent that we write vectors in bold typeface but matrices in ordinary typeface. However, it's a common convention and we'll stick with it.

New matrices from old

There are several algebraic operations on matrices. To define them, I'll use the convention that the (i,j)-entry of a matrix M is written as M_{ij}.

- **Addition**. Given $m \times n$ matrices A and B, we define an $m \times n$ matrix $A+B$ by
$$(A+B)_{ij} = A_{ij} + B_{ij}$$
$(1 \leq i \leq m,\ 1 \leq j \leq n)$. That is, the (i,j)-entry of $A+B$ is the (i,j)-entry of A plus the (i,j)-entry of B.

  Two matrices can only be added if they have the same number of rows and columns.

- **Scalar multiplication**. Given an $m \times n$ matrix A and a scalar $c \in \mathbb{R}$, we define an $m \times n$ matrix cA by
$$(cA)_{ij} = c \cdot A_{ij}$$
$(1 \leq i \leq m,\ 1 \leq j \leq n)$. That is, the (i,j)-entry of cA is c times the (i,j)-entry of A.

- **Matrix multiplication**. Given an $m \times n$ matrix A and an $n \times p$ matrix B, we define an $m \times p$ matrix AB by
$$(AB)_{ik} = \sum_{j=1}^{n} A_{ij} B_{jk}$$
$(1 \leq i \leq m,\ 1 \leq k \leq p)$. That is, the (i,k)-entry of AB is
$$A_{i1}B_{1k} + A_{i2}B_{2k} + \cdots + A_{in}B_{nk}.$$
Two matrices can only be multiplied if the number of columns in the first is equal to the number of rows in the second.

One particularly important matrix is the $m \times n$ matrix all of whose entries are 0. We call this matrix 0, too. Another important matrix is the $n \times n$

identity matrix

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

When it's clear which n is meant, we write I_n as just I.

Since vectors are matrices of a special kind, you can add up vectors of the same dimension and multiply them by scalars. This is exactly the same as the addition and scalar multiplication of vectors that we met in the last section.

But matrix multiplication also gives us something new to do with vectors. If A is an $m \times n$ matrix and $\mathbf{x}$ is an n-dimensional vector (thought of as an $n \times 1$ matrix), the definition above gives us the matrix product $A\mathbf{x}$. This is an $m \times 1$ matrix, that is, an m-dimensional vector. It has entries

$$(A\mathbf{x})_i = \sum_{j=1}^{n} A_{ij} x_j \tag{A:11}$$

(for $1 \leq i \leq m$).

Here's a crucial point about matrices:

Matrices of different sizes can't be equal!

By 'size' I mean the numbers of rows and columns. For instance, a 3×2 matrix stands no chance of being equal to a 2×3 matrix, because they have different sizes.

So if you're given matrices A and B and asked to prove that they're equal, you have to do two things:

- prove that they have the same size;
- prove that they have the same entries.

We'll see an example of this in the proof of the following lemma, which lists some fundamental facts about matrix algebra.

Lemma A4.1 *i. $(A+B)+C = A+(B+C)$ for any $m \times n$ matrices A, B and C;*

ii. $A+B = B+A$ for any $m \times n$ matrices A and B;

iii. $A+0 = A$ for any matrix A;

iv. $c(A+B) = cA+cB$ for any $m \times n$ matrices A and B and scalar c:

v. $(AB)C = A(BC)$ for any $m \times n$ matrix A, $n \times p$ matrix B, and $p \times q$ matrix C;

vi. $AI_n = A = I_m A$ for any $m \times n$ matrix A;

vii. $A(B+C) = AB+AC$ for any $m \times n$ matrix A and $n \times p$ matrices B and C;

viii. $(A+B)C = AC+BC$ *for any* $m \times n$ *matrices* A *and* B *and* $n \times p$ *matrix* C*;*

ix. $c(AB) = (cA)B = A(cB)$ *for any* $m \times n$ *matrix* A*,* $n \times p$ *matrix* B*, and scalar* c*.*

Proof Again, these are just algebraic checks using the definitions. I'll prove (v) as an example, using the convention that the (i, j)-entry of a matrix M is written as M_{ij}. You should try some others yourself.

To prove that $(AB)C = A(BC)$, we have to show first that the two matrices have the same size, and second that they have the same entries.

First, A is an $m \times n$ matrix and B is an $n \times p$ matrix, so AB is an $m \times p$ matrix. Also C is a $p \times q$ matrix, so $(AB)C$ is an $m \times q$ matrix. On the other hand, B is $n \times p$ and C is $p \times q$, so BC is $n \times q$. Also A is $m \times n$, so $A(BC)$ is $m \times q$. We have now shown that $(AB)C$ and $A(BC)$ are both $m \times q$ matrices, so they are the same size.

Second, we have to show that the $m \times q$ matrices $(AB)C$ and $A(BC)$ have the same entries. So, let $i \in \{1, \ldots, m\}$ and $\ell \in \{1, \ldots, q\}$; we have to show that $((AB)C)_{i\ell} = (A(BC))_{i\ell}$. On the one hand,

$$\begin{aligned} ((AB)C)_{i\ell} &= \sum_{k=1}^{p} (AB)_{ik} C_{k\ell} && \text{by definition of matrix multiplication} \\ &= \sum_{k=1}^{p} \left(\sum_{j=1}^{n} A_{ij} B_{jk} \right) C_{k\ell} && \text{by definition of matrix multiplication} \\ &= \sum_{k=1}^{p} \sum_{j=1}^{n} A_{ij} B_{jk} C_{k\ell}. \end{aligned}$$

On the other hand,

$$\begin{aligned} (A(BC))_{i\ell} &= \sum_{j=1}^{n} A_{ij} (BC)_{j\ell} && \text{by definition of matrix multiplication} \\ &= \sum_{j=1}^{n} A_{ij} \sum_{k=1}^{p} B_{jk} C_{k\ell} && \text{by definition of matrix multiplication} \\ &= \sum_{j=1}^{n} \sum_{k=1}^{p} A_{ij} B_{jk} C_{k\ell}. \end{aligned}$$

(Reminder: if you're in doubt about any of these steps, write it out in full using $+$ and $\cdots$ signs instead of summation notation.) And since we can change the order of summation, it follows that

$$((AB)C)_{i\ell} = (A(BC))_{i\ell},$$

as required. □

Lemma A4.1 is a result about matrices in general, but we can apply it in the special case where some of the matrices are vectors. For instance, it tells us that

$$A(\mathbf{x}+\mathbf{y}) = A\mathbf{x} + A\mathbf{y}, \qquad A(c\mathbf{x}) = cA\mathbf{x}, \qquad (A+B)\mathbf{x} = A\mathbf{x} + B\mathbf{x} \tag{A:12}$$

whenever A, B, $\mathbf{x}$ and $\mathbf{y}$ are matrices and vectors of the appropriate sizes, and c is a scalar.

Warning A4.2 Matrix multiplication is not **commutative**! That is, it's *not* always true that $AB = BA$ for matrices A and B. For example, you can check that if

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

then $AB = A$ and $BA = B$, so $AB \neq BA$. Sometimes, the order that we do things in matters. Opening a window and sticking your head out is quite different from sticking your head out of the window then opening it.

In this course, I'll write $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ for the vectors in $\mathbb{R}^n$ defined by

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{A:13}$$

(Beware that not everyone uses this notation.) I'll also write

$$A = (\mathbf{x}_1 | \mathbf{x}_2 | \cdots | \mathbf{x}_n)$$

to mean that the columns of the matrix A are $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$. For instance, if A is an $m \times n$ matrix with (i, j)-entry written as A_{ij}, then $\mathbf{x}_1 \in \mathbb{R}^m$ is given by

$$\mathbf{x}_1 = \begin{pmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{m1} \end{pmatrix}.$$

Here's a lemma that will be repeatedly useful later on.

Lemma A4.3 *Let $A = (\mathbf{x}_1 | \mathbf{x}_2 | \cdots | \mathbf{x}_n)$ be an $m \times n$ matrix. Then:*

i. $A\mathbf{y} = y_1\mathbf{x}_1 + y_2\mathbf{x}_2 + \cdots + y_n\mathbf{x}_n$ for any vector $\mathbf{y} \in \mathbb{R}^n$;

ii. $A\mathbf{e}_j = \mathbf{x}_j$ for any $j \in \{1, \dots, n\}$; that is, $A\mathbf{e}_j$ is the jth column of A;

iii. $A(\mathbf{y}_1 | \mathbf{y}_2 | \cdots | \mathbf{y}_p) = (A\mathbf{y}_1 | A\mathbf{y}_2 | \cdots | A\mathbf{y}_p)$ for any $\mathbf{y}_1, \dots, \mathbf{y}_p \in \mathbb{R}^n$.

Proof I will prove (i) and (ii), and leave (iii) to you.

For (i), we first have to show that the two sides are the same size. On the left, A is an $m \times n$ matrix and $\mathbf{y}$ is an $n \times 1$ matrix, so $A\mathbf{y}$ is an $m \times 1$ matrix, that is, an m-dimensional vector. On the right, each y_j is a scalar and each $\mathbf{x}_j$ is an m-dimensional vector, so each $y_j\mathbf{x}_j$ is an m-dimensional vector; hence $\sum_{j=1}^n y_j\mathbf{x}_j$ is an m-dimensional vector too. So both sides are m-dimensional vectors.

Now we have to check that these two vectors have the same entries. Let $1 \leq i \leq m$. Then, using equation (A:11) (page 33), the ith entry of $A\mathbf{y}$ is

$$(A\mathbf{y})_i = \sum_{j=1}^n A_{ij} y_j.$$

On the other hand, the ith entry of $\mathbf{x}_1$ is A_{i1}, so the ith entry of $y_1\mathbf{x}_1$ is y_1A_{i1}, and similarly for $y_2\mathbf{x}_2, \dots, y_n\mathbf{x}_n$. Hence the ith entry of $y_1\mathbf{x}_1 + y_2\mathbf{x}_2 + \cdots + y_n\mathbf{x}_n$ is

$$y_1A_{i1} + y_2A_{i2} + \cdots + y_nA_{in} = \sum_{j=1}^{n} A_{ij}y_j.$$

So the ith entries of the two sides are equal, as required.

To prove (ii), just put $\mathbf{y} = \mathbf{e}_j$ in (i). □

A5 Inverse and transpose matrices

Inverses

Definition A5.1 An $m \times n$ matrix A is **invertible** if there exists an $n \times m$ matrix B such that $AB = I_m$ and $BA = I_n$. It is called **singular** if it is not invertible.

In the definition of invertibility, *both equations are needed.* It's possible to find an example of a matrix A such that there does exist a B satisfying $AB = I$, but there does not exist a B satisfying $BA = I$.

Let A be an invertible $m \times n$ matrix. Then there can be only one matrix B such that $AB = I_m$ and $BA = I_n$, since if B' is another one then

$$B = BI_m = B(AB') = (BA)B' = I_nB' = B'.$$

We call this matrix B the **inverse** of A, and write B as A^{-1}.

Examples A5.2 i. The matrix $\begin{pmatrix} 1 & -2 \\ -3 & 5 \end{pmatrix}$ is invertible, with inverse $\begin{pmatrix} -5 & -2 \\ -3 & -1 \end{pmatrix}$. To prove this, you have to check that

$$\begin{aligned} \begin{pmatrix} 1 & -2 \\ -3 & 5 \end{pmatrix}\begin{pmatrix} -5 & -2 \\ -3 & -1 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \text{and} \quad \begin{pmatrix} -5 & -2 \\ -3 & -1 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -3 & 5 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

This check can be made quicker by using Theorem A5.3(ii) below.

ii. Take scalars $a, b, c \neq 0$. Then

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}^{-1} = \begin{pmatrix} 1/a & 0 & 0 \\ 0 & 1/b & 0 \\ 0 & 0 & 1/c \end{pmatrix}.$$

Exercise: check that this is true, as in (i).

iii. The matrix $A = \begin{pmatrix} 1 & -2 & 4 \\ -3 & 5 & 6 \end{pmatrix}$ is not invertible. To verify this, you have to show that there is no matrix $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$ such that $AB = I_2$ and $BA = I_3$. This can be done by a long and tedious calculation—but also follows immediately from Theorem A5.3(i) below.

Neither part of the following theorem is obvious. (If you think it is, try proving it!) We'll only prove it later, once we've developed some theory that will make it easy.

Theorem A5.3 *i. Every invertible matrix is square. That is, if A is an invertible $m \times n$ matrix then $m = n$.*

ii. Let A and B be $n \times n$ matrices. Then $AB = I_n \iff BA = I_n$.

Part (ii) is only true for *square* matrices. For instance, you should be able to find a 1×2 matrix A and a 2×1 matrix B with $AB = I_1$ but $BA \neq I_2$. The truth of part (ii) is not obvious even for 2×2 matrices. That is, can you prove directly that if A and B are 2×2 matrices such that $AB = I$, then also $BA = I$? If you can do it for 2×2, can you do it for 3×3, or, generally, $n \times n$? I know of no easy way.

Example A5.4 Consider a 2×2 matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. If $ad - bc \neq 0$ then A is invertible and

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Write B for the right-hand side. To check that B really is the inverse of A, in principle we need to check that $AB = I_2$ and $BA = I_2$. However, Theorem A5.3(ii) means that we only need to check one of these. This is straightforward (try it!).

The number $ad - bc$ is called the **determinant** of A. (Questions: if $ad - bc = 0$, is A still invertible? If so, why? If not, why not?) We'll do more on determinants later in the course.

Lemma A5.5 *i. Let A and B be invertible $n \times n$ matrices. Then AB is also invertible, and $(AB)^{-1} = B^{-1}A^{-1}$.*

ii. The identity matrix I_n is invertible, with inverse I_n.

Proof For (i), we have

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.$$

Also $(B^{-1}A^{-1})(AB) = I_n$, either by a similar argument or by Theorem A5.3(ii). So AB is invertible (by definition of invertibility), with inverse $B^{-1}A^{-1}$ (by definition of inverse).

Part (ii) follows immediately from the fact that $I_nI_n = I_n$, which is a special case of Lemma A4.1(vi). □

Warning A5.6 In Lemma A5.5(i), note the reversal of order! The inverse of AB is $B^{-1}A^{-1}$, not $A^{-1}B^{-1}$. It's like this: at the beginning of the day, you put on your socks and then put on your shoes. At the end of the day, you take off your shoes and then take off your socks—not the other way round!

Transposes

The **transpose** of an $m \times n$ matrix A is the $n \times m$ matrix A^T whose (j, i)-entry is the (i, j)-entry of A (for $1 \leq i \leq m$, $1 \leq j \leq n$). For instance,

$$\begin{pmatrix} 1 & -2 & 4 \\ -3 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & -3 \\ -2 & 5 \\ 4 & 6 \end{pmatrix}.$$

So a matrix and its transpose have different sizes, unless the matrix is square.

A square matrix $A = (a_{ij})$ is **symmetric** if $A^T = A$, that is, $a_{ji} = a_{ij}$ for all i and j. It is **skew symmetric** (or **antisymmetric**) if $A^T = -A$, that is, $a_{ji} = -a_{ij}$ for all i and j. For instance, the first of these matrices is symmetric and the second is skew symmetric:

$$\begin{pmatrix} 4 & -7 & 2 \\ -7 & 99 & 12 \\ 2 & 12 & -10 \end{pmatrix}, \qquad \begin{pmatrix} 0 & -7 & 2 \\ 7 & 0 & 12 \\ -2 & -12 & 0. \end{pmatrix}.$$

The diagonal entries a_{ii} of a skew symmetric matrix $A = (a_{ij})$ must all be 0, since $a_{ii} = -a_{ii}$ for all i.

We can express the dot product in terms of matrix multiplication and transpose:

Lemma A5.7 *Let* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. *Then* $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T \mathbf{y}$.

Proof The right-hand side is the 1×1 matrix

$$\begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = (x_1 y_1 + x_2 y_2 + \cdots + x_n y_n).$$

But a 1×1 matrix is just a scalar, and the scalar here is exactly the dot product $\mathbf{x} \cdot \mathbf{y}$. □

Here are some further useful properties of transposes.

Lemma A5.8 *i.* $(A + B)^T = A^T + B^T$ *for all* $m \times n$ *matrices* A *and* B;

ii. $(cA)^T = cA^T$ *for every matrix* A *and scalar* c;

iii. $(AB)^T = B^T A^T$ *for every* $m \times n$ *matrix* A *and* $n \times p$ *matrix* B.

iv. $(A^T)^T = A$

Proof Once again, I will leave most of these to you, just giving you the proof of the hardest one: in this case, (iii). Again, I will use the convention that the (i, j)-entry of a matrix M is written as M_{ij}.

First we have to show that the two matrices have the same size. On the left-hand side, A is an $m \times n$ matrix and B is an $n \times p$ matrix, so AB is an $m \times p$ matrix, so $(AB)^T$ is a $p \times m$ matrix. On the right, B^T is $p \times n$ and A^T is $n \times m$, so $B^T A^T$ is $p \times m$. So, both sides are $p \times m$ matrices.

Second, we show that the entries of these two $p \times m$ matrices are equal. Let $k \in \{1, \ldots, p\}$ and $i \in \{1, \ldots, m\}$. On the left-hand side,

$$\begin{aligned} ((AB)^T)_{ki} &= (AB)_{ik} && \text{by definition of transpose} \\ &= \sum_{j=1}^{n} A_{ij} B_{jk} && \text{by definition of matrix multiplication.} \end{aligned}$$

On the right-hand side,

$$\begin{aligned} (B^T A^T)_{ki} &= \sum_{j=1}^{n} (B^T)_{kj} (A^T)_{ji} && \text{by definition of matrix multiplication} \\ &= \sum_{j=1}^{n} B_{jk} A_{ij} && \text{by definition of transpose.} \end{aligned}$$

So the (k, i)-entries of $(AB)^T$ and $B^T A^T$ are equal, as required. □

Warning A5.9 Just as for inverses, note the reversal of order in (iii): $(AB)^T$ is $B^T A^T$, not $A^T B^T$. Transpose is not the same as inverse, but in some respects the two operations behave similarly.

Our final lemma connects these two concepts, transpose and inverse:

Lemma A5.10 *Let A be an invertible matrix. Then A^T is also invertible, and $(A^T)^{-1} = (A^{-1})^T$.*

Proof By definition of invertibility and of inverse, it is enough to show that $(A^{-1})^T A^T = I$ and $A^T (A^{-1})^T = I$. For the first, we have

$$\begin{aligned} (A^{-1})^T A^T &= (AA^{-1})^T && \text{by Lemma A5.8(iii)} \\ &= I^T && \text{by definition of inverse} \\ &= I. \end{aligned}$$

The proof of the second equation is very similar. □

A6 Complex numbers

Almost everything in this course is about the *real* numbers. However, it's an amazing fact—first discovered in 16th century Italy—that some statements about real numbers are most easily proved using *complex* numbers. We'll meet an example of this phenomenon right at the end of the course. In preparation, we now gather together some of the most important facts about the complex numbers.

To obtain $\mathbb{C}$ from $\mathbb{R}$, we begin by adjoining to $\mathbb{R}$ a new element i with the property that $i^2 = -1$. Because we want to be able to add and multiply complex numbers, we also have to throw in elements $a + bi$ for each $a, b \in \mathbb{R}$. Every complex number looks like this. More precisely, for every complex number z, there are unique real numbers a and b such that $z = a + bi$.

We constructed $\mathbb{C}$ deliberately so that the equation $x^2 + 1 = 0$, which has no solution in $\mathbb{R}$, *does* have a solution in $\mathbb{C}$. We simply invented (or 'imagined')

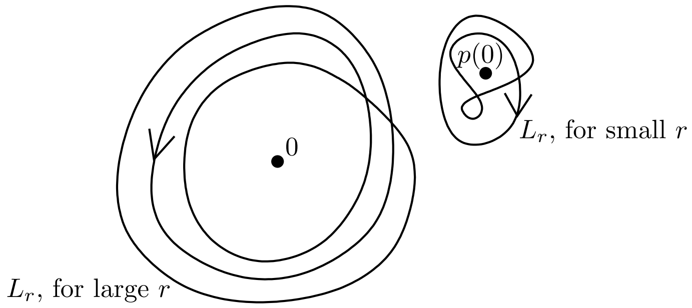


Figure A.2: The proof of the fundamental theorem of algebra, shown for a polynomial of degree 3

a solution i. But there are lots of other polynomial equations that have no solution in $\mathbb{R}$, such as $x^6 + 1 = 0$ or $2x^4 + x + 1 = 0$. What if we try to perform the same trick again, expanding $\mathbb{C}$ further so that it contain solutions to these equations too?

The miracle is that *we don't need to.* Although $\mathbb{C}$ was only designed in order to contain a solution to $x^2 + 1 = 0$, it actually **already contains solutions to all other polynomial equations!** This fact is so miraculous and wonderful and unexpected that it has a very grand name:

Theorem A6.1 (Fundamental theorem of algebra) *Every non-constant polynomial over $\mathbb{C}$ has at least one root in $\mathbb{C}$.*

Proof (sketch; non-examinable) There are many known proofs of this theorem, but none is very simple as far as I know. Here is an outline of my favourite. If you want to know how to make it precise, you should take the 4th year course Algebraic Topology.

Let $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$, with $n \geq 1$ and $a_n \neq 0$. Suppose for a contradiction that p has no root.

For $r \geq 0$, write C_r for the circle in the complex plane with centre 0 and radius r. As z moves one revolution around C_r, $p(z)$ traces out a loop L_r in $\mathbb{C}$. It cannot pass *through* 0, since p has no root, but we can ask whether L_r *winds around* 0.

(See Figure A.2. Picture a pole sticking up at 0 and the loop L_r as made of string. The question is whether the string is wound around the pole.)

When $r = 0$, the loop L_r is constant at $p(0)$, so it does not wind around 0. When r is small, the whole loop L_r lies close to $p(0)$, so again it does not wind around 0. As r gradually increases, L_r gradually changes position, and so L_r cannot wind around 0 for any value of r. (It takes some work to make that step precise, but the intuitive idea is that if L_r doesn't wind around 0 and L_s does, then L_t must actually pass *through* 0 for some t between r and s, contradicting our hypothesis.)

However, when r is large, $p(z)$ behaves roughly like its leading term $a_n z^n$, so the loop L_r winds n times round 0. Since $n \neq 0$, this is a contradiction. □

Corollary A6.2 *Let $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ be a polynomial over $\mathbb{C}$. Then*

$$p(z) = a_n (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n)$$

for some $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$.

Proof (sketch; non-examinable) This follows by induction from Theorem A6.1. □

This corollary fails if you replace $\mathbb{C}$ by $\mathbb{R}$. For example, the polynomial $p(z) = z^2 + 1$ cannot be written as $(z - \lambda_1)(z - \lambda_2)$ for any *real* λ_1 and λ_2, since if it could be, λ_1 and λ_2 would be real square roots of -1. You might think that the complex numbers are harder than the real numbers, but results like this mean that in some ways, they're easier.

Coming down from the high of these towering results, we finish by summarizing some of the facts about complex conjugates.

Let z be a complex number. Then $z = x + iy$ for unique $x, y \in \mathbb{R}$. The **complex conjugate** of z is defined to be $\overline{z} = x - iy$. Graphically, complex conjugation is reflection in the real axis. A complex number z is real if and only if $z = \overline{z}$ (that is, z lies on the axis of reflection).

A fundamental feature of complex conjugation is that it preserves addition and multiplication:

$$\overline{z + w} = \overline{z} + \overline{w}, \qquad \overline{z \cdot w} = \overline{z} \cdot \overline{w} \tag{A:14}$$

for all $z, w \in \mathbb{C}$. You should check these properties yourself if you haven't done so before.

Another convenient feature is that the modulus $|z| = \sqrt{x^2 + y^2}$ of a complex number $z = x + iy$ can be expressed in terms of conjugates:

$$|z| = \sqrt{z\overline{z}}. \tag{A:15}$$

Geometrically, $|z|$ is the distance between 0 and z. Much as in Lemma A2.4(i), modulus has the property that $|z| \geq 0$ for all z, with $|z| = 0 \iff z = 0$.

Beginners tend to express every complex number in terms of its real and imaginary parts: $z = x+iy$. Actually, it's often most graceful to avoid doing this and think in terms of conjugates instead, only going to the real and imaginary parts as a last resort.

Next time: we lay our hands on the slippery concept of dimension.

Summary of Chapter A

This is for you to fill in.

The most important definitions and ideas in this chapter

The most important results in this chapter

Points I didn't understand

Chapter B

Dimension

To be read before the lecture of Monday, 8 October 2018

There is a romantic image of mathematicians as people who spend their time struggling to solve immensely difficult problems, maybe trying to prove some conjecture that was made hundreds of years ago, maybe trying to solve some more recent problem. It's true that we spend a lot of our time trying to solve difficult problems, but often the problem is not to find a proof: it's to find the right *definition* of something. There have been countless historical episodes where finding the right definition has taken decades of work by dozens of people.

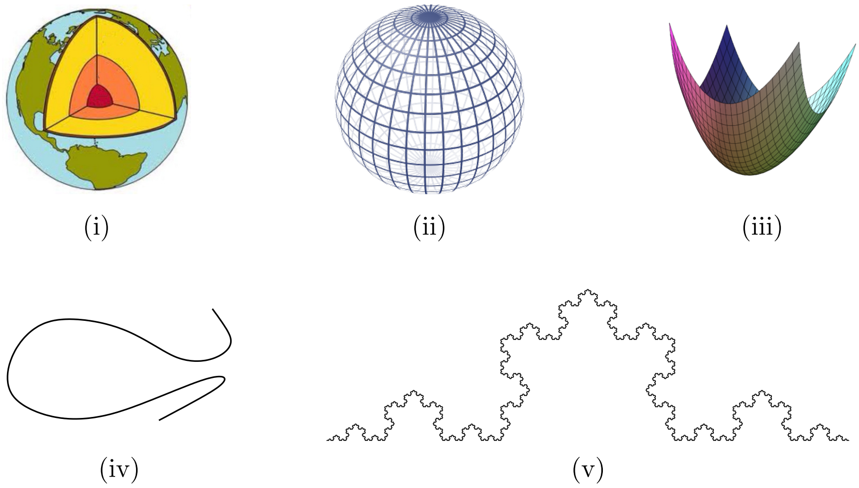
What do I mean by the 'right' definition? A definition can't really be right or wrong, but what I mean here is a definition that is simple and elegant, that captures the examples that ought to be captured, and that gives rise to useful theorems. The challenge is to turn vague intuition into precise mathematics.

This chapter is about the quest for the definition of dimension. We all know *roughly* what it means for a shape to be 1-, 2- or 3-dimensional, but what does it mean *precisely*? And what about dimensions higher than 3?

Consider the shapes in Figure B.1. Shape (i) is obviously 3-dimensional. Shape (ii) could be interpreted as either a solid or a hollow shape. If it's solid (like planet Earth) then it's 3-dimensional. If it's hollow (like the surface of the Earth) then a mathematician would say that it's 2-dimensional. You might want to call it 3-dimensional, because it lives in $\mathbb{R}^3$, but the reason why it's 2-dimensional is that any point on the surface of the Earth can be specified by just 2 coordinates, longitude and latitude.

For similar reasons, the surface in (iii) is said to be 2-dimensional and the curve in (iv) is 1-dimensional. Shape (v) is called the Koch curve (and if you glue three of them together to make a star-like shape, what you get is called the Koch snowflake). Its dimension is $(\log 4)/(\log 3) = 1.261\ldots$. Obviously that can't be explained in terms of 'how many coordinates you need in order to specify a point on it', as it's not a whole number. Explaining why its dimension is $(\log 4)/(\log 3)$ is not too hard, but since it has nothing to do with linear algebra, I'll leave it as a mystery. (You can easily find the answer on the web.) Roughly speaking, the Koch curve is too wiggly to be 1-dimensional, but not substantial enough to be 2-dimensional, so its dimension is between 1 and 2.

A great deal can be said about definitions of dimension, far more than would



(i) (ii) (iii)

(iv) (v)

Figure B.1: Shapes of various dimensions. (Image credits: (i), (ii), (v) Wikimedia Commons; (iii) thephysicsmill.com.)

fit into a course of this length even if we studied nothing else. We are going to restrict ourselves to something very humble. All we will study is the dimension of flat, linear shapes: lines and planes and their higher-dimensional cousins.

We will use the 'longitude and latitude' idea: the dimension of a shape should be how many numbers you need in order to specify the position of a point in that shape. But it's not so obvious how to take the idea of 'how many numbers do you need?' and make it precise. For instance, it was shown over a century ago that there is a bijection (one-to-one correspondence) between $\mathbb{R}^2$ and $\mathbb{R}$...which means that two real numbers can be described using just one real number! Nonetheless, we will find a way.

B1 Subspaces

What things are we going to define the dimension *of*? I said just now that it would be 'lines and planes and their higher-dimensional cousins'. But what does that mean, exactly? What things in 100-dimensional space are like lines and planes in our familiar 3-dimensional space? This section is devoted to answering these questions.

To make life easier, we're only going to consider lines, planes, etc. *through the origin.* Consider a plane P through the origin in $\mathbb{R}^3$ (Figure B.2). As the figure shows, if we take any two points $\mathbf{x}$ and $\mathbf{y}$ on the plane P, then their sum $\mathbf{x} + \mathbf{y}$ is also on P. (Compare Figure A.1(i).) There is some jargon for this: one says that P is **closed under addition**. Also, P is **closed under scalar multiplication**, meaning that if we take any point $\mathbf{x}$ on P and any scalar c, then $c\mathbf{x}$ is on P too.

So, any plane P through the origin has three properties:

i. $\mathbf{0} \in P$ (obviously!);

ii. P is closed under addition; that is, for all $\mathbf{x}, \mathbf{y} \in P$, we have $\mathbf{x} + \mathbf{y} \in P$;

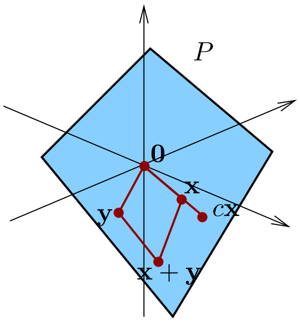


Figure B.2: Algebra in a plane P

iii. P is closed under scalar multiplication; that is, for all $\mathbf{x} \in P$ and $c \in \mathbb{R}$, we have $c\mathbf{x} \in P$.

Now consider a line L through the origin in $\mathbb{R}^3$ or $\mathbb{R}^2$. You should be able to convince yourself that L, too, has these three properties: it contains $\mathbf{0}$ and is closed under addition and scalar multiplication.

You might be able to think of other properties of lines and planes through the origin in $\mathbb{R}^2$ and $\mathbb{R}^3$, but it turns out that these three properties alone are very powerful. We now take these properties and use them as a *definition*, valid in all dimensions.

Definition B1.1 A **linear subspace** of $\mathbb{R}^n$ is a subset V of $\mathbb{R}^n$ with the following properties:

i. $\mathbf{0} \in V$;

ii. for all $\mathbf{x}, \mathbf{y} \in V$, we have $\mathbf{x} + \mathbf{y} \in V$;

iii. for all $\mathbf{x} \in V$ and $c \in \mathbb{R}$, we have $c\mathbf{x} \in V$.

In short: a linear subspace of $\mathbb{R}^n$ is a subset of $\mathbb{R}^n$ containing $\mathbf{0}$ and closed under addition and scalar multiplication.

Mathematicians use the word 'space' in many ways. For example, in the next few years you may meet vector spaces, measure spaces, metric spaces, and topological spaces. Similarly, the word 'subspace' is used in many ways. But in the context of this linear algebra course, we will only be concerned with *linear* subspaces, so it is safe if we call them **subspaces** for short.

Examples B1.2 i. Any plane or line through the origin in $\mathbb{R}^3$ is a linear subspace of $\mathbb{R}^3$, and any line through the origin in $\mathbb{R}^2$ is a linear subspace of $\mathbb{R}^2$. These were our motivating examples.

ii. $\{\mathbf{0}\}$ is a subspace of $\mathbb{R}^n$ (for any $n \geq 0$), called the **trivial subspace**.

iii. $\mathbb{R}^n$ is a subspace of $\mathbb{R}^n$ (for any $n \geq 0$). The subspaces $\{\mathbf{0}\}$ and $\mathbb{R}^n$ are not tremendously interesting, but they satisfy the definition and shouldn't be forgotten!

iv. We now know of three kinds of subspace of $\mathbb{R}^2$: the trivial subspace $\{\mathbf{0}\}$, lines through the origin, and $\mathbb{R}^2$ itself. In fact, these are the *only* subspaces of $\mathbb{R}^2$. You can try proving this for yourself now, but it will be easier once we have developed some technology in the next few sections.

v. Similarly, there are exactly four kinds of subspace of $\mathbb{R}^3$: $\{\mathbf{0}\}$, lines through the origin, planes through the origin, and $\mathbb{R}^3$ itself.

Here is a rather general way of creating linear subspaces of $\mathbb{R}^n$:

Definition B1.3 Let A be an $m \times n$ real matrix. The **kernel** of A is

$$\ker(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

Some people (including Poole) call the kernel the **null space**.

Lemma B1.4 *For any $m \times n$ matrix A, the kernel $\ker(A)$ is a linear subspace of $\mathbb{R}^n$.*

Proof We have to check that the three conditions of Definition B1.1 are satisfied. To do this, we will use some of the algebraic laws in Lemma A4.1.

For condition (i), we have $A\mathbf{0} = \mathbf{0}$, so $\mathbf{0} \in \ker(A)$.

For (ii), if $\mathbf{x}, \mathbf{y} \in \ker(A)$ then

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \mathbf{0} + \mathbf{0} = \mathbf{0},$$

so $\mathbf{x} + \mathbf{y} \in \ker(A)$.

For (iii), if $\mathbf{x} \in \ker(A)$ and $c \in \mathbb{R}$ then

$$A(c\mathbf{x}) = c(A\mathbf{x}) = c\mathbf{0} = \mathbf{0},$$

so $c\mathbf{x} \in \ker(A)$. □

Example B1.5 Let

$$A = \begin{pmatrix} 1 & -2 & 3 \\ -4 & 5 & -6 \end{pmatrix}.$$

This is a 2×3 matrix, so its kernel is a subspace of $\mathbb{R}^3$. It is given by

$$\begin{aligned} \ker(A) &= \{\mathbf{x} \in \mathbb{R}^3 : A\mathbf{x} = \mathbf{0}\} \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} x_1 - 2x_2 + 3x_3 \\ -4x_1 + 5x_2 - 6x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : x_1 - 2x_2 + 3x_3 = -4x_1 + 5x_2 - 6x_3 = 0 \right\}. \end{aligned}$$

Since $\ker(A)$ is a linear subspace of $\mathbb{R}^3$, it must be the trivial subspace $\{\mathbf{0}\}$, a line through $\mathbf{0}$, a plane through $\mathbf{0}$, or the whole of $\mathbb{R}^3$ (by Example B1.2(v)). In fact, it is a line through $\mathbf{0}$.

Given a subspace V of $\mathbb{R}^n$, you can repeatedly use the three conditions in Definition B1.1 to show that if $\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ then $2\mathbf{w} - 7\mathbf{x} - 4\mathbf{y} + 15\mathbf{z} \in V$, etc. We finish this section with a definition and a lemma that capture the general principle.

Definition B1.6 Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ and $\mathbf{y}$ be vectors in $\mathbb{R}^n$. Then $\mathbf{y}$ is a **linear combination** of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ if there exist $c_1, \ldots, c_m \in \mathbb{R}$ such that

$$\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m.$$

Examples B1.7 i. In $\mathbb{R}^3$, the vector $\mathbf{y} = \begin{pmatrix} 5 \\ 9 \\ 2 \end{pmatrix}$ is a linear combination of the vectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} -2 \\ 4 \\ 3 \end{pmatrix}$ and $\mathbf{v}_3 = \begin{pmatrix} 1 \\ -3 \\ -2 \end{pmatrix}$, since

$$\begin{pmatrix} 5 \\ 9 \\ 2 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} + 0 \begin{pmatrix} -2 \\ 4 \\ 3 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ -3 \\ -2 \end{pmatrix}.$$

ii. For any $m \geq 0$ and any $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$, the zero vector $\mathbf{0}$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$, since

$$\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_m.$$

This is true even if $m = 0$! In that case, the right-hand side of the equation is the sum of no things. The sum of no things should always be interpreted as zero. (In this case, ‘thing’ means vector, so ‘zero’ means the zero vector.) Why? Well, at a supermarket, the price you pay is always the sum of the prices of the items in your basket. If you put *no* items in your basket, the price you pay is zero. You’d be upset otherwise!

Now we establish the general principle described just before Definition B1.6.

Lemma B1.8 *Let V be a linear subspace of $\mathbb{R}^n$, let $m \geq 0$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$. Then every linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ also belongs to V.*

Proof If $m = 0$ then (by Example B1.7(ii)) the only linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is $\mathbf{0}$, which belongs to V by definition of subspace.

Now suppose that $m \geq 1$, and assume inductively that the lemma holds for $m - 1$. Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ and let $c_1, \ldots, c_m \in \mathbb{R}$. Put

$$\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_{m-1}\mathbf{v}_{m-1}.$$

Then $\mathbf{w} \in V$ by inductive hypothesis, and $c_m\mathbf{v}_m \in V$ since $\mathbf{v}_m \in V$ and V is closed under scalar multiplication. Also V is closed under addition, so $\mathbf{w} + c_m\mathbf{v}_m \in V$. But $\sum_{i=1}^m c_i\mathbf{v}_i = \mathbf{w} + c_m\mathbf{v}_m$, so $\sum_{i=1}^m c_i\mathbf{v}_i \in V$, completing the induction. □

Example B1.9 I claim that the vector $\mathbf{x} = \begin{pmatrix} 5 \\ 9 \\ 3 \end{pmatrix}$ is *not* a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ of Example B1.7(i). To prove this, we will use the lemma just proved. Put $A = \begin{pmatrix} 1 & -1 & 2 \end{pmatrix}$, a 1×3 matrix. Observe that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \ker(A)$ but $\mathbf{x} \notin \ker(A)$. Since $\ker(A)$ is a linear subspace of $\mathbb{R}^3$, it follows from Lemma B1.8 that $\mathbf{x}$ cannot be expressed as a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$.

You might be wondering what made me think of that particular matrix A. More generally, you might be wondering how, if someone gave you a specific vector $\mathbf{y}$ and specific vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$, you could find out whether or not $\mathbf{y}$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$. We will meet a general method for doing this in the next chapter; it relies on the theory that we are about to develop.

B2 Spanning sets

Let $\mathbf{v}$ and $\mathbf{w}$ be two points in $\mathbb{R}^3$, and suppose that neither is a scalar multiple of the other. Then there is exactly one plane P passing through $\mathbf{v}$, $\mathbf{w}$ and the origin. Since P is a subspace, every linear combination of $\mathbf{v}$ and $\mathbf{w}$ also belongs to P (Lemma B1.8). Conversely, as Figure B.2 suggests, every point of P can be expressed as a linear combination of $\mathbf{v}$ and $\mathbf{w}$. So, P is the set of linear combinations of $\mathbf{v}$ and $\mathbf{w}$.

More generally, for any list $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of vectors in $\mathbb{R}^n$, it turns out to be useful to consider the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Here is the definition.

Definition B2.1 Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$. The **span** of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the set

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \text{ is a linear combination of } \mathbf{v}_1, \ldots, \mathbf{v}_m\}.$$

When $V = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$, we sometimes say that $\mathbf{v}_1, \ldots, \mathbf{v}_m$ **span** V, or that $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a **spanning set** for V.

Examples B2.2 i. In the introductory paragraph to this section, $P = \operatorname{span}\{\mathbf{v}, \mathbf{w}\}$.

ii. Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ be the vectors defined in equations (A:13) (page 35). Then $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ span $\mathbb{R}^n$, since for every $\mathbf{y} \in \mathbb{R}^n$ we have

$$\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \cdots + y_n\mathbf{e}_n \in \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}.$$

iii. In $\mathbb{R}^3$, we have

$$\operatorname{span}\left\{\begin{pmatrix}1\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\-1\end{pmatrix}\right\} = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}.$$

To prove this, we have to show that each side is a subset of the other (following the strategy explained on page 14). Write $\mathbf{v}_1 = \left(\begin{smallmatrix}1\\-1\\0\end{smallmatrix}\right)$ and $\mathbf{v}_2 = \left(\begin{smallmatrix}0\\1\\-1\end{smallmatrix}\right)$, and write $V = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$.

To show that LHS $\subseteq$ RHS, let $\mathbf{v} \in \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Then $\mathbf{v}$ is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. Certainly $\mathbf{v}_1, \mathbf{v}_2 \in V$ (since $1 + (-1) + 0 = 0 = 0 + 1 + (-1)$). But V is a subspace of $\mathbb{R}^3$, since it is the kernel of the 1×3 matrix $(1 \quad 1 \quad 1)$. Hence $\mathbf{v} \in V$ by Lemma B1.8.

(Alternatively, you could write $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$ and show by explicit calculation that the coordinates of $\mathbf{v}$ sum to 0. But I encourage you to minimize calculation whenever possible.)

To show that RHS $\subseteq$ LHS, let $\mathbf{x} \in V$. Then

$$\mathbf{x} = x_1\mathbf{v}_1 - x_3\mathbf{v}_2$$

(check!). This is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$, so $\mathbf{x} \in \operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

iv. A similar argument shows that

$$\operatorname{span}\left\{\begin{pmatrix}1\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\-1\end{pmatrix}, \begin{pmatrix}2\\5\\-7\end{pmatrix}, \begin{pmatrix}-4\\1\\3\end{pmatrix}\right\} = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}.$$

v. The span of a single vector $\mathbf{v}$ is just the set of all scalar multiples of $\mathbf{v}$. This is a line if $\mathbf{v} \neq \mathbf{0}$, and is the trivial subspace $\{\mathbf{0}\}$ if $\mathbf{v} = \mathbf{0}$.

vi. What does Definition B2.1 say if $m = 0$, so that there are no $\mathbf{v}_i$s at all? As we saw in Example B1.7(ii), the vector $\mathbf{0}$ is the one and only linear combination of no vectors. So, the span of the empty list of vectors is $\{\mathbf{0}\}$.

In all of these examples, $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is not merely a *subset* of $\mathbb{R}^n$, but a *linear subspace.* This is no coincidence:

Lemma B2.3 *Let* $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$*. Then* $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ *is a linear subspace of* $\mathbb{R}^n$*.*

Proof We verify the three conditions of Definition B1.1. Write $V = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$.

For (i), we have $\mathbf{0} = \sum_{i=1}^m 0\mathbf{v}_i$, so $\mathbf{0} \in V$.

For (ii), let $\mathbf{x}, \mathbf{y} \in V$. Then $\mathbf{x} = \sum_{i=1}^m c_i\mathbf{v}_i$ and $\mathbf{y} = \sum_{i=1}^m d_i\mathbf{v}_i$ for some scalars $c_1, \ldots, c_m, d_1, \ldots, d_m$. Hence

$$\mathbf{x} + \mathbf{y} = \sum_{i=1}^m c_i\mathbf{v}_i + \sum_{i=1}^m d_i\mathbf{v}_i = \sum_{i=1}^m (c_i\mathbf{v}_i + d_i\mathbf{v}_i) = \sum_{i=1}^m (c_i + d_i)\mathbf{v}_i,$$

so $\mathbf{x} + \mathbf{y} \in V$.

For (iii), let $\mathbf{x} \in V$ and $c \in \mathbb{R}$. We have $\mathbf{x} = \sum_{i=1}^m d_i\mathbf{v}_i$ for some scalars $d_1, \ldots, d_m$. Hence

$$c\mathbf{x} = c\sum_{i=1}^m d_i\mathbf{v}_i = \sum_{i=1}^m cd_i\mathbf{v}_i,$$

so $c\mathbf{x} \in V$. □

So, we can manufacture subspaces of $\mathbb{R}^n$ simply by choosing a few vectors and taking their span.

The span of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is the smallest linear subspace containing $\mathbf{v}_1, \ldots, \mathbf{v}_m$, in the following sense:

Lemma B2.4 *Let* $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$*, and let* V *be a subspace of* $\mathbb{R}^n$*. Then* $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V \iff \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq V$*.*

Proof Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$. By Lemma B1.8, any linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ also belongs to V. So V contains the set of all such linear combinations, which is exactly $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$. Conversely, if $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq V$ then certainly $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$, since $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$. □

Look back at Examples B2.2(iii) and (iv). In going from (iii) to (iv), we have put in two extra vectors on the left-hand side. You might expect this to make the span bigger; but it does not, since they were already in the span of the original two vectors. Here is a general result explaining this:

Lemma B2.5 *Let* $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m+k} \in \mathbb{R}^n$*. Then:*

i. $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subseteq \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m+k}\}$.

ii. If $\mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m+k} \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ *then*

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m+k}\}.$$

Proof For (i), the right-hand side is a subspace of $\mathbb{R}^n$ containing $\mathbf{v}_1, \ldots, \mathbf{v}_m$, so by Lemma B2.4, it also contains $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$.

For (ii), it remains to prove that RHS $\subseteq$ LHS. But $\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a subspace of $\mathbb{R}^n$ containing each of $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_{m+k}$, so by Lemma B2.4, it contains the right-hand side. □

Matrices give rise to linear subspaces in several ways. We have already seen that the kernel of a matrix is a subspace (Lemma B1.4). Here are two more ways.

Let A be an $m \times n$ matrix. Each column of A is an element of $\mathbb{R}^m$, and the span of the n columns of A is called the **column space** of A, written as $\operatorname{col}(A)$. It is a linear subspace of $\mathbb{R}^m$.

Each row of A is an n-dimensional row vector. Strictly speaking, it is not quite an element of $\mathbb{R}^n$, since we made the convention that the elements of $\mathbb{R}^n$ are n-dimensional *column* vectors (page 21). But the transpose of each row is an n-dimensional column vector. The span of the transposes of the m rows of A is called the **row space** of A, and is written as $\operatorname{row}(A)$. It is a linear subspace of $\mathbb{R}^n$.

Example B2.6 Let

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

Then

$$\operatorname{col}(A) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\} = \mathbb{R}^2$$

and

$$\operatorname{row}(A) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\} = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\},$$

where the last equality comes from Example B2.2(iii).

Here is another way of looking at the column space, which will be useful later:

Lemma B2.7 *Let A be an $m \times n$ matrix. Then*

$$\operatorname{col}(A) = \{\mathbf{y} \in \mathbb{R}^m : \mathbf{y} = A\mathbf{x} \textit{ for some } \mathbf{x} \in \mathbb{R}^n\}.$$

Proof Let $\mathbf{y} \in \mathbb{R}^m$. We have to prove that $\mathbf{y} \in \operatorname{col}(A)$ if and only if $\mathbf{y} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$.

Write the columns of A as $\mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^m$. By definition of column space, $\mathbf{y} \in \operatorname{col}(A)$ if and only if

$$\mathbf{y} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$$

for some $x_1, \ldots, x_n \in \mathbb{R}$. But by Lemma A4.3(i),

$$x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Hence $\mathbf{y} \in \operatorname{col}(A)$ if and only if $\mathbf{y} = A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$, as required. □

Remember that our mission in this chapter is to find a good definition of the dimension of a linear subspace $V \subseteq \mathbb{R}^n$. We *could* attempt to define the dimension of V as the smallest number of vectors needed in order to span V. This would give the 'right' answer in the examples we have met so far, but as a general definition, there are several potential problems.

For instance, how do we know that there is *any* finite set of vectors that spans V? If no such set exists, our 'definition' wouldn't make sense. And, supposing that you have found some vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ spanning V, and you *believe* that V can't be spanned by fewer than m vectors, how are you ever going to prove it?

We will answer these questions, and find a good definition of dimension, in the next few sections.

B3 Linear independence

There are infinitely many elements of $\mathbb{R}^n$. However, if you start writing down vectors

$$\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots,$$

you'll eventually find that you start repeating yourself—not in the literal sense that you're mentioning vectors that you've mentioned before, but in the sense that you're mentioning vectors that are *a linear combination of* vectors you've mentioned before.

To a linear algebraist, vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are 'truly different' if none of them is a linear combination of the others. (This is a much stronger requirement than merely asking that none of them is *equal* to any of the others.) The formal definition is as follows.

Definition B3.1 Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ are **linearly independent** if for $c_1, \ldots, c_m \in \mathbb{R}$,

$$c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m = \mathbf{0} \implies c_1 = \cdots = c_m = 0.$$

If $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are not linearly independent, they are said to be **linearly dependent**. So, for $\mathbf{v}_1, \ldots, \mathbf{v}_m$ to be linearly dependent means that there exist scalars $c_1, \ldots, c_m$, not all zero, such that $\sum_i c_i\mathbf{v}_i = \mathbf{0}$.

Definition B3.1 doesn't immediately look as if it's saying 'none of them can be written as a linear combination of the others.' But we will prove later that it is equivalent to that condition (Lemma B3.4). First, some examples:

Examples B3.2 i. Consider the case of a single vector ($m = 1$ in Definition B3.1). The definition says that $\mathbf{v}_1 \in \mathbb{R}^n$ is linearly independent if and only if, for scalars c,

$$c_1\mathbf{v}_1 = \mathbf{0} \implies c_1 = 0.$$

This is true as long as $\mathbf{v}_1 \neq \mathbf{0}$. If $\mathbf{v}_1 = \mathbf{0}$ then $\mathbf{v}_1$ is not linearly independent; see (vi) below.

ii. For two vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ to be linearly independent means that for scalars c_1 and c_2,

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = \mathbf{0} \implies c_1 = c_2 = 0.$$

Linear *dependence* means that there exist scalars c_1 and c_2, not both 0, such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = 0$. Equivalently, $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent if there exist scalars a and b, not both 0, such that $a\mathbf{v}_1 = b\mathbf{v}_2$. If $a \neq 0$ then we can divide through to get $\mathbf{v}_1 = (b/a)\mathbf{v}_2$, and similarly if $b \neq 0$. So, $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent if and only if one is a scalar multiple of the other. (Note that $\mathbf{v}_1$ can be a scalar multiple of $\mathbf{v}_2$ without $\mathbf{v}_2$ being a scalar multiple of $\mathbf{v}_1$. This happens if $\mathbf{v}_1 = \mathbf{0} \neq \mathbf{v}_2$.)

Geometrically, then, $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent if and only if $\mathbf{0}$, $\mathbf{v}_1$ and $\mathbf{v}_2$ are collinear. So in the Cauchy–Schwarz inequality $|\mathbf{v}_1 \cdot \mathbf{v}_2| \leq \|\mathbf{v}_1\| \, \|\mathbf{v}_2\|$ (Lemma A3.2), equality holds if and only if $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent.

iii. The vectors

$$\mathbf{v}_1 = \begin{pmatrix} -1 \\ -8 \\ 8 \\ -1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 4 \\ 0 \\ -2 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 1 \\ -4 \\ 2 \end{pmatrix}$$

in $\mathbb{R}^4$ are linearly dependent, since $2\mathbf{v}_1 + 3\mathbf{v}_2 + 4\mathbf{v}_3 = \mathbf{0}$. In the next chapter, we will establish a general method for deciding whether a given list of vectors is linearly dependent or independent.

iv. The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^n$ (defined in equations (A:13), page 35) are linearly independent. For let $c_1, \ldots, c_n \in \mathbb{R}$ with

$$c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n = \mathbf{0}.$$

The left-hand side of this equation is simply

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix},$$

so $c_1 = \cdots = c_n = 0$, as required.

v. The empty list of vectors is linearly independent, for logical reasons that are fundamental but slightly subtle. You could just take this as part of the definition, or use the following argument.

We have to show, for $n = 0$, that $\sum_{i=1}^n c_i\mathbf{v}_i = \mathbf{0} \Rightarrow c_1 = \cdots = c_n = 0$. Any implication $P \Rightarrow Q$ holds when Q is true, so it is enough to show, when $n = 0$, that $c_i = 0$ for all $i \in \{1, \ldots, n\}$. This says that for all $i \in \varnothing$ we have $c_i = 0$. But any statement beginning 'for all $i \in \varnothing$' is true, since it could only be false if it failed for some i, and there are no i at all.

vi. Any list of vectors containing $\mathbf{0}$ is linearly dependent. Indeed, given vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$ with $\mathbf{v}_1 = \mathbf{0}$, we have $1\mathbf{v}_1 + 0\mathbf{v}_2 \cdots + 0\mathbf{v}_m = \mathbf{0}$, and not all the coefficients on the left-hand side are zero.

Warning B3.3 It is easy to get the definition of linear (in)dependence wrong. For instance, linear dependence of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ does *not* say that $\sum c_i\mathbf{v}_i = \mathbf{0}$ for all $c_1, \ldots, c_m$, nor that $\sum c_i\mathbf{v}_i = \mathbf{0}$ for some $c_1, \ldots, c_m$. Linear independence does *not* say that there are no $c_1, \ldots, c_m$ satisfying $\sum c_i\mathbf{v}_i = \mathbf{0}$.

There are some useful ways of restating the definition of linear independence:

Lemma B3.4 *Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$. The following are equivalent:*

i. *$\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent;*

ii. *for all $i \in \{1, \ldots, m\}$, we have $\mathbf{v}_i \notin \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_m\}$ (that is, none of the vectors is a linear combination of the others);*

iii. *for all $\mathbf{x} \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$, there are* unique *$c_1, \ldots, c_m \in \mathbb{R}$ such that $\mathbf{x} = \sum_{i=1}^m c_i \mathbf{v}_i$.*

Proof We prove that (iii)⇒(ii)⇒(i)⇒(iii).

(iii)⇒(ii): assume (iii). Let $i \in \{1, \ldots, m\}$ and suppose for a contradiction that $\mathbf{v}_i \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_m\}$ for some i. Then we can write

$$\mathbf{v}_i = c_1\mathbf{v}_1 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_m\mathbf{v}_m$$

for some scalars $c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_m$. But then

$$\begin{aligned}&0\mathbf{v}_1 + \cdots + 0\mathbf{v}_{i-1} + 1\mathbf{v}_i + 0\mathbf{v}_{i+1} + \cdots + 0\mathbf{v}_m\\ &= c_1\mathbf{v}_1 + \cdots + c_{i-1}\mathbf{v}_{i-1} + 0\mathbf{v}_i + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_m\mathbf{v}_m,\end{aligned}$$

and the coefficients of $\mathbf{v}_i$ are different on the left- and right-hand sides, contradicting the uniqueness in (iii).

(ii)⇒(i): assume (ii). Let $c_1, \ldots, c_m \in \mathbb{R}$ with $\sum_i c_i\mathbf{v}_i = \mathbf{0}$, and assume for a contradiction that $c_i \neq 0$ for some i. Then for that i, we have

$$\mathbf{v}_i = -\frac{1}{c_i}(c_1\mathbf{v}_1 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_m\mathbf{v}_m)$$

and so $\mathbf{v}_i \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_m\}$, a contradiction.

(i)⇒(iii): assume (i). Let $\mathbf{x} \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$, and suppose that $\mathbf{x} = \sum_i c_i\mathbf{v}_i = \sum_i d_i\mathbf{v}_i$. Subtracting, $\sum_i (c_i - d_i)\mathbf{v}_i = 0$. But then by linear independence, $c_i - d_i = 0$ for all i, or equivalently $c_i = d_i$ for all i. □

Lemma B2.5(ii) tells us that if we start with a spanning set for a subspace V of $\mathbb{R}^n$, then put some more elements of V into it, the result stills spans V. On the other hand, if we start with a linearly independent set of vectors, then remove some of them, the result is still linearly independent. In summary: if you make a spanning set bigger, it's still spanning, and if you make a linearly independent set smaller, it's still linearly independent.

But so far we know nothing about how the sizes of spanning sets and of linearly independent sets are related to one another. The next result describes the relationship, and is the foundation stone of the theory of dimension.

Proposition B3.5 (Steinitz exchange lemma) *Let V be a subspace of $\mathbb{R}^n$. Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be linearly independent vectors in V and let $\mathbf{w}_1, \ldots, \mathbf{w}_m$ be vectors spanning V. Then $k \leq m$.*

Before I give the proof, here is the intuitive idea. In a d-dimensional space, any linearly independent set has $\leq d$ elements, since there 'isn't enough room' for more than d linearly independent vectors inside V. (For instance, you can't 'fit' 4 linearly independent vectors into $\mathbb{R}^3$.) On the other hand, in a d-dimensional space, any spanning set must have $\geq d$ elements. So in the situation of Proposition B3.5, if V is d-dimensional then $k \leq d$ and $m \geq d$. It follows that $k \leq m$.

However, right now we can't make this intuitive idea precise, since we don't yet have a definition of dimension! So the following proof of Proposition B3.5 takes another path.

Proof In fact, we prove something stronger: that it is possible to choose k members of the list $\mathbf{w}_1, \ldots, \mathbf{w}_m$ in such a way that when these members are replaced by $\mathbf{v}_1, \ldots, \mathbf{v}_k$, the resulting list still spans V. (For instance, if $k = 3$ and $m = 5$ then it may be that $\mathbf{w}_1, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_4, \mathbf{v}_3$ span V.) It will follow immediately that $k \leq m$, since otherwise it would not be possible to choose k members of the list $\mathbf{w}_1, \ldots, \mathbf{w}_m$ at all.

We choose the members to be replaced one by one. Let $0 \leq i < k$. Suppose inductively that $m \geq i$ and that we have chosen i members of the list $\mathbf{w}_1, \ldots, \mathbf{w}_m$ in such a way that when these members are replaced by $\mathbf{v}_1, \ldots, \mathbf{v}_i$, the resulting list spans V. (Clearly this is possible when $i = 0$.) We may assume without loss of generality that the i members of the list replaced so far are the first i; thus, we are assuming that $\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{w}_{i+1}, \ldots, \mathbf{w}_m$ span V.

Since $\mathbf{v}_{i+1} \in V$ and $V = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{w}_{i+1}, \ldots, \mathbf{w}_m\}$, we can write

$$\mathbf{v}_{i+1} = c_1\mathbf{v}_1 + \cdots + c_i\mathbf{v}_i + c_{i+1}\mathbf{w}_{i+1} + \cdots + c_m\mathbf{w}_m \tag{B:1}$$

for some scalars $c_1, \ldots, c_m$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent, not all of $c_{i+1}, \ldots, c_m$ can be zero (by Lemma B3.4(ii)). In particular, this implies that $m \geq i + 1$. Assume without loss of generality that $c_{i+1} \neq 0$. Then we can rearrange equation (B:1) to show that

$$\mathbf{w}_{i+1} \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{v}_{i+1}, \mathbf{w}_{i+2}, \ldots, \mathbf{w}_m\}. \tag{B:2}$$

Write $W = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{v}_{i+1}, \mathbf{w}_{i+2}, \ldots, \mathbf{w}_m\}$. Then W is a linear subspace of V containing each of $\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{w}_{i+1}, \mathbf{w}_{i+2}, \ldots, \mathbf{w}_m$, by (B:2). (Notice the difference between the lists of vectors in the last two sentences! That's a crucial point in the proof.) By Lemma B2.4, W therefore contains the span of this list, which is V. So W is a subspace of V containing V; that is, $W = V$. Hence $\mathbf{v}_1, \ldots, \mathbf{v}_i, \mathbf{v}_{i+1}, \mathbf{w}_{i+2}, \ldots, \mathbf{w}_m$ span V, completing the induction. □

Corollary B3.6 *If $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ are linearly independent then $k \leq n$.*

Proof The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ span $\mathbb{R}^n$ (Example B2.2(ii)), so the result follows from Proposition B3.5 with $V = \mathbb{R}^n$. □

Put another way, if $k > n$ then any k vectors in $\mathbb{R}^n$ are linearly dependent. For instance, any 4 vectors in $\mathbb{R}^3$ are linearly dependent: one must be in the span of the others.

We *could* define the dimension of a subspace V of $\mathbb{R}^n$ as the largest number of linearly independent vectors in V that it is possible to find. But in particular examples, how could we ever verify that a larger linearly independent set could not be found? And does this agree with the 'definition' of dimension proposed at the end of Section B2? We are getting closer to being able to answer these questions. . .

B4 Bases

We have seen that the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ span $\mathbb{R}^n$ and are linearly independent. Another way of saying this is that every vector $\mathbf{x} \in \mathbb{R}^n$ can be written as a linear combination

$$\mathbf{x} = c_1\mathbf{e}_1 + \cdots + c_n\mathbf{e}_n$$

in *exactly one* way. Indeed, for $\mathbf{e}_1, \ldots, \mathbf{e}_n$ to span $\mathbb{R}^n$ means that every $\mathbf{x} \in \mathbb{R}^n$ can be written as a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in *at least one* way, and for them to be linearly independent means that every $\mathbf{x} \in \mathbb{R}^n$ can be written as a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ in *at most one* way (Lemma B3.4(iii)).

In the terminology we are about to introduce, $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a 'basis' of $\mathbb{R}^n$. But remember that we are trying to find the right definition of dimension not just for $\mathbb{R}^n$ itself, but for arbitrary linear subspaces. So we make the following definition.

Definition B4.1 Let V be a subspace of $\mathbb{R}^n$. A **basis** of V is a list $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of elements of V that is linearly independent and spans V.

In other words, a basis of V is a linearly independent spanning set. The plural of basis is **bases** (pronounced 'base-eez').

Examples B4.2 i. Take V to be $\mathbb{R}^n$ itself. The list of vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis of $\mathbb{R}^n$, called the **standard basis** of $\mathbb{R}^n$.

ii. For any nonzero $x \in \mathbb{R}$, the one-element list x is a basis of $\mathbb{R}$. It is linearly independent by Example B3.2(i), and spans $\mathbb{R}$ because every element of $\mathbb{R}$ is a multiple of x.

iii. Let $\mathbf{v} \in \mathbb{R}^n$ be any nonzero vector, and consider the line $V = \operatorname{span}\{\mathbf{v}\}$. For any scalar $c \neq 0$, the one-element list $c\mathbf{v}$ is a basis of V (for the same reasons as in (ii)).

iv. The empty list of vectors is a basis of the trivial subspace $\{\mathbf{0}\}$ of $\mathbb{R}^n$, since it spans $\{\mathbf{0}\}$ (Example B2.2(vi)) and is linearly independent (Example B3.2(v)). (Poole's book wrongly states that $\{\mathbf{0}\}$ has no basis.)

v. Let $V = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$. As observed in Example B2.2(iii), V is a linear subspace of $\mathbb{R}^3$ spanned by the vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

It is straightforward to show that $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. (As always, when you read a sentence like that, you should treat it as an exercise!) It follows that the two-element list $\mathbf{v}_1, \mathbf{v}_2$ is a basis of V.

vi. With the same V as in the last example, another basis is

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \qquad \mathbf{w}_2 = \begin{pmatrix} -3 \\ -5 \\ 8 \end{pmatrix}.$$

Our next lemma extends the observations about $\mathbf{e}_1, \dots, \mathbf{e}_n$ in the first paragraph of this section.

Lemma B4.3 *Let V be a linear subspace of $\mathbb{R}^n$ and let $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$.*

i. $\mathbf{v}_1, \dots, \mathbf{v}_m$ *span V $\iff$ for all $\mathbf{x} \in V$, there is* at least one *list of scalars $c_1, \dots, c_m$ such that $\mathbf{x} = \sum c_i \mathbf{v}_i$.*

ii. $\mathbf{v}_1, \dots, \mathbf{v}_m$ *are linearly independent $\iff$ for all $\mathbf{x} \in V$, there is* at most one *list of scalars $c_1, \dots, c_m$ such that $\mathbf{x} = \sum c_i \mathbf{v}_i$.*

iii. $\mathbf{v}_1, \dots, \mathbf{v}_m$ *is a basis of V $\iff$ for all $\mathbf{x} \in V$, there is* exactly one *list of scalars $c_1, \dots, c_m$ such that $\mathbf{x} = \sum c_i \mathbf{v}_i$.*

Proof Part (i) is the definition of spanning, part (ii) follows from Lemma B3.4, and part (iii) follows from parts (i) and (ii). □

Warning B4.4 Examples B4.2(v) and (vi) demonstrate that the same subspace can have multiple bases. *If you ever find yourself writing the words 'the basis', you're probably making a mistake.*

Remark B4.5 The list of vectors $\mathbf{e}_1, \dots, \mathbf{e}_n$ is called the 'standard' basis of $\mathbb{R}^n$ because it is in some sense the obvious one. It is not the *only* basis of $\mathbb{R}^n$; even $\mathbb{R}^1$ has multiple bases, as Example B4.2(ii) shows. But $\mathbb{R}^n$ has the property that an obvious choice of basis exists.

However, it is important to appreciate that for most subspaces V of $\mathbb{R}^n$, there is no obvious choice of basis. For instance, consider the plane $V = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$. It is unlikely that anyone would claim that the basis in Example B4.2(vi) is 'obvious', but perhaps the basis in Example B4.2(v) looks a bit more natural. However, can you really claim that it is superior to the basis $\mathbf{v}_1, \mathbf{w}_1$, for instance, or the basis $-\mathbf{w}_1, -\mathbf{v}_2$?

Although the same subspace can have many different bases, you may have noticed that in all the examples so far, all the bases of a given subspace have the same number of elements. This is a general truth:

Proposition B4.6 *Let V be a linear subspace of $\mathbb{R}^n$. Then any two bases of V have the same number of elements.*

Proof Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ and $\mathbf{w}_1, \dots, \mathbf{w}_m$ be bases of V. Since any basis of V is linearly independent *and* spans V, the Steinitz exchange lemma (Proposition B3.5) implies that both $k \leq m$ and $m \leq k$. Hence $k = m$. □

We could now attempt to define the dimension of V as the number of elements in a basis of V... but we still have a problem! That 'definition' only makes sense if every subspace has at least one basis. And so far, we don't know that. We will, however, prove it soon.

First we show that given any list of vectors spanning V, we can get a basis by deleting some of them.

Lemma B4.7 (Deletion) *Let V be a subspace of $\mathbb{R}^n$ and let $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$ be vectors spanning V. Then some subset of $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ is a basis of V.*

This subset could consist of all, some or none of $\mathbf{v}_1, \ldots, \mathbf{v}_m$. 'All' would be the case if $\mathbf{v}_1, \ldots, \mathbf{v}_m$ was already a basis. 'None' would be the case if $V = \{\mathbf{0}\}$ (Example B4.2(iv)).

Proof Consider all subsets of $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ that span V. Choose one with the smallest possible number of elements: k, say. Without loss of generality, we may assume that it is of the form $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$.

I claim that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a basis of V. Certainly it spans V, so it only remains to show that it is linearly independent. Suppose not. Then by Lemma B3.4, there exists $i \in \{1, \ldots, k\}$ such that $\mathbf{v}_i \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k\}$. But then $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k$ span V, by Lemma B2.5(ii). This is a subset of $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ that spans V and has fewer than k elements, a contradiction. $\square$

Examples B4.8 Let $V = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$, a subspace of $\mathbb{R}^3$.

i. The vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 4 \\ -5 \end{pmatrix}$$

span V. Hence by Lemma B4.7, some subset of this list must be a basis of V. In fact, any two of them form a basis of V. (Check!)

ii. The vectors

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{w}_3 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$$

also span V. This time, $\mathbf{w}_1, \mathbf{w}_2$ is a basis and $\mathbf{w}_1, \mathbf{w}_3$ is a basis, but $\mathbf{w}_2, \mathbf{w}_3$ is not a basis. (Again, check!)

The mirror image of Lemma B4.7 is the next result, which states that any linearly independent set can be extended to make a basis.

Lemma B4.9 (Extension) *Let V be a subspace of $\mathbb{R}^n$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be linearly independent vectors in V. Then there is some basis of V containing all of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.*

Proof Consider all lists of linearly independent vectors in V containing $\mathbf{v}_1, \ldots, \mathbf{v}_k$. By Corollary B3.6, no such list contains more than n elements. We can therefore choose one with the largest number of elements (m, say) and call it $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_m$.

I claim that $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a basis of V. Certainly it is linearly independent, so it only remains to show that it spans V. Let $\mathbf{v} \in V$. The list of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}$ has more than m elements, so these vectors are linearly dependent. Hence there exist scalars $c_1, \ldots, c_m, c$, not all zero, such that

$$c_1\mathbf{v}_1 + \cdots + c_m\mathbf{v}_m + c\mathbf{v} = 0.$$

Since $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent, we must have $c \neq 0$. Rearranging the equation then shows that $\mathbf{v}$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$, as required.$\square$

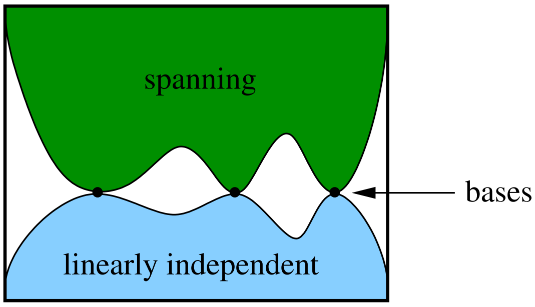


Figure B.3: Schematic diagram of the main results of this section

Example B4.10 Let $V = \mathbb{R}^3$, and put

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}.$$

Then $\mathbf{v}_1, \mathbf{v}_2$ are linearly independent. To extend to a basis, we simply pick any vector $\mathbf{v}_3$ not in $\operatorname{span}\{\mathbf{v}_1, \mathbf{v}_2\}$ (such as $\mathbf{v}_3 = \left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$); then $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is a basis.

We can now easily deduce the result that we hoped for above:

Proposition B4.11 *Every linear subspace of* $\mathbb{R}^n$ *has at least one basis.*

Proof This follows from Lemma B4.9 by taking $\mathbf{v}_1, \ldots, \mathbf{v}_k$ to be the empty list (i.e. the list with $k = 0$), which is linearly independent by Example B3.2(v). □

The main results of this section are summarized in Figure B.3. Don't take it too literally; this is not a picture of $\mathbb{R}^n$. It is to be interpreted as follows.

Let V be a subspace of $\mathbb{R}^n$. Each point in Figure B.3 represents a set of vectors in V, with higher-up points representing larger sets. As remarked just before Proposition B3.5, if you take a spanning set for V and put some more elements of V into it, it still spans V. Correspondingly, in Figure B.3, if you take a point representing a spanning set and move upwards, you are still in the zone of spanning sets. Similarly, moving downwards in Figure B.3 from a point representing a linearly independent set always gives you another point representing a linearly independent set, because if you remove some vectors from a linearly independent set then it's still linearly independent.

Moreover, if we start with a spanning set then it is always possible to move downwards to reach a basis. This is the deletion lemma (Lemma B4.7). And if we start with a linearly independent set, it is always possible to move upwards to reach a basis; that's the extension lemma (Lemma B4.9).

Finally, all the bases are shown at the same height, because all bases of V have the same number of elements (Proposition B4.11).

So, many of our results are encapsulated in Figure B.3. But the diagram is also misleading in some ways. For instance, if V is nontrivial then V has infinitely many bases, even though Figure B.3 shows only three.

B5 The definition of dimension

We are finally ready to define the dimension of a linear subspace of $\mathbb{R}^n$, and to show that it enjoys all the good properties that we might hope for.

Definition B5.1 Let V be a linear subspace of $\mathbb{R}^n$. The **dimension** of V, written as $\dim V$, is the number of elements in any basis of V.

In order for this definition to make sense, we need to know two things: that every subspace V of $\mathbb{R}^n$ has at least one basis, and that any two bases of V have the same number of elements. We proved these in Propositions B4.11 and B4.6, respectively.

Examples B5.2 i. $\dim(\mathbb{R}^n) = n$, since $\mathbb{R}^n$ has a basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ with n elements (Example B4.2(i)).

ii. $\dim(\{\mathbf{0}\}) = 0$, since the empty list is a basis of $\{\mathbf{0}\}$ (Example B4.2(iv)).

iii. The subspace
$$V = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$$
of $\mathbb{R}^3$ has dimension 2, since it has a basis with 2 elements (Example B4.2(v)). Two-dimensional subspaces of $\mathbb{R}^n$ are often called **planes**, and one-dimensional subspaces are **lines**.

iv. The subspace
$$V = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0,\ x_1 = x_2\}$$
of $\mathbb{R}^3$ has dimension 1 (it's a line), since the one-element list $\left(\begin{smallmatrix} 1 \\ 1 \\ -2 \end{smallmatrix}\right)$ is a basis (exercise!).

At the end of Section B2, we proposed a definition of dimension using spanning sets, and at the end of Section B3, we proposed another definition of dimension using linearly independent sets. At the time, it wasn't clear that the definitions made sense or agreed with each other. But we can now show that they do indeed make sense, and that they both agree with the definition of dimension just given.

Proposition B5.3 *Let V be a linear subspace of $\mathbb{R}^n$.*

i. $\dim V$ is the smallest number of elements in any spanning set of V. That is, V has a spanning set with exactly $\dim V$ elements, and every spanning set has $\geq \dim V$ elements.

ii. $\dim V$ is the largest number of elements in a linearly independent subset of V. That is, there is a linearly independent subset of V with exactly $\dim V$ elements, and every linearly independent subset of V has $\leq \dim V$ elements.

Proof Choose a basis $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of V. Then $\dim V = m$.

For (i), $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a spanning set of V with $m = \dim V$ elements. Now let $\mathbf{w}_1, \ldots, \mathbf{w}_k$ be a spanning set of V. By the deletion lemma (Lemma B4.7), some subset of $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ is a basis of V; but any basis has m elements, so $k \geq m = \dim V$.

For (ii), $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a linearly independent subset of V with $m = \dim V$ elements. Now let $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ be a linearly independent subset of V. By the extension lemma (Lemma B4.9), there is some basis of V containing all of $\mathbf{w}_1, \ldots, \mathbf{w}_k$; but any basis of V has m elements, so $k \leq m = \dim V$. □

Let V be an m-dimensional subspace of $\mathbb{R}^n$. What if we have a list of m vectors in V and want to know whether it is a basis of V? In that situation, the following result is very useful.

Proposition B5.4 *Let V be an m-dimensional linear subspace of $\mathbb{R}^n$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be m vectors in V. Then*

$$\begin{aligned} & \mathbf{v}_1, \ldots, \mathbf{v}_m \textit{ is a basis of } V \\ \iff & \mathbf{v}_1, \ldots, \mathbf{v}_m \textit{ span } V \\ \iff & \mathbf{v}_1, \ldots, \mathbf{v}_m \textit{ are linearly independent.} \end{aligned}$$

Proof Any basis of V is certainly linearly independent and spans V, so it remains to prove the converses: if $\mathbf{v}_1, \ldots, \mathbf{v}_m$ span V or are linearly independent then they are a basis of V.

First suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_m$ span V. By the deletion lemma (Lemma B4.7), some subset of $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ is a basis of V. But any basis of V has exactly m elements, so this subset must be the whole of $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$. Hence $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a basis of V.

The proof in the case that $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent is very similar, using the extension lemma instead (Lemma B4.9): exercise. □

Warning B5.5 It's certainly not true that an arbitrary spanning set or linearly independent set is a basis. The crucial point in Proposition B5.4 is that the number of vectors in the list is the same as the dimension of the subspace.

Proposition B5.4 can also be seen in Figure B.3: any spanning set or linearly independent set at the same height as the dots representing the bases is itself a basis.

Example B5.6 Once more, consider the subspace $V = \{\mathbf{x} \in \mathbb{R}^3 : x_1+x_2+x_3 = 0\}$ of $\mathbb{R}^3$. We have already shown that V is 2-dimensional, that is, a plane (Example B5.2(iii)).

Suppose you wanted to show that the vectors

$$\mathbf{x}_1 = \begin{pmatrix} -2 \\ 3 \\ -1 \end{pmatrix}, \qquad \mathbf{x}_2 = \begin{pmatrix} 1 \\ 5 \\ -6 \end{pmatrix}$$

form a basis of V. (It's clear that they do belong to V, since $-2+3+(-1) = 0 = 1+5+(-6)$.) You *could* do this by calculating explicitly that $\mathbf{x}_1$ and $\mathbf{x}_2$ span V, and that they are linearly independent. However, the theory we have developed enables us to do it with almost no calculation at all, as follows.

The vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are not scalar multiples of one another, which by Example B3.2(ii) implies that they are linearly independent. Since $\dim V = 2$, Proposition B5.4 then implies that $\mathbf{x}_1, \mathbf{x}_2$ is a basis of V. And that's it!

Remark B5.7 Proposition B5.4, or something like it, is also useful in the context of differential equations. Suppose we have proved that the space of solutions of a 2nd order linear ordinary differential equation is 2-dimensional, and we have found two linearly independent solutions of the equation. Then it follows that our two solutions *span* the space of solutions: in other words, every solution is a linear combination of our two.

What's needed here is not quite Proposition B5.4, but a very similar result in the theory of vector spaces. These are just beyond the scope of this course.

Proposition B5.4 is also used to prove the following useful result.

Proposition B5.8 *Let* $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$*. Then* $\dim(\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}) \leq m$*, with equality if and only if* $\mathbf{v}_1, \ldots, \mathbf{v}_m$ *are linearly independent.*

Proof Exercise. □

Let V and W be subspaces of $\mathbb{R}^n$ with $V \subseteq W$. Then we would expect $\dim V$ to be no greater than $\dim W$. Moreover, $\dim V$ should be strictly smaller than $\dim W$ unless $V = W$; for instance, you can't have one plane in $\mathbb{R}^3$ being a proper subset of another. Our intuition is, on this occasion, correct:

Lemma B5.9 *Let* V *and* W *be linear subspaces of* $\mathbb{R}^n$ *with* $V \subseteq W$*. Then* $\dim V \leq \dim W$*, with equality if and only if* $V = W$*.*

Proof Choose a basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of V. Then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent vectors in W, so by the extension lemma (Lemma B4.9), we can extend this list to a basis $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_m$ of W. In particular, $k \leq m$, that is, $\dim V \leq \dim W$. If $\dim V = \dim W$ then $k = m$, so $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a basis of both V and W, so

$$V = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = W.$$

Conversely, if $V = W$ then $\dim V = \dim W$ immediately. □

This lemma provides a useful strategy for showing that two subspaces V and W of $\mathbb{R}^n$ are equal: first show that $V \subseteq W$, then show that $\dim V = \dim W$.

We use this strategy in the second part of the following lemma. Recall from Examples B5.2 that $\{\mathbf{0}\}$ is a 0-dimensional subspace of $\mathbb{R}^n$ and that $\mathbb{R}^n$ is an n-dimensional subspace of $\mathbb{R}^n$. In fact, these are the *only* 0- and n-dimensional subspaces of $\mathbb{R}^n$:

Lemma B5.10 *The only 0-dimensional subspace of* $\mathbb{R}^n$ *is* $\{\mathbf{0}\}$*, and the only* n*-dimensional subspace of* $\mathbb{R}^n$ *is* $\mathbb{R}^n$*.*

Proof Let V be a 0-dimensional subspace of $\mathbb{R}^n$. Then V has a basis with no elements; that is, $\varnothing$ is a basis of V. Hence $V = \operatorname{span}\varnothing = \{\mathbf{0}\}$ by Example B2.2(vi).

Now let V be an n-dimensional subspace of $\mathbb{R}^n$. Since $V \subseteq \mathbb{R}^n$ and $\dim V = \dim \mathbb{R}^n$, Lemma B5.9 implies that $V = \mathbb{R}^n$. □

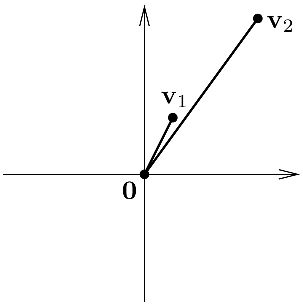


Figure B.4: A basis of $\mathbb{R}^2$

B6 Orthonormal bases

The easiest example of a basis is the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ of $\mathbb{R}^n$. However, as an example it is somewhat misleading, since it has the special properties that the basis vectors all have length 1 and are all orthogonal to each other. In general, basis vectors can be more or less any length and at more or less any angle to each other. For instance (Figure B.4), any two vectors $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^2$ form a basis of $\mathbb{R}^2$, just as long as neither is a scalar multiple of the other (by the argument of Example B5.6).

In this section, we will consider bases that, like the standard basis, enjoy the special properties just mentioned.

Recall from Section A3 that two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are said to be orthogonal if $\mathbf{x} \cdot \mathbf{y} = 0$. Now we extend this definition to lists of any number of vectors.

Definition B6.1 Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$.

i. We say that $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are **orthogonal** if $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ for all $i, j \in \{1, \ldots, m\}$ with $i \neq j$.

ii. We say that they are **orthonormal** if they are orthogonal and $\|\mathbf{v}_i\| = 1$ for all $i \in \{1, \ldots, m\}$.

Examples B6.2 i. The standard basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are orthonormal, since $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ for all $i \neq j$ and $\|\mathbf{e}_i\| = 1$ for all i.

ii. The vectors

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}, \qquad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}$$

are not orthogonal, as even though $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$ and $\mathbf{v}_1 \cdot \mathbf{v}_3 = 0$, we have $\mathbf{v}_2 \cdot \mathbf{v}_3 \neq 0$. (Suggestion: draw a picture showing $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$.)

iii. The vectors $\binom{1}{1}, \binom{3}{-3} \in \mathbb{R}^2$ are orthogonal. But they are not orthonormal, as their lengths are not 1.

iv. The vectors $\begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}, \begin{pmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{pmatrix}$ are orthonormal (Figure B.5). This example was obtained from the previous example by rescaling the two vectors to make them have length 1.

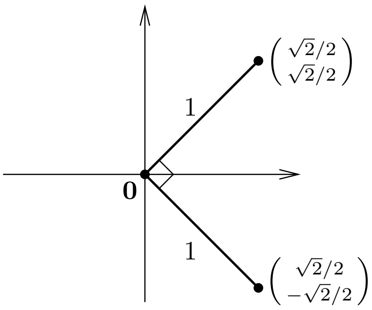

$\begin{pmatrix}\sqrt{2}/2\\ \sqrt{2}/2\end{pmatrix}$
1
$\mathbf{0}$
1
$\begin{pmatrix}\sqrt{2}/2\\ -\sqrt{2}/2\end{pmatrix}$

Figure B.5: Two orthonormal vectors in $\mathbb{R}^2$

The definition of orthonormality can be rephrased neatly as follows.

Lemma B6.3 *Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$. Then $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are orthonormal if and only if for all $i, j \in \{1, \ldots, m\}$,*

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 0 & \text{if } i \neq j, \\ 1 & \text{if } i = j. \end{cases}$$

Proof This follows from the fact that $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}}$ for all $\mathbf{v} \in \mathbb{R}^n$. □

We now show that orthonormal vectors are always linearly independent. Orthonormality is a much stronger condition than linear independence: linearly independent vectors need not even be orthogonal, let alone orthonormal (Figure B.4).

Lemma B6.4 *Let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$. If $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are orthogonal and all nonzero then they are linearly independent. In particular, orthonormal vectors are linearly independent.*

Proof Let $c_1, \ldots, c_m$ be scalars such that

$$\sum_{i=1}^{m} c_i \mathbf{v}_i = \mathbf{0}. \tag{B:3}$$

Taking the dot product of each side of equation (B:3) with $\mathbf{v}_1$ gives

$$\left(\sum_{i=1}^{m} c_i \mathbf{v}_i \right) \cdot \mathbf{v}_1 = 0.$$

By Lemma A3.1 and induction, the left-hand side is $\sum_i c_i(\mathbf{v}_i \cdot \mathbf{v}_1)$. Then by orthogonality, the left-hand side is $c_1\|\mathbf{v}_1\|^2$. Hence $c_1\|\mathbf{v}_1\|^2 = 0$; but $\mathbf{v}_1 \neq \mathbf{0}$, so $c_1 = 0$. Similarly, $c_2 = \cdots = c_m = 0$. So $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent. □

Let V be a subspace of $\mathbb{R}^n$. An **orthonormal basis** of V is an orthonormal list of vectors that is a basis of V. For example, the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is an orthonormal basis of $\mathbb{R}^n$.

Corollary B6.5 *Let V be an m-dimensional subspace of $\mathbb{R}^n$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be m orthonormal vectors in V. Then $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is an orthonormal basis of V.*

Proof By Lemma B6.4, $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent. But $\dim V = m$, so by Proposition B5.4, they are a basis of V. □

Geometrically, then, an orthonormal basis of an m-dimensional subspace V consists of m unit-length vectors in V all at right angles to each other.

When $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is a basis of V (not necessarily orthonormal), we can write any vector $\mathbf{x} \in V$ as a linear combination

$$\mathbf{x} = \sum_{i=1}^{m} c_i \mathbf{v}_i$$

for a unique list $c_1, \ldots, c_m$ of scalars. (We saw this in Lemma B4.3(iii).) But how do we actually *find* these scalars? For instance, the vectors

$$\begin{pmatrix} -2 \\ 1 \\ 5 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 7 \\ -4 \end{pmatrix}$$

form a basis of $\mathbb{R}^3$, so there are unique scalars c_1, c_2, c_3 such that

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = c_1 \begin{pmatrix} -2 \\ 1 \\ 5 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} + c_3 \begin{pmatrix} 2 \\ 7 \\ -4 \end{pmatrix}.$$

How would we go about finding the coefficients c_1, c_2 and c_3? We'd have to solve a set of simultaneous equations. In the next chapter, we'll see how to do that sort of computation efficiently. But right now, the point is that *when the basis is orthonormal, this task is much easier.* In fact, the coefficients are given by a very simple formula:

Lemma B6.6 *Let V be a linear subspace of $\mathbb{R}^n$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be an orthonormal basis of V. Then for all $\mathbf{x} \in V$,*

$$\mathbf{x} = \sum_{i=1}^{m} (\mathbf{x} \cdot \mathbf{v}_i) \mathbf{v}_i.$$

Proof By definition of basis, $\mathbf{x} = \sum_i c_i \mathbf{v}_i$ for some scalars $c_1, \ldots, c_m$. Let $j \in \{1, \ldots, m\}$. Taking the dot product of each side of this equation with $\mathbf{v}_j$ gives

$$\mathbf{x} \cdot \mathbf{v}_j = \left(\sum_{i=1}^{m} c_i \mathbf{v}_i \right) \cdot \mathbf{v}_j = \sum_{i=1}^{m} c_i (\mathbf{v}_i \cdot \mathbf{v}_j) = c_j,$$

by Lemma B6.3. So $c_j = \mathbf{x} \cdot \mathbf{v}_j$, and the result follows. □

Examples B6.7 i. Take $V = \mathbb{R}^n$ with its standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$, which we have already noted is orthonormal. For $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$, we have $\mathbf{x} \cdot \mathbf{e}_i = x_i$. So in this case, Lemma B6.6 simply states that $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$.

ii. We observed in Example B6.2(iv) that the vectors

$$\mathbf{v}_1 = \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{pmatrix}$$

are orthonormal. It follows from Corollary B6.5 that they form a basis of $\mathbb{R}^2$. How can we find the unique scalars c_1, c_2 such that

$$\begin{pmatrix} 5 \\ 3 \end{pmatrix} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2?$$

Lemma B6.6 gives the answer:

$$c_1 = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \cdot \mathbf{v}_1 = 4\sqrt{2}, \qquad c_2 = \begin{pmatrix} 5 \\ 3 \end{pmatrix} \cdot \mathbf{v}_2 = \sqrt{2},$$

and so

$$\begin{pmatrix} 5 \\ 3 \end{pmatrix} = 4\sqrt{2}\,\mathbf{v}_1 + \sqrt{2}\,\mathbf{v}_2.$$

You can check directly that this is correct.

Back in Proposition B4.11, we showed that every subspace of $\mathbb{R}^n$ has at least one basis. But so far, we don't know whether every subspace of $\mathbb{R}^n$ has an *orthonormal* basis. For all we know so far, it could be that some subspaces have no orthonormal basis at all. However, we will soon show that every subspace of $\mathbb{R}^n$ *does* have an orthonormal basis. To do this, we introduce the concept of 'orthogonal complement'.

B7 Orthogonal complements

Given a line through the origin in $\mathbb{R}^3$, you can take the plane through the origin orthogonal to it. Similarly, given a plane through the origin in $\mathbb{R}^3$, you can take the line through the origin orthogonal to it. Here is the general definition.

Definition B7.1 Let V be a linear subspace of $\mathbb{R}^n$. The **orthogonal complement** of V is

$$V^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{v} = 0 \text{ for all } \mathbf{v} \in V\}.$$

$V^\perp$ is pronounced 'V-perp'. In order for a vector $\mathbf{x}$ to belong to $V^\perp$, it must be orthogonal to *everything* in V.

Lemma B7.2 *Let V be a linear subspace of $\mathbb{R}^n$. Then $V^\perp$ is also a linear subspace of $\mathbb{R}^n$.*

Proof We verify the three conditions of Definition B1.1.

For (i), certainly $\mathbf{0} \cdot \mathbf{v} = 0$ for all $\mathbf{v} \in V$.

For (ii), let $\mathbf{x}, \mathbf{y} \in V^\perp$. Then for each $\mathbf{v} \in V$, we have

$$(\mathbf{x} + \mathbf{y}) \cdot \mathbf{v} = \mathbf{x} \cdot \mathbf{v} + \mathbf{y} \cdot \mathbf{v} = 0 + 0 = 0,$$

so $\mathbf{x} + \mathbf{y} \in V^\perp$.

The proof of (iii) is similar and left as an exercise. □

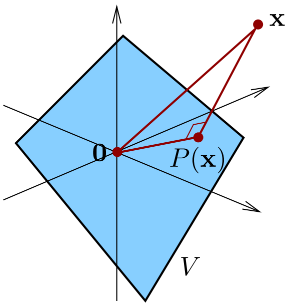


Figure B.6: Projecting a vector $\mathbf{x}$ onto a subspace V

To check that an element $\mathbf{x} \in \mathbb{R}^n$ belongs to $V^\perp$, in principle we need to check that $\mathbf{x} \cdot \mathbf{v} = 0$ for *every* $\mathbf{v} \in V$. That might seem very hard: typically V has infinitely many elements, so in principle there are infinitely many checks to do. The following lemma makes life much easier.

Lemma B7.3 *Let V be a linear subspace of $\mathbb{R}^n$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_m \in V$ be vectors spanning V. Then*

$$V^\perp = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \cdot \mathbf{v}_1 = \cdots = \mathbf{x} \cdot \mathbf{v}_m = 0\}.$$

Proof Exercise. □

Example B7.4 Let $V = \operatorname{span}\left\{\left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right)\right\}$, which is a line in $\mathbb{R}^3$. Then by Lemma B7.3,

$$V^\perp = \left\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} \cdot \left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right) = 0\right\} = \{\mathbf{x} \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}.$$

Thus, $V^\perp$ is the plane that has appeared in several recent examples.

Let V be a subspace of $\mathbb{R}^n$. Any vector $\mathbf{x} \in \mathbb{R}^n$ can be 'resolved' into a component in V and a component orthogonal to V. What this means is that we can write $\mathbf{x} = \mathbf{v} + \mathbf{w}$ for some $\mathbf{v} \in V$ and some $\mathbf{w} \in V^\perp$. In fact, there is only one possible choice of $\mathbf{v}$ and $\mathbf{w}$ with this property. We spend most of the rest of this section proving this.

Figure B.6 illustrates the first of our lemmas:

Lemma B7.5 *Let V be a subspace of $\mathbb{R}^n$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be an orthonormal basis of V. For $\mathbf{x} \in \mathbb{R}^n$, write*

$$P(\mathbf{x}) = \sum_{i=1}^{m} (\mathbf{x} \cdot \mathbf{v}_i)\mathbf{v}_i.$$

Then $P(\mathbf{x}) \in V$ and $\mathbf{x} - P(\mathbf{x}) \in V^\perp$.

We haven't proved yet that every subspace of $\mathbb{R}^n$ has an orthonormal basis (though we will soon). And the statement of this lemma doesn't assume that every subspace has an orthonormal basis. It merely says that *if* $\mathbf{v}_1, \ldots, \mathbf{v}_m$ is an orthonormal basis of V then the stated result holds.

Proof Evidently $P(\mathbf{x}) \in \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} = V$. To prove that $\mathbf{x} - P(\mathbf{x}) \in V^\perp$, it suffices (by Lemma B7.3) to show that $(\mathbf{x} - P(\mathbf{x})) \cdot \mathbf{v}_j = 0$ for each $j \in \{1, \ldots, m\}$. Let $j \in \{1, \ldots, m\}$. Then

$$P(\mathbf{x}) \cdot \mathbf{v}_j = \left(\sum_{i=1}^{m} (\mathbf{x} \cdot \mathbf{v}_i)\mathbf{v}_i\right) \cdot \mathbf{v}_j = \sum_{i=1}^{m} (\mathbf{x} \cdot \mathbf{v}_i)(\mathbf{v}_i \cdot \mathbf{v}_j) = \mathbf{x} \cdot \mathbf{v}_j$$

where the last step is by orthonormality. Hence $(\mathbf{x} - P(\mathbf{x})) \cdot \mathbf{v}_j = 0$, as required. □

In Section B4, we proved the extension lemma: every linearly independent set in a subspace V can be extended to a basis of V. Now we prove an analogous fact in the orthonormal context:

Lemma B7.6 (Orthonormal extension) *Let V be a subspace of $\mathbb{R}^n$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be orthonormal vectors in V. Then there is some orthonormal basis of V containing all of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.*

Proof Write $m = \dim V$ and $W = \operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. If $k = m$ then $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a basis of V (by Corollary B6.5), so we are done. Otherwise, $k < m$ and W is a proper subset of V, so we can choose some $\mathbf{y} \in V \setminus W$. Define

$$\mathbf{w} = \sum_{i=1}^{k} (\mathbf{y} \cdot \mathbf{v}_i)\mathbf{v}_i.$$

Then $\mathbf{w} \in W$, so $\mathbf{y} \neq \mathbf{w}$, so we can define

$$\mathbf{v}_{k+1} = \frac{\mathbf{y} - \mathbf{w}}{\|\mathbf{y} - \mathbf{w}\|}.$$

We have $\mathbf{y}, \mathbf{w} \in V$, so $\mathbf{v}_{k+1} \in V$. Now Lemma B7.5 implies that $\mathbf{v}_{k+1} \in W^\perp$, so $\mathbf{v}_{k+1} \cdot \mathbf{v}_i = 0$ for all $i \leq k$. Moreover, $\|\mathbf{v}_{k+1}\| = 1$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are orthonormal, it follows that $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}$ are orthonormal too.

We have now extended our original list $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of orthonormal vectors in V to a list $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}$ of orthonormal vectors in V. After doing this construction $m - k$ times, we obtain a list

$$\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_m$$

of orthonormal vectors in V. By Corollary B6.5, it is an orthonormal basis of V. □

Just as for non-orthonormal bases, we can use this extension lemma to prove that every subspace has at least one orthonormal basis—something we didn't know until now.

Proposition B7.7 *Every linear subspace of $\mathbb{R}^n$ has at least one orthonormal basis.*

Proof This follows from Lemma B7.6 by taking $\mathbf{v}_1, \ldots, \mathbf{v}_k$ to be the empty list (i.e. the list with $k = 0$), which is orthonormal. □

Given a subspace of $\mathbb{R}^n$, how could we actually *construct* an orthonormal basis of it? For instance, if I gave you a matrix and ask you to find an orthonormal of its kernel, how could you do it? Careful examination of the proof of Lemma B7.6 reveals an algorithm for this. We will meet this algorithm in the next chapter.

Just before Lemma B7.5, I promised we would show that, for a subspace V of $\mathbb{R}^n$, every vector in $\mathbb{R}^n$ can be uniquely resolved into a component in V and a component orthogonal to V. This is part (ii) of our next result.

Proposition B7.8 *Let V be a linear subspace of $\mathbb{R}^n$. Then:*

i. $V \cap V^{\perp} = \{\mathbf{0}\}$*;*

ii. for each $\mathbf{x} \in \mathbb{R}^n$*, there are unique* $\mathbf{v} \in V$ *and* $\mathbf{w} \in V^{\perp}$ *such that* $\mathbf{x} = \mathbf{v}+\mathbf{w}$*;*

iii. $\dim V + \dim V^{\perp} = n$.

Proof For (i), certainly $\mathbf{0} \in V \cap V^{\perp}$, since both V and $V^{\perp}$ are subspaces. Now take any element $\mathbf{x} \in V \cap V^{\perp}$. We have $\mathbf{x} \cdot \mathbf{v} = 0$ for all $\mathbf{v} \in V$, and in particular this holds when $\mathbf{v} = \mathbf{x}$. Hence $\mathbf{x} \cdot \mathbf{x} = 0$, that is, $\|\mathbf{x}\|^2 = 0$, so $\mathbf{x} = \mathbf{0}$.

For (ii), let $\mathbf{x} \in \mathbb{R}^n$. By Proposition B7.7, we can choose an orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of V. Now put

$$P(\mathbf{x}) = \sum_{i=1}^{m} (\mathbf{x} \cdot \mathbf{v}_i)\mathbf{v}_i.$$

Evidently

$$\mathbf{x} = P(\mathbf{x}) + (\mathbf{x} - P(\mathbf{x})),$$

and by Lemma B7.5, $P(\mathbf{x}) \in V$ and $\mathbf{x} - P(\mathbf{x}) \in V^{\perp}$. Putting $\mathbf{v} = P(\mathbf{x})$ and $\mathbf{w} = \mathbf{x} - P(\mathbf{x})$, we have $\mathbf{x} = \mathbf{v} + \mathbf{w}$ with $\mathbf{v} \in V$ and $\mathbf{w} \in V^{\perp}$.

To prove uniqueness, let $\mathbf{v}' \in V$ and $\mathbf{w}' \in V^{\perp}$ with $\mathbf{x} = \mathbf{v}' + \mathbf{w}'$. We must show that $\mathbf{v}' = \mathbf{v}$ and $\mathbf{w}' = \mathbf{w}$. Now,

$$\mathbf{v} + \mathbf{w} = \mathbf{x} = \mathbf{v}' + \mathbf{w}',$$

so $\mathbf{v} - \mathbf{v}' = \mathbf{w}' - \mathbf{w}$. But $\mathbf{v} - \mathbf{v}' \in V$ and $\mathbf{w}' - \mathbf{w} \in V^{\perp}$, so both belong to $V \cap V^{\perp}$, which is $\{\mathbf{0}\}$ by (i). Hence $\mathbf{v} = \mathbf{v}'$ and $\mathbf{w} = \mathbf{w}'$, as required.

For (iii), choose orthonormal bases $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of V and $\mathbf{w}_1, \ldots, \mathbf{w}_k$ of $V^{\perp}$ (as Proposition B7.7 allows us to do). I claim that $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{w}_1, \ldots, \mathbf{w}_k$ is an orthonormal basis of $\mathbb{R}^n$.

First we prove orthonormality. Since both $\mathbf{v}_1, \ldots, \mathbf{v}_m$ and $\mathbf{w}_1, \ldots, \mathbf{w}_k$ are orthonormal lists of vectors, all these vectors have unit length, the vectors $\mathbf{v}_i$ are orthogonal to each other, and the vectors $\mathbf{w}_j$ are orthogonal to each other. It only remains to show that $\mathbf{v}_i \cdot \mathbf{w}_j = 0$ for all i and j; but this is true because $\mathbf{v}_i \in V$ and $\mathbf{w}_j \in V^{\perp}$.

Now we show that $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{w}_1, \ldots, \mathbf{w}_k$ span $\mathbb{R}^n$. Let $\mathbf{x} \in \mathbb{R}^n$. By (ii), there exist $\mathbf{v} \in V$ and $\mathbf{w} \in V^{\perp}$ such that $\mathbf{x} = \mathbf{v} + \mathbf{w}$. But $\mathbf{v}$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m$ and $\mathbf{w}$ is a linear combination of $\mathbf{w}_1, \ldots, \mathbf{w}_k$, so $\mathbf{x}$ is a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{w}_1, \ldots, \mathbf{w}_k$, as required.

Hence $\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{w}_1, \ldots, \mathbf{w}_k$ is an orthonormal spanning set of $\mathbb{R}^n$. Since orthonormality implies linear independence (Lemma B6.4), it is a basis. But $m = \dim V$ and $k = \dim V^{\perp}$, and all bases of $\mathbb{R}^n$ have n elements, so $\dim V + \dim V^{\perp} = n$. □

For example, take a line L through the origin in $\mathbb{R}^3$. Then $L^\perp$ is a plane, so $\dim L + \dim L^\perp = 1 + 2 = 3$, in accordance with part (iii) of Proposition B7.8.

What if we take the orthogonal complement of an orthogonal complement? In the example just mentioned, $(L^\perp)^\perp$ is simply L again. In fact, this is a general phenomenon:

Corollary B7.9 *Let V be a linear subspace of $\mathbb{R}^n$. Then $(V^\perp)^\perp = V$.*

Proof We use the strategy described just after Lemma B5.9, proving that $V \subseteq (V^\perp)^\perp$ and $\dim V = \dim(V^\perp)^\perp$.

First we prove that $V \subseteq (V^\perp)^\perp$. Let $\mathbf{v} \in V$. To show that $\mathbf{v} \in (V^\perp)^\perp$, we have to show that $\mathbf{v} \cdot \mathbf{w} = 0$ for all $\mathbf{w} \in V^\perp$; but this is immediate from the definition of $V^\perp$.

By Proposition B7.8(iii) applied to V, we have $\dim V + \dim V^\perp = n$. But we can also apply Proposition B7.8(iii) to $V^\perp$, giving $\dim V^\perp + \dim(V^\perp)^\perp = n$. Hence $\dim V = \dim(V^\perp)^\perp$.

It follows from Lemma B5.9 that $V = (V^\perp)^\perp$. □

We have now shown that subspaces of $\mathbb{R}^n$ come in pairs. Each subspace has a kind of 'partner', its orthogonal complement, and Corollary B7.9 says that its partner's partner is itself. The dimensions of a subspace of $\mathbb{R}^n$ and its partner add up to n. For instance, in $\mathbb{R}^{100}$, the 23-dimensional subspaces are partnered with the 77-dimensional subspaces.

Remark B7.10 Orthogonal complements work a bit like set-theoretic complements. Write $E = \{1, \ldots, n\}$. Then every subset V of E has a complement or 'partner' $E \setminus V = \{x \in E : x \notin V\}$. The complement of the complement of V is V (that is, $E \setminus (E \setminus V) = V$), much as $(V^\perp)^\perp = V$ for linear subspaces. Moreover, $|E \setminus V| = n - |V|$ (where the bars mean 'number of elements'), much as $\dim V^\perp = n - \dim V$ in the linear situation.

This is only an analogy, but it is quite a fruitful one. Visualizing orthogonal complements in dimensions higher than three is difficult, and analogies like this can help our intuition.

Next time: we use all this theory to calculate stuff.

Summary of Chapter B

This is for you to fill in.

The most important definitions and ideas in this chapter

The most important results in this chapter

Points I didn't understand

Chapter C

Matrices and linear systems

To be read before the lecture of Monday, 22 October 2018

A **linear system** is a system of simultaneous linear equations, like this:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned} \tag{C:1}$$

Here a_{ij}, x_j and b_i all represent scalars, but we think of a_{ij} and b_i as ‘known’ and x_j as ‘unknown’. The fundamental questions about a linear system are these: are there any solutions? If so, how many? And how can we compute them?

Equations (C:1) can be written much more compactly as

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

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

It turns out that viewing a linear system in this way not only saves space, it is also very useful mathematically. A great deal is known about matrices, and this knowledge can be applied to help us analyse linear systems.

The first few sections of this chapter develop some of the theory of matrices. The last few are computational. Building on the theory developed, they provide methods that will enable you to answer natural computational questions about vectors, matrices and linear systems. For instance, how do you tell whether a given list of vectors is linearly independent? Given a subspace of $\mathbb{R}^n$, how do you find an orthogonal basis of it, or even any basis at all? How do you tell whether a given matrix is invertible, and if it is, how do you find its inverse?

By the end of this chapter, you will know how to answer all these questions, as well as the ‘fundamental questions’ about a linear system mentioned above.

On page 2, I wrote that school mathematics tends to emphasize computation and university mathematics tends to emphasize concepts. This is the most computational chapter of the course, but I hope you will notice how much the computational methods later in the chapter depend on the conceptual, theoretical developments that come first.

C1 Rank

As you know, there are two whole numbers associated with any matrix: the number of rows and the number of columns. In this section we will see that there is also a third, called its rank. Very crudely indeed, it indicates ‘how much stuff’ there is in the matrix. Later, we will see that it can also be interpreted in terms of linear systems.

Definition C1.1 Let A be an $m \times n$ matrix.

i. The **column rank** of A is $\dim(\mathrm{col}(A))$.

ii. The **row rank** of A is $\dim(\mathrm{row}(A))$.

iii. The **nullity** of A is $\dim(\ker(A))$.

Recall that $\mathrm{col}(A)$ denotes the column space of A, $\mathrm{row}(A)$ denotes the row space of A, and $\ker(A)$ denotes the kernel of A. They are subspaces of $\mathbb{R}^m$ or $\mathbb{R}^n$, as follows:

$$\mathrm{col}(A) \subseteq \mathbb{R}^m, \qquad \mathrm{row}(A) \subseteq \mathbb{R}^n, \qquad \ker(A) \subseteq \mathbb{R}^n.$$

Example C1.2 Consider the 4×3 matrix

$$A = \begin{pmatrix} 1 & 3 & 4 \\ 0 & 0 & 0 \\ 2 & 5 & 7 \\ 12 & 35 & 47 \end{pmatrix}.$$

Write the rows of A as $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4$ and the columns as $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$.

Looking at the row space, we have

$$\mathbf{x}_2 = 0\mathbf{x}_1 + 0\mathbf{x}_3, \quad \mathbf{x}_4 = 10\mathbf{x}_1 + \mathbf{x}_3,$$

so $\mathbf{x}_1^T$ and $\mathbf{x}_3^T$ span $\mathrm{row}(A)$. (Recall from page 50 that strictly speaking, the row space of A is spanned by the *transposes of* the rows, rather than the rows themselves.) Moreover, $\mathbf{x}_1^T$ and $\mathbf{x}_3^T$ are linearly independent (by Example B3.2(ii)), so they form a basis of $\mathrm{row}(A)$. Hence the row rank of A is 2.

Now looking at the column space, we have

$$\mathbf{y}_3 = \mathbf{y}_1 + \mathbf{y}_2,$$

and $\mathbf{y}_1, \mathbf{y}_2$ are linearly independent, so they form a basis of $\mathrm{col}(A)$. Hence the column rank of A is also 2.

A short calculation shows that the kernel of A is spanned by $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$. Hence $\ker(A)$ is 1-dimensional, that is, the nullity of A is 1.

In this example, we have calculated the row rank, column rank and nullity in a rather improvised way. Later in the chapter, we will see how to do calculations of this type systematically.

We now seem to have not just three but *five* whole numbers associated with any matrix: the number of rows, the number of columns, the row rank, the column rank, and the nullity. However, this is not really the case. We will prove that the row rank is equal to the column rank. This is perhaps a surprise, as $\mathrm{row}(A)$ is a subspace of $\mathbb{R}^n$ but $\mathrm{col}(A)$ is a subspace of $\mathbb{R}^m$. Once we've done that, we'll refer to both as just 'rank'. So there are really only four numbers in play.

We will also prove that rank + nullity = number of columns. So, if you know the number of columns of a matrix and its rank then you know its nullity. So there are actually only *three* numbers in play.

Before we prove these big results, we make a small but important observation.

Lemma C1.3 *Let A be an $m \times n$ matrix. Then the row rank and column rank of A are both less than or equal to $\min\{m, n\}$.*

Once we have shown that the row rank is equal to the column rank, this statement will simply say that $\mathrm{rank}(A) \leq \min\{m, n\}$.

Proof We prove it for column rank. The proof for row rank is similar.

Write the columns of A as $\mathbf{y}_1, \ldots, \mathbf{y}_n \in \mathbb{R}^m$. Since $\mathbf{y}_1, \ldots, \mathbf{y}_n$ span $\mathrm{col}(A)$ (by definition of $\mathrm{col}(A)$), Proposition B5.8 implies that $\dim(\mathrm{col}(A)) \leq n$. On the other hand, $\mathrm{col}(A)$ is a subspace of $\mathbb{R}^m$, so $\dim(\mathrm{col}(A)) \leq m$ by Lemma B5.9.□

Now we prove the first major result about rank. It is called the rank-nullity theorem (or the rank theorem). Figure C.1 illustrates the proof. Don't take the figure too literally: it shows various relationships between the vectors and subspaces involved, but it's not a *geometric* diagram.

Theorem C1.4 (Rank-nullity, column version) *For any matrix A,*

$$\textit{column-rank}(A) + \textit{nullity}(A) = \textit{number of columns of } A.$$

Proof Let A be an $m \times n$ matrix. Choose a basis $\mathbf{w}_1, \ldots, \mathbf{w}_\ell$ of $\mathrm{col}(A)$ and a basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of $\ker(A)$. Then ℓ is the column-rank of A and k is the nullity of A, so we have to prove that $k + \ell = n$.

By Lemma B2.7, we can choose $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_{k+\ell} \in \mathbb{R}^n$ such that

$$A\mathbf{v}_{k+1} = \mathbf{w}_1,\ A\mathbf{v}_{k+2} = \mathbf{w}_2,\ \ldots,\ A\mathbf{v}_{k+\ell} = \mathbf{w}_\ell.$$

I claim that $\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_{k+\ell}$ is a basis of $\mathbb{R}^n$. If we can show this then it will follow that $k + \ell = n$ (since all bases of $\mathbb{R}^n$ have n elements), and so the proof will be finished.

First, we prove that $\mathbf{v}_1, \ldots, \mathbf{v}_{k+\ell}$ span $\mathbb{R}^n$. Let $\mathbf{x} \in \mathbb{R}^n$. Then $A\mathbf{x} \in \mathrm{col}(A)$ by Lemma B2.7. But $\mathbf{w}_1, \ldots, \mathbf{w}_\ell$ span $\mathrm{col}(A)$, so

$$A\mathbf{x} = d_1\mathbf{w}_1 + \cdots + d_\ell\mathbf{w}_\ell$$

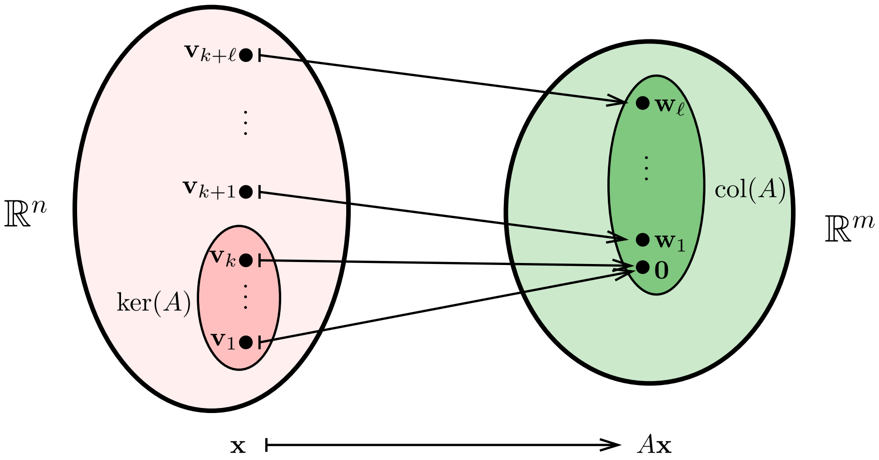


Figure C.1: Schematic diagram of the proof of Theorem C1.4

for some scalars $d_1, \dots, d_\ell$. Since $\mathbf{w}_1 = A\mathbf{v}_{k+1}$ etc.,

$$A\mathbf{x} = d_1 A\mathbf{v}_{k+1} + \cdots + d_\ell A\mathbf{v}_{k+\ell} = A(d_1\mathbf{v}_{k+1} + \cdots + d_\ell \mathbf{v}_{k+\ell}).$$

Put $\hat{\mathbf{x}} = d_1\mathbf{v}_{k+1} + \cdots + d_\ell\mathbf{v}_{k+\ell} \in \mathbb{R}^n$. Then $A\mathbf{x} = A\hat{\mathbf{x}}$, so $A(\mathbf{x} - \hat{\mathbf{x}}) = \mathbf{0}$, or equivalently $\mathbf{x} - \hat{\mathbf{x}} \in \ker(A)$. But $\mathbf{v}_1, \dots, \mathbf{v}_k$ span $\ker(A)$, so

$$\mathbf{x} - \hat{\mathbf{x}} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

for some scalars $c_1, \dots, c_k$. Substituting in the definition of $\hat{\mathbf{x}}$ and rearranging gives

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + d_1\mathbf{v}_{k+1} + \cdots + d_\ell\mathbf{v}_{k+\ell}.$$

So $\mathbf{x} \in \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k+\ell}\}$, as required.

Second, we prove that $\mathbf{v}_1, \dots, \mathbf{v}_{k+\ell}$ are linearly independent. Let $c_1, \dots, c_k, d_1, \dots, d_\ell$ be scalars such that

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + d_1\mathbf{v}_{k+1} + \cdots + d_\ell\mathbf{v}_{k+\ell} = \mathbf{0}. \tag{C:2}$$

Multiplying each side by A gives

$$c_1A\mathbf{v}_1 + \cdots + c_kA\mathbf{v}_k + d_1A\mathbf{v}_{k+1} + \cdots + d_\ell A\mathbf{v}_{k+\ell} = \mathbf{0}.$$

But $\mathbf{v}_1, \dots, \mathbf{v}_k \in \ker(A)$ and $A\mathbf{v}_{k+i} = \mathbf{w}_i$, so this gives

$$d_1\mathbf{w}_1 + \cdots + d_\ell\mathbf{w}_\ell = \mathbf{0}.$$

Now $\mathbf{w}_1, \dots, \mathbf{w}_\ell$ are linearly independent, so $d_1 = \cdots = d_\ell = 0$. Hence equation (C:2) reduces to

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

But $\mathbf{v}_1, \dots, \mathbf{v}_k$ are also linearly independent, so $c_1 = \cdots = c_k = 0$. We have now shown that all the scalars c_j and d_i are zero. Hence $\mathbf{v}_1, \dots, \mathbf{v}_{k+\ell}$ are linearly independent, completing the proof. □

This result is very important, but so far we have proved nothing about *row* rank. The key to understanding the row space is the following lemma. It says that the row space and kernel of a matrix are 'partners' in the sense of page 69: each is the orthogonal complement of the other.

Lemma C1.5 $\ker(A) = \mathrm{row}(A)^\perp$ *and* $\mathrm{row}(A) = \ker(A)^\perp$*, for any matrix* A*.*

Proof Let A be an $m \times n$ matrix. We first show that $\ker(A) = \mathrm{row}(A)^\perp$. Both sides are subspaces of $\mathbb{R}^n$, so we have to show that for $\mathbf{x} \in \mathbb{R}^n$,

$$A\mathbf{x} = \mathbf{0} \iff \mathbf{x} \in \mathrm{row}(A)^\perp.$$

Let $\mathbf{x} \in \mathbb{R}^n$. Write the rows of A as $\mathbf{r}_1, \ldots, \mathbf{r}_m$; thus, $\mathbf{r}_i$ is a row vector and $\mathbf{r}_i^T$ is a column vector. By definition of matrix multiplication and dot product,

$$A\mathbf{x} = \begin{pmatrix} \mathbf{r}_1^T \cdot \mathbf{x} \\ \vdots \\ \mathbf{r}_m^T \cdot \mathbf{x} \end{pmatrix}.$$

Hence

$$\begin{aligned} A\mathbf{x} = \mathbf{0} &\iff \mathbf{r}_1^T \cdot \mathbf{x} = \cdots = \mathbf{r}_m^T \cdot \mathbf{x} = 0 \\ &\iff \mathbf{x} \in \left(\mathrm{span}\{\mathbf{r}_1^T, \ldots, \mathbf{r}_m^T\}\right)^\perp && \text{by Lemma B7.3} \\ &\iff \mathbf{x} \in \mathrm{row}(A)^\perp && \text{by definition of } \mathrm{row}(A). \end{aligned}$$

This proves that $\ker(A) = \mathrm{row}(A)^\perp$. Applying $(\ )^\perp$ to each side and using Corollary B7.9 then gives $\ker(A)^\perp = \mathrm{row}(A)$. □

We have now done all the work and can read off two more big theorems.

Theorem C1.6 (Rank-nullity, row version) *For any matrix* A*,*

$$\textit{row-rank}(A) + \textit{nullity}(A) = \textit{number of columns of } A.$$

Proof Let A be an $m \times n$ matrix. Then $\mathrm{row}(A)$ is a subspace of $\mathbb{R}^n$, so $\dim(\mathrm{row}(A)) + \dim(\mathrm{row}(A)^\perp) = n$ by Proposition B7.8. By Lemma C1.5, this is equivalent to $\dim(\mathrm{row}(A)) + \dim(\ker(A)) = n$, which is exactly what the theorem states. □

Theorem C1.7 *The row rank of a matrix is equal to its column rank.*

Proof This is immediate from Theorems C1.4 and C1.6. □

We can now define the **rank** of a matrix A, written as $\mathrm{rank}(A)$, to be either the row rank or the column rank of A: they're the same! So both versions of the rank-nullity theorem state that

$$\mathrm{rank} + \mathrm{nullity} = \text{number of columns}.$$

But we needed to prove both versions in order to deduce that the two kinds of rank were the same.

Example C1.8 Let us check that these theorems hold for the 4×3 matrix A of Example C1.2. First, the row rank is indeed equal to the column rank: both are 2. So, $\text{rank}(A) = 2$. Second, the rank plus the nullity is $2 + 1 = 3$, which is indeed equal to the number of columns of A.

At the start of this section, I said that very crudely, you could think of the rank of a matrix as the 'amount of stuff in it'. A more refined interpretation is that rank is the 'effective number of rows', or equivalently the 'effective number of columns'. For instance, our matrix A has 3 columns, but the third is in the span of the first two (which are linearly independent), so there are 'effectively' only two columns. Correspondingly, the rank is 2.

Phrases such as 'amount of stuff' and 'effective number' are informal and unrigorous, but may help you to understand what rank means.

C2 Invertibility

Back in Theorem A5.3, we made two nontrivial statements about invertible matrices: that they are always square, and that for square matrices A and B of the same size, $AB = I \iff BA = I$. But we didn't prove either statement. In fact, we observed that *even for* 2×2 *matrices*, proving that $AB = I \iff BA = I$ isn't a pushover.

But we now have the technology to prove the theorem. We can prove the first part immediately:

Theorem C2.1 *Every invertible matrix is square.*

Proof Let A be an $m \times n$ invertible matrix. We must prove that $m = n$.

The kernel of A is trivial, since if $\mathbf{x} \in \mathbb{R}^n$ with $A\mathbf{x} = \mathbf{0}$ then $A^{-1}A\mathbf{x} = A^{-1}\mathbf{0}$, so $\mathbf{x} = \mathbf{0}$. Hence A has nullity 0. The rank-nullity theorem then implies that $\text{rank}(A) = n$. But $\text{rank}(A) \leq m$ by Lemma C1.3, so $n \leq m$.

We have just shown that the number of columns in an invertible matrix is less than or equal to the number of rows. Applying this to the $n \times m$ invertible matrix A^{-1} tells us that $m \leq n$. Hence $m = n$. □

For the rest of this section, we will build up to the proof of the second part of Theorem A5.3: that for square matrices A and B of the same size, $AB = I \iff BA = I$. Along the way, we will establish a large number of equivalent conditions for invertibility.

We will soon show that the following four conditions are equivalent to invertibility, but for now we just show that they are equivalent to each other:

Lemma C2.2 *Let A be an $n \times n$ matrix. The following are equivalent:*

i. the columns of A are linearly independent;

ii. the columns of A span $\mathbb{R}^n$;

iii. the columns of A are a basis of $\mathbb{R}^n$;

iv. $\text{rank}(A) = n$.

Proof The equivalence of conditions (i)–(iii) follows from Proposition B5.4. Also, (ii) is equivalent to (iv) since by Lemma B5.10, $\text{col}(A) = \mathbb{R}^n$ if and only if $\dim(\text{col}(A)) = n$. □

Now we can prove that all those conditions, and more besides, are equivalent to invertibility.

Theorem C2.3 (Equivalent conditions for invertibility, part 1) *Let A be an $n \times n$ matrix. The following are equivalent:*

i. *A is invertible;*

ii. *there exists an $n \times n$ matrix A' such that $A'A = I$;*

iii. $\ker(A) = \{\mathbf{0}\}$*;*

iv. $\operatorname{nullity}(A) = 0$*;*

v. *the columns of A are linearly independent;*

vi. *the columns of A span $\mathbb{R}^n$;*

vii. *the columns of A are a basis of $\mathbb{R}^n$;*

viii. $\operatorname{rank}(A) = n$*;*

ix. *for all $\mathbf{b} \in \mathbb{R}^n$, there is exactly one $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \mathbf{b}$;*

x. *for all integers $p \geq 0$ and all $n \times p$ matrices B, there is exactly one $n \times p$ matrix X such that $AX = B$.*

Proof We have just shown that (v)–(viii) are equivalent, and Lemma B5.10 implies that (iii) is equivalent to (iv). So, it suffices to prove that (i)$\Longrightarrow$(ii)$\Longrightarrow$(iii)$\Longrightarrow$(v) and (vii)$\Longrightarrow$(ix)$\Longrightarrow$(x)$\Longrightarrow$(i). (Suggestion: draw a diagram showing all the implications.)

(i)$\Longrightarrow$(ii) is immediate from the definition of invertibility.

(ii)$\Longrightarrow$(iii): take some A' such that $A'A = I$, and let $\mathbf{x} \in \ker(A)$. Then $A'(A\mathbf{x}) = A'\mathbf{0} = \mathbf{0}$. But also $A'(A\mathbf{x}) = (A'A)\mathbf{x} = \mathbf{x}$, so $\mathbf{x} = \mathbf{0}$.

(iii)$\Longrightarrow$(v): assume that $\ker(A) = \{\mathbf{0}\}$, and write the columns of A as $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Let $c_1, \ldots, c_n$ be scalars such that $\sum c_i \mathbf{v}_i = \mathbf{0}$. By Lemma A4.3(i),

$$A\mathbf{c} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

where $\mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. Hence $\mathbf{c} \in \ker(A) = \{\mathbf{0}\}$, and so $c_1 = \cdots = c_n = 0$. Thus, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent.

(vii)$\Longrightarrow$(ix): again, write the columns of A as $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Let $\mathbf{b} \in \mathbb{R}^n$. By Lemma A4.3(i), we must prove that there is exactly one $\mathbf{x} \in \mathbb{R}^n$ such that

$$x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \mathbf{b}.$$

But $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of $\mathbb{R}^n$, so this follows from Lemma B4.3(iii).

(ix)$\Longrightarrow$(x): assume (ix), and let B be an $n \times p$ matrix. Write the columns of B as $\mathbf{b}_1, \ldots, \mathbf{b}_p \in \mathbb{R}^n$. By (ix), there are unique $\mathbf{x}_1, \ldots, \mathbf{x}_p \in \mathbb{R}^n$ such that

$$A\mathbf{x}_1 = \mathbf{b}_1, \ \ldots, \ A\mathbf{x}_p = \mathbf{b}_p.$$

By Lemma A4.3(iii), an equivalent statement is that there is a unique $n \times p$ matrix X such that $AX = B$. This proves (x).

(x)$\Longrightarrow$(i): assume (x). Taking $B = I_n$, there is a unique matrix A' such that $AA' = I$. Observe now that $A(A'A) = (AA')A = IA = A$ and $AI = A$. So, the equation $AX = A$ holds for both $X = A'A$ and $X = I$. But taking $B = A$ in (x) tells us that there is only one matrix X such that $AX = A$. Hence $A'A = I$. But also $AA' = I$, so A is invertible. □

What's the point of having so many equivalent conditions for invertibility? First, if you want to show that some matrix is invertible, you can verify *any* of these conditions—whichever one is easiest. Second, if you want to show that some matrix is not invertible, you can prove the failure of whichever one of these conditions you like. Third, if you have some matrix that you already know to be invertible, then you can immediately deduce that it has *all* these properties... and fourth, if you have some matrix that you know not to be invertible, then it fails every one of them.

Further equivalent conditions for invertibility can be obtained by swapping the roles of the rows and columns. The neatest way to handle this is to use transposes, as follows.

Theorem C2.4 (Equivalent conditions for invertibility, part 2) *Let A be an $n \times n$ matrix. The following are equivalent:*

i. A is invertible;

ii. there exists an $n \times n$ matrix A' such that $AA' = I$;

iii. the rows of A are linearly independent;

iv. the rows of A span $\mathbb{R}^n$;

v. the rows of A are a basis of $\mathbb{R}^n$.

Proof We apply Theorem C2.3 to the matrix A^T.

Condition (i) of the present theorem states that A is invertible, which by Lemma A5.10 is equivalent to A^T being invertible. This is condition (i) of Theorem C2.3 applied to A^T.

Condition (ii) of the present theorem holds if and only if there is some A' satisfying $(AA')^T = I$, if and only if there is some A' satisfying $(A')^T A^T = I$, if and only if there is some A'' satisfying $A''A^T = I$. This is condition (ii) of Theorem C2.3 applied to A^T.

The rows of A are the columns of A^T, so condition (iii) of the present theorem is equivalent to condition (v) of Theorem C2.3 applied to A^T. Similarly, conditions (iv) and (v) of the present theorem are equivalent to conditions (vi) and (vii) of Theorem C2.3 applied to A^T.

So each of the five conditions in the present theorem is equivalent to one of the conditions of Theorem C2.3 applied to A^T. Since all the conditions of that theorem are equivalent, so too are the five conditions above. □

We now achieve our goal, easily deducing the second half of Theorem A5.3:

Corollary C2.5 *Let A and B be $n \times n$ matrices. Then $AB = I \iff BA = I$. Moreover, if $AB = I$ or $BA = I$ then A is invertible and $B = A^{-1}$.*

Proof Suppose that $AB = I$. Then A is invertible by Theorem C2.4, so $B = A^{-1}AB = A^{-1}I = A^{-1}$, so $BA = A^{-1}A = I$. A similar argument applies if $BA = I$. □

So although AB and BA are not usually equal for $n \times n$ matrices A and B, they *are* equal if AB or BA is I.

C3 Determinants

This course swings back and forth between geometric and algebraic viewpoints. The section on invertibility was purely algebraic. This section involves both viewpoints: determinants are defined algebraically, but have a strong geometric interpretation in terms of volume.

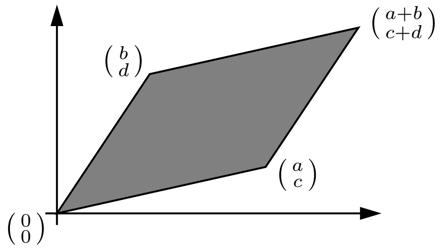
You've already met determinants for 2×2 and maybe 3×3 matrices, but here we'll consider arbitrary $n \times n$ matrices.

What's the point of determinants? First, they'll give us yet another condition for invertibility of a matrix, and a (very inefficient) method for computing the inverse. Second, they turn out to be essential for changing variables in several-variable integrals. (We won't cover that in this course, but the word to watch out for elsewhere is *Jacobian*.) This is closely related to the third reason: a determinant can be understood as a kind of volume scale factor.

Intuitive background

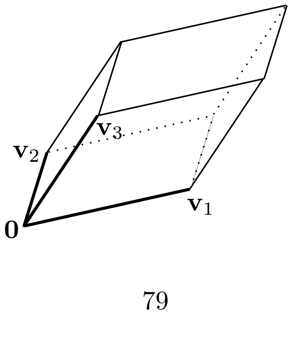
Let us begin with 2×2 matrices $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. As you know from Example A5.4 or elsewhere, the determinant of A is defined as $\det(A) = ad - bc$. But what does $ad - bc$ mean geometrically?

Consider the two columns of our matrix, $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$. Draw a parallelogram:


$\left(\begin{smallmatrix} a+b \\ c+d \end{smallmatrix}\right)$
$\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$
$\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$
$\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$

It's not too hard to see that the area of this parallelogram is $ad - bc$ (exercise). So, the determinant of A is the area of the parallelogram whose edge-vectors are the columns of A.

We can try something similar in dimension three. Let A be a 3×3 matrix. The three columns $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ of A are vectors in $\mathbb{R}^3$, and we can think about the volume of the 'squashed cube' whose edges are those three vectors:



A 'squashed cube' is called a **parallelepiped**. So to continue the pattern that we observed in two dimensions, we'd like to define the determinant of a 3×3 matrix A to be the volume of the parallelepiped whose edge-vectors are the columns of A.

Below, we give an algebraic definition of the determinant of an $n \times n$ matrix. It can be shown that in the 3×3 case, the determinant really is equal to the volume of our parallelepiped. And once suitable definitions of 'higher-dimensional volume' and 'higher-dimensional parallelepiped' have been made, it can be shown that the same pattern continues in all dimensions.

There is a subtlety concerning signs. Determinants can sometimes be negative, whereas areas and volumes cannot. For example, $\det\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) = -1$, and yet the parallelogram with edge-vectors $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ is the unit square

$$\left\{\mathbf{x} \in \mathbb{R}^2 : 0 \leq x_1 \leq 1 \text{ and } 0 \leq x_2 \leq 1\right\},$$

which has area $+1$. Actually, the area of the parallelogram (in the two-dimensional case) or volume of the parallelepiped (in the three-dimensional case) is *the absolute value of* the determinant of the matrix. The sign of the determinant depends on the order in which the edge-vectors are listed, as Proposition C3.4(i) makes clear.

Definition and examples

We define the determinant $\det(A)$ of an $n \times n$ matrix A by induction on n.

The **determinant** of a 1×1 matrix (a) is a.

Now let $n \geq 2$, and suppose inductively that we have already defined the determinant of any $(n-1) \times (n-1)$ matrix. Let $A = (A_{ij})$ be an $n \times n$ matrix. Whenever $i, j \in \{1, \ldots, n\}$, write $A[i,j]$ for the $(n-1) \times (n-1)$ matrix obtained from A by deleting the ith row and jth column. (This is not standard notation.) The **determinant** of A is

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \det(A[1,j]).$$

Examples C3.1 i. Let $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Then $A[1,1] = (d)$ and $A[1,2] = (c)$, so

$$\det(A) = (-1)^{1+1} a \det((d)) + (-1)^{1+2} b \det((c)) = ad - bc.$$

So, our new definition agrees with the definition in Example A5.4.

ii. Let

$$A = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

be a 3×3 matrix. Then

$$\det(A) = A_{11} \det \begin{pmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{pmatrix} - A_{12} \det \begin{pmatrix} A_{21} & A_{23} \\ A_{31} & A_{33} \end{pmatrix} + A_{13} \det \begin{pmatrix} A_{21} & A_{22} \\ A_{31} & A_{32} \end{pmatrix}$$

and each of the individual 2×2 determinants can be worked out using the formula in the previous example.

iii. Similar calculations can be performed for larger matrices. For instance, if A is a 4×4 matrix then

$$\det(A) = A_{11}\det(A[1,1]) - A_{12}\det(A[1,2]) + A_{13}\det(A[1,3]) - A_{14}\det(A[1,4]),$$

and $A[1,1]$, $A[1,2]$, $A[1,3]$ and $A[1,4]$ are 3×3 matrices, so their determinants can be computed as in the previous example.

iv. An easy proof by induction shows that $\det(I_n) = 1$ for all n.

v. More generally, you can show by induction that if

$$A = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_n \end{pmatrix}$$

then $\det(A) = c_1 c_2 \cdots c_n$.

The definition of determinant appears to give a special status to the first row. The next result says that this is an illusion: the same kind of expansion can be done along any row, or even any column. Pay attention to the is and js!

Proposition C3.2 *Let A be an $n \times n$ matrix. Then for each $i \in \{1, \ldots, n\}$,*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(A[i,j]).$$

Moreover, for each $j \in \{1, \ldots, n\}$,

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det(A[i,j]).$$

Proof Omitted. □

We can use this result to speed up calculations of determinants, as in the next example. To handle the signs $(-1)^{i+j}$, it is useful to notice that they form a chessboard pattern:

$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Example C3.3 Suppose we want to compute the determinant of

$$A = \begin{pmatrix} 3 & 1 & 0 & 2 \\ 10 & 4 & 3 & -9 \\ 4 & -1 & 0 & 4 \\ -7 & 2 & 0 & 0 \end{pmatrix}.$$

If we simply use the definition of determinant then we will have three summands to deal with, one for each nonzero entry in the top row. However, Proposition C3.2 allows us to expand along any other row or column. We make the calculation easier by expanding down the third column, which gives:

$$\det(A) = -3\det\begin{pmatrix} 3 & 1 & 2 \\ 4 & -1 & 4 \\ -7 & 2 & 0 \end{pmatrix}.$$

(The minus sign on the right-hand side is because $(-1)^{2+3} = -1$, as in the chessboard pattern.) To compute the determinant of this 3×3 matrix, it makes life easier if we expand along its third row (or column). This gives

$$\begin{aligned} \det(A) &= -3\left[-7\det\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} - 2\det\begin{pmatrix} 3 & 2 \\ 4 & 4 \end{pmatrix}\right] \\ &= -3\big[-7(1 \times 4 + 2 \times 1) - 2(3 \times 4 - 2 \times 4)\big] \\ &= -3[-7 \times 6 - 2 \times 4] = 150. \end{aligned}$$

Properties of determinants

Here are some properties of determinants, stated without proof.

Proposition C3.4 *Let A be an $n \times n$ matrix, with rows $\mathbf{r}_1, \ldots, \mathbf{r}_n$.*

i. Let B be the matrix obtained from A by swapping rows i and j (where $i \neq j$). Then $\det(B) = -\det(A)$.

ii. Let B be the matrix obtained from A by multiplying the ith row by a scalar c. Then $\det(B) = c\det(A)$.

iii. If some row $\mathbf{r}_i$ is $\mathbf{0}$ then $\det(A) = 0$.

iv. Let $\mathbf{r}'_i$ be an n-dimensional row vector, write A' for the matrix with rows

$$\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}'_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_n,$$

and write B for the matrix with rows

$$\mathbf{r}_1, \ldots, \mathbf{r}_{i-1}, \mathbf{r}_i + \mathbf{r}'_i, \mathbf{r}_{i+1}, \ldots, \mathbf{r}_n.$$

Then $\det(B) = \det(A) + \det(A')$.

v. $\det(A^T) = \det(A)$.

vi. $\det(AB) = \det(A)\det(B)$ for any $n \times n$ matrix B.

Proof Omitted. □

Part (v) implies that for every property of determinants involving rows, there is a similar property involving columns. For example, if some column of A is zero then $\det(A) = 0$.

Corollary C3.5 *Let A be an $n \times n$ matrix. If two rows of A are identical, or two columns of A are identical, then $\det(A) = 0$.*

Proof Suppose that two rows are identical. Then $\det(A) = -\det(A)$ by Proposition C3.4(i), so $\det(A) = 0$.

If two columns of A are identical then two rows of A^T are identical, so $\det(A^T) = 0$, so $\det(A) = 0$ by Proposition C3.4(v). □

Corollary C3.6 *Every invertible matrix A has nonzero determinant, and $\det(A^{-1}) = 1/\det(A)$.*

Proof Let A be an invertible matrix. Then $AA^{-1} = I$, so $\det(A)\det(A^{-1}) = \det(I) = 1$ by Proposition C3.4(vi) and Example C3.1(iv). The result follows.□

We will also prove the converse: a matrix with nonzero determinant is invertible. For this, we introduce some terminology. The (i, j)-**cofactor** of A is $C_{ij} = (-1)^{i+j}\det(A[i, j])$. So by Proposition C3.2,

$$\det(A) = \sum_{j=1}^{n} A_{ij}C_{ij}$$

for any $i \in \{1, \ldots, n\}$.

The **adjugate** of A is the $n \times n$ matrix $\operatorname{adj}(A)$ whose (i, j)-entry is C_{ji}. Note the reversal of the indices! (The adjugate is sometimes called the **classical adjoint**, or simply the **adjoint**. This terminology is problematic, since the word ‘adjoint’ also has another, different, meaning in linear algebra.)

Example C3.7 The adjugate of a 2×2 matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is

$$\operatorname{adj}(A) = \begin{pmatrix} C_{11} & C_{21} \\ C_{12} & C_{22} \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Notice that in this case, $A\operatorname{adj}(A) = \det(A)I$. We will now show that this is actually true for all square matrices, not just 2×2.

Proposition C3.8 *$A\operatorname{adj}(A) = \det(A)I$ for all square matrices A.*

Proof Let A be an $n\times n$ matrix. Both $A\operatorname{adj}(A)$ and $\det(A)I$ are $n\times n$ matrices, so it remains to prove that they have the same entries. We use the convention that the (i, j)-entry of a matrix M is written as M_{ij}.

Let $i, k \in \{1, \ldots, n\}$. We must show that $(A\operatorname{adj}(A))_{ik}$ is equal to $\det(A)$ if $i = k$, or 0 if $i \neq k$. We have

$$(A\operatorname{adj}(A))_{ik} = \sum_{j=1}^{n} A_{ij}C_{kj} = \sum_{j=1}^{n} (-1)^{k+j} A_{ij}\det(A[k, j]).$$

If $i = k$ then this sum is equal to $\det(A)$, by Proposition C3.2, as required.

Now suppose that $i \neq k$. Let A' be the $n \times n$ matrix obtained from A by replacing the kth row by the ith row (and leaving all the other rows alone, including the ith). The rows of A' and A are the same apart from the kth, so $A'[k, j] = A[k, j]$ for all j. Also, $A'_{kj} = A_{ij}$ for all j. Hence

$$(A\operatorname{adj}(A))_{ik} = \sum_{j=1}^{n} (-1)^{k+j} A'_{kj}\det(A'[k, j]),$$

which is equal to $\det(A')$ by Proposition C3.2. But two rows of A' are equal, so $\det(A') = 0$ by Corollary C3.5. Hence $(A\operatorname{adj}(A))_{ik} = 0$, as required. □

Theorem C3.9 (Equivalent conditions for invertibility, part 3) *Let A be an $n \times n$ matrix. Then A is invertible if and only if $\det A \neq 0$, and in that case $A^{-1} = \frac{1}{\det A} \operatorname{adj}(A)$.*

Proof 'Only if' is part of Corollary C3.6. For 'if', suppose that $\det A \neq 0$. Then $A\big(\frac{1}{\det A} \operatorname{adj}(A)\big) = I$ by Proposition C3.8, so by Corollary C2.5, A is invertible with inverse $\frac{1}{\det A} \operatorname{adj}(A)$. □

Example C3.10 By Theorem C3.9, a 2×2 matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is invertible if and only if $ad \neq bc$, and in that case,

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Let us finish by seeing how Theorem C3.9 fits with the volume-based understanding of determinants. We introduced determinants for 3×3 matrices by saying that $\det(A)$ (or really $|\det(A)|$) is the volume of the parallelepiped whose edge-vectors are the columns $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ of A. According to the theorem we just proved, A is invertible exactly when that volume is nonzero. Does that make sense?

The volume of the parallelepiped is zero if and only if $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ all lie on some plane. This happens if and only if $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly dependent. But we know from Theorem C2.3 that they are linearly dependent if and only if A is not invertible. So, the volume of the parallelepiped is *non*zero if and only if A *is* invertible. Since the volume is $|\det(A)|$, that fits exactly with Theorem C3.9.

C4 Linear systems

We now begin our study of linear systems (defined on page 71). Linear systems include collections of equations like this:

$$\begin{aligned} 2x - 3y &= z + 6 \\ y + 5 &= -2z - 3(x - 2) \\ 0 &= x + y + z \\ 3x + 4 &= 2y. \end{aligned}$$

Although this is not literally in the form of equations (C:1) (page 71), it is easily put into that form by tidying up the equations so that all the variables are on the left and all the constants are on the right.

In the rest of this chapter, we will address the 'fundamental questions' listed after equations (C:1), using our earlier observation that a linear system can be expressed in the form $A\mathbf{x} = \mathbf{b}$. But let us start in a very elementary way, by considering some small examples.

Example C4.1 A 2×2 linear system consists of equations

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{aligned}$$

in variables x and y. Assuming that a_{11} and a_{12} are not both zero, the first equation represents a straight line in the (x, y)-plane. Assuming that a_{21} and a_{22} are not both zero, the second represents a straight line too.

The set of solutions is the set of points that lie on both lines. There are several possibilities:

- The lines intersect at a single point. Then the system has exactly one solution.
- The lines are parallel but not the same. Then the system has no solutions.
- The lines are the same. Then the system has infinitely many solutions.

How can we actually solve the equations? Multiply the first by a_{21} and the second by a_{11}, then subtract. This gives

$$(a_{11}a_{22} - a_{12}a_{21})y = a_{11}b_2 - a_{21}b_1.$$

Assuming that $a_{11}a_{22} - a_{12}a_{21} \neq 0$, this gives

$$y = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{12}a_{21}}$$

from which it follows that

$$x = \frac{a_{22}b_1 - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}.$$

So as long as $a_{11}a_{22} - a_{12}a_{21} \neq 0$, there is a unique solution.

As you may have noticed,

$$a_{11}a_{22} - a_{12}a_{21} = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

So, we have just shown that if this matrix is invertible then the system has a unique solution. We will come back to this connection later.

I leave you the exercise of investigating the case where $a_{11}a_{22} - a_{12}a_{21} = 0$.

Example C4.2 We want a methodical way of calculating the solutions of any linear system. A 2 × 2 system isn't big enough to illustrate the method, so let's try a 3 × 3 system, say

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

We start by trying to eliminate the xs from all but one equation. We could do this by subtracting suitable multiples of the first equation from each of the other two, but the numbers will be easier if we use the second equation instead. So let's begin by swapping the first two equations:

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

Obviously, changing the order of the equations makes no difference to the solutions! Now let's subtract 2 times the first equation from the second equation and subtract 3 times the first equation from the third equation. This gives:

$$\begin{aligned} x + y + z &= 4 \\ 0x + y - 3z &= -5 \\ 0x - 7y - 2z &= -11. \end{aligned}$$

Doing this doesn't change the solutions either: numbers x, y and z satisfy the old equations if and only if they satisfy the new ones.

We're now done with x, having eliminated it from all but one equation. Next let's eliminate y, by subtracting 1 times the second equation from the first and adding 7 times the second equation to the third. This gives:

$$\begin{aligned} x + 0y + 4z &= 9 \\ 0x + y - 3z &= -5 \\ 0x + 0y - 23z &= -46. \end{aligned}$$

Again, this has the same solutions as the original equations. We might as well simplify the third equation by taking out a factor of -23:

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

(which again, doesn't change the solutions). Finally, we eliminate the z by subtracting 4 times the third equation from the first equation and adding 3 times the third to the second. This gives

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

or equivalently $x = 1$, $y = 1$ and $z = 2$. So this particular linear system has a unique solution.

Remark C4.3 You might think that the method above misses out a move commonly used in solving simultaneous equations: make one variable (x, say) the subject of an equation, then substitute it into the other equations in order to eliminate x. But in fact, this is really the same as the move we used repeatedly above, where we eliminated a variable by subtracting suitable multiples of one equation from all the others. It only looks different because we are insisting on keeping all the variables on the left-hand side and all the constants on the right.

Example C4.4 Consider a different 3×3 linear system:

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

Subtract 3 times the first equation from the second, and 4 times the first equation from the third, to get:

$$\begin{aligned} x + 2y + 3z &= 1 \\ 0x - y - 11z &= -1 \\ 0x - y - 11z &= -1. \end{aligned}$$

To eliminate y, first multiply the second equation by -1 so that y has a coefficient of 1:

$$\begin{aligned} x + 2y + 3z &= 1 \\ 0x + y + 11z &= 1 \\ 0x - y - 11z &= -1. \end{aligned}$$

Then subtract 2 times the second equation from the first and add 1 times the second equation to the third:

$$\begin{aligned} x + 0y - 19z &= -1 \\ 0x + y + 11z &= 1 \\ 0x + 0y + 0z &= 0. \end{aligned}$$

The last equation tells us nothing and can therefore be ignored. In the first two, we can choose z freely, say by putting $z = t$ for an arbitrary scalar t. Thus, the solutions of the system are:

$$x = -1 + 19t, \quad y = 1 - 11t, \quad z = t \quad (t \in \mathbb{R}).$$

We call z a **free variable** (since we can choose it freely) and x and y **leading variables**. The method for solving linear systems illustrated in the last two examples is called **Gaussian elimination**.

It is useful to consider the special kind of linear system where all the constants on the right-hand side are 0:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ \vdots \qquad\qquad\qquad & \;\;\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0. \end{aligned} \tag{C:3}$$

A linear system with this property is called **homogeneous**. Otherwise, it is called **inhomogeneous**.

We observed on page 71 that any linear system (C:1) can be expressed in matrix notation:

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

So in matrix notation, a homogeneous linear system is an equation of the form

$$A\mathbf{x} = \mathbf{0}$$

where A is a given $m \times n$ matrix, $\mathbf{x} \in \mathbb{R}^n$ is an unknown vector, and $\mathbf{0}$ is the zero vector in $\mathbb{R}^m$. The set of solutions of this homogeneous linear system is exactly the kernel of A (immediately from the definition of kernel). In other words:

> *The kernel of A is the set of solutions $\mathbf{x}$ of the homogeneous linear system $A\mathbf{x} = \mathbf{0}$.*

This is one way of understanding kernels. By Lemma B1.4, then, the solution-set of a homogeneous linear system in n variables is always a linear subspace of $\mathbb{R}^n$.

We saw in Example C4.1 that a linear system need not have any solutions at all. But a homogeneous linear system always has at least one, the **trivial solution** $\mathbf{x} = \mathbf{0}$.

Now consider an arbitrary linear system $A\mathbf{x} = \mathbf{b}$ (not necessarily homogeneous). I claim that if we can find just one solution of the system, then all other solutions can be obtained by adding on any solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$. That is:

Lemma C4.5 *Consider a linear system $A\mathbf{x} = \mathbf{b}$, where A is an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$. Let $\mathbf{x} = \mathbf{u}$ be a solution. Then the set of all solutions is*

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{u} + \mathbf{w} \textit{ for some } \mathbf{w} \in \ker(A)\}.$$

Proof We have to prove two things: that every element of this set is a solution, and that every solution belongs to this set.

First, let $\mathbf{w} \in \ker(A)$; we must prove that $\mathbf{x} = \mathbf{u} + \mathbf{w}$ satisfies $A\mathbf{x} = \mathbf{b}$. And indeed,

$$A(\mathbf{u} + \mathbf{w}) = A\mathbf{u} + A\mathbf{w} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Second, let $\mathbf{x} \in \mathbb{R}^n$ be a solution of $A\mathbf{x} = \mathbf{b}$; we must prove that $\mathbf{x} = \mathbf{u} + \mathbf{w}$ for some $\mathbf{w} \in \ker(A)$. Put $\mathbf{w} = \mathbf{x} - \mathbf{u}$. Then $\mathbf{x} = \mathbf{u} + \mathbf{w}$, and

$$A\mathbf{w} = A(\mathbf{x} - \mathbf{u}) = A\mathbf{x} - A\mathbf{u} = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

so $\mathbf{w} \in \ker(A)$. □

Example C4.6 (See Figure C.2.) Suppose that our system consists of a single equation in two variables,

$$2x - 3y = 7.$$

(In the notation above: $m = 1$, $n = 2$, $A = \begin{pmatrix} 2 & -3 \end{pmatrix}$, and $\mathbf{b} = (7)$.) The associated homogeneous system is

$$2x - 3y = 0.$$

One solution of the original system is $x = 5$, $y = 1$. The general solution of the homogeneous system is $\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} 3t \\ 2t \end{smallmatrix}\right)$ $(t \in \mathbb{R})$. So by Lemma C4.5, the general solution of the original system is

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix} + \begin{pmatrix} 3t \\ 2t \end{pmatrix} \qquad (t \in \mathbb{R}).$$

Just before Lemma C4.5, we saw that the solution-set of a *homogeneous* linear system in n variables is always a linear subspace of $\mathbb{R}^n$. The solution-set S of an inhomogeneous linear system is never a subspace, since $\mathbf{x} = \mathbf{0}$ is not a solution. However, Lemma C4.5 tells us that S is a subspace translated by some vector (called $\mathbf{u}$ in the statement of the lemma).

For instance, in Example C4.6, the set S of solutions of $2x - 3y = 7$ is the line $2x - 3y = 0$ translated by the vector $\left(\begin{smallmatrix} 5 \\ 1 \end{smallmatrix}\right)$. The solution-sets of the homogeneous and inhomogeneous systems are both lines, but only for the homogeneous system is it a line through the origin.

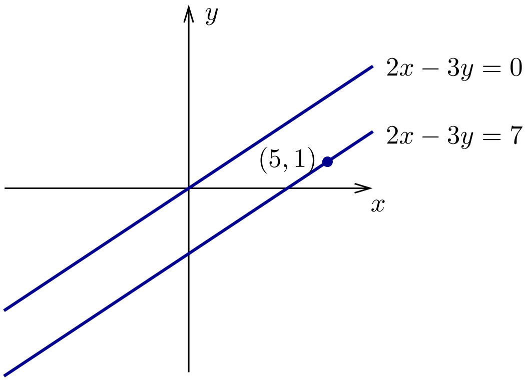


Figure C.2: The solution-sets of an inhomogeneous linear system and its associated homogeneous linear system (Example C4.6)

Remark C4.7 (Non-examinable.) A very similar result holds in the theory of linear differential equations. Consider, for instance, the differential equation

$$2f'(\theta) - 3f(\theta) = \cos\theta$$

(where f is an unknown function and the equation is to hold for all θ). Suppose we have found one solution, say $f = u$. Then the other solutions are the functions of the form $u+w$ where w is a solution of the *homogeneous* differential equation

$$2w'(\theta) - 3w(\theta) = 0.$$

It's not hard to show this directly; the proof is very similar to that of Lemma C4.5. (Maybe you've already shown this in SVCDE. In traditional terminology, u is called a 'particular integral' and w is called the/a 'complementary function'.) A crucial fact here is that differentiation is linear, meaning that $(f + g)' = f' + g'$ and $(cf)' = cf'$ for any scalar c.

In Example C4.6, the homogeneous system had infinitely many solutions. There are other examples of homogeneous systems that have only one solution, namely, the trivial solution $\mathbf{0}$. (Can you think of an example?) In fact, these are the only possibilities:

Lemma C4.8 *A homogeneous linear system has either just one solution (the trivial solution* $\mathbf{0}$*) or infinitely many solutions.*

Proof The set of solutions of a homogeneous system $A\mathbf{x} = \mathbf{0}$ is $\ker(A)$, which is a linear subspace of $\mathbb{R}^n$. So if the system has a nontrivial solution $\mathbf{x} \neq \mathbf{0}$ then $c\mathbf{x}$ is a solution for all $c \in \mathbb{R}$. But the vectors $c\mathbf{x}$ are different for different values of $c \in \mathbb{R}$ (since $\mathbf{x} \neq \mathbf{0}$), so there are infinitely many solutions. □

We saw in Example C4.1 that some inhomogeneous linear systems have no solutions, some have exactly one solution, and some have infinitely many solutions. Again, these are the only possibilities:

Lemma C4.9 *A linear system has no solutions, exactly one solution, or infinitely many solutions.*

Proof Consider a linear system $A\mathbf{x} = \mathbf{b}$. If it has any solutions at all, then by Lemma C4.5, the set of all solutions has exactly as many elements as the set of solutions of the homogeneous system $A\mathbf{x} = \mathbf{0}$. But by Lemma C4.8, this set has either one element or infinitely many. □

So if, for instance, you have found two different solutions of a linear system, you can immediately deduce that it has infinitely many solutions.

C5 How to solve a linear system

In Examples C4.2 and C4.4, we solved some linear systems in a more or less methodical way. We used three operations repeatedly:

- interchange two equations;
- multiply an equation by a nonzero scalar (on both sides);
- add a multiple of one equation to another equation.

We've seen that a linear system can be expressed most efficiently in matrix form. So, let's now translate these operations into matrix terms.

When dealing with a linear system $A\mathbf{x} = \mathbf{b}$, it is often convenient to write the $m \times n$ matrix A next to the m-dimensional vector $\mathbf{b}$, making a single matrix with a vertical bar separating A from $\mathbf{b}$:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right).$$

Here $A = (a_{ij})$.

This is called the **augmented matrix** of the linear system. The equations of the linear system correspond to the rows of the augmented matrix, with the bar separating the left-hand side from the right-hand side. The three operations on equations correspond to the following three operations on a matrix, which are called the **elementary row operations**:

- interchange two rows;
- multiply a row by a nonzero scalar;
- add a scalar multiple of one row to another row.

Since we will be using these operations repeatedly, we set up some notation for them. Interchanging rows i and j is written as $R_i \leftrightarrow R_j$, multiplying row i by a nonzero scalar c is written as $R_i \to cR_i$, and adding c times row j to row i is written as $R_i \to R_i + cR_j$.

Example C5.1 The augmented matrix of the linear system in Example C4.2 is

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

When we solved this system in Example C4.2, we did it explicitly in terms of equations and variables. Here it is again: exactly the same argument as in Example C4.2, but this time in terms of elementary row operations.

$$
\begin{aligned}
\left(\begin{array}{ccc|c} 2 & 3 & -1 & 3 \\ 1 & 1 & 1 & 4 \\ 3 & -4 & 1 & 1 \end{array}\right)
&\xrightarrow[R_1\leftrightarrow R_2]{\hspace{8em}}
\left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 2 & 3 & -1 & 3 \\ 3 & -4 & 1 & 1 \end{array}\right) \\
&\xrightarrow[\substack{R_2\to R_2-2R_1\\ R_3\to R_3-3R_1}]{\hspace{8em}}
\left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 1 & -3 & -5 \\ 0 & -7 & -2 & -11 \end{array}\right) \\
&\xrightarrow[\substack{R_1\to R_1-R_2\\ R_3\to R_3+7R_2}]{\hspace{8em}}
\left(\begin{array}{ccc|c} 1 & 0 & 4 & 9 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & -23 & -46 \end{array}\right) \\
&\xrightarrow[R_3\to(-1/23)R_3]{\hspace{8em}}
\left(\begin{array}{ccc|c} 1 & 0 & 4 & 9 \\ 0 & 1 & -3 & -5 \\ 0 & 0 & 1 & 2 \end{array}\right) \\
&\xrightarrow[\substack{R_1\to R_1-4R_3\\ R_2\to R_2+3R_3}]{\hspace{8em}}
\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right).
\end{aligned}
$$

(Notice how we cleared the columns one by one, working from left to right.) We conclude that the unique solution of the linear system is

$$
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.
$$

Warning C5.2 It's *sometimes* OK to do multiple row operations at once, as shown in the last example. But in any one step, don't do more than one operation *to any individual row*, as that's where mistakes can arise.

Warning C5.3 When you're doing row reductions, *always say which row you're operating on.* I say this because Poole uses a form of notation that I discourage you from using; for instance, he writes '$R_1 \to R_1 + R_2$' as just '$R_1 + R_2$'. So, for him, '$R_1 + R_2$' and '$R_2 + R_1$' mean different things! *Don't do this.* Stick to the notation above.

How do we know this method gives the correct set of solutions? It is the same method that we used as in Examples C4.2 and C4.4, just in different notation. Back then, we used the principle that changing the order of the equations, or multiplying an equation by a nonzero constant, or adding a multiple of one equation to another, doesn't change the solution-set. The first two are clear, but the last perhaps less so, and in any case we should state the principle formally:

Lemma C5.4 *Let A be an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$. Let A' be the matrix obtained from A by performing a single row operation, and let $\mathbf{b}'$ be the vector obtained from $\mathbf{b}$ by performing the same row operation. Then for $\mathbf{x} \in \mathbb{R}^n$,*

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

Proof There are three cases to prove, corresponding to the three types of elementary row operation. I will do the third type only; the first two are similar, easier, and left to you as an exercise.

So, suppose that A' is obtained from A by the row operation $R_i \to R_i + cR_j$ (where $i, j \in \{1, \ldots, m\}$ and $c \in \mathbb{R}$), and similarly $\mathbf{b}'$ from $\mathbf{b}$. Assume without loss of generality that $i = 1$ and $j = 2$. Write the rows of A as $\mathbf{r}_1, \ldots, \mathbf{r}_m$.

Let $\mathbf{x} \in \mathbb{R}^n$. By definition of matrix multiplication,

$$A\mathbf{x} = \begin{pmatrix} \mathbf{r}_1\mathbf{x} \\ \mathbf{r}_2\mathbf{x} \\ \vdots \\ \mathbf{r}_m\mathbf{x} \end{pmatrix}.$$

(Each row $\mathbf{r}_p$ is a $1 \times n$ matrix, and $\mathbf{x}$ is an $n \times 1$ matrix, so $\mathbf{r}_p\mathbf{x}$ is a 1×1 matrix, that is, a scalar.) The rows of A' are $\mathbf{r}_1 + c\mathbf{r}_2, \mathbf{r}_2, \ldots, \mathbf{r}_m$, so

$$A'\mathbf{x} = \begin{pmatrix} (\mathbf{r}_1 + c\mathbf{r}_2)\mathbf{x} \\ \mathbf{r}_2\mathbf{x} \\ \vdots \\ \mathbf{r}_m\mathbf{x} \end{pmatrix} = \begin{pmatrix} \mathbf{r}_1\mathbf{x} + c\mathbf{r}_2\mathbf{x} \\ \mathbf{r}_2\mathbf{x} \\ \vdots \\ \mathbf{r}_m\mathbf{x} \end{pmatrix}.$$

Hence

$$\begin{aligned} A'\mathbf{x} = \mathbf{b}' &\iff \begin{pmatrix} \mathbf{r}_1\mathbf{x} + c\mathbf{r}_2\mathbf{x} \\ \mathbf{r}_2\mathbf{x} \\ \vdots \\ \mathbf{r}_m\mathbf{x} \end{pmatrix} = \begin{pmatrix} b_1 + cb_2 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \\ &\iff \begin{pmatrix} \mathbf{r}_1\mathbf{x} \\ \mathbf{r}_2\mathbf{x} \\ \vdots \\ \mathbf{r}_m\mathbf{x} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \iff A\mathbf{x} = \mathbf{b} \end{aligned}$$

where the second '$\iff$' follows from the observation that

$$\big(\mathbf{r}_1\mathbf{x} + c\mathbf{r}_2\mathbf{x} = b_1 + cb_2 \text{ and } \mathbf{r}_2\mathbf{x} = b_2\big) \iff \big(\mathbf{r}_1\mathbf{x} = b_1 \text{ and } \mathbf{r}_2\mathbf{x} = b_2\big).$$

This completes the proof. □

Now that we have shown our method to be valid, here is another example.

Example C5.5 The augmented matrix of the linear system in Example C4.4 is

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

The reductions performed in Example C4.4 can be translated into row operations. (Exercise: try it!) The end result is

$$\left(\begin{array}{ccc|c} 1 & 0 & -19 & -1 \\ 0 & 1 & 11 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

As in Example C4.4, we conclude that the set of solutions is

$$\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 19 \\ -11 \\ 1 \end{pmatrix} : t \in \mathbb{R} \right\}.$$

The matrices we ended up with in both of the last two examples were of a certain special type. Here is the terminology.

Definition C5.6 A matrix is in **row echelon form (REF)** if:

i. any rows consisting entirely of zeros are at the bottom; and

ii. in each nonzero row, the first nonzero entry (called the **leading entry**) is to the left of all the leading entries below it.

Example C5.7 The matrix

$$\begin{pmatrix} \underline{1} & 2 & 3 & 4 & 5 \\ 0 & \underline{6} & 7 & 8 & 9 \\ 0 & 0 & 0 & \underline{10} & 11 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

is in row echelon form. The leading entries are underlined.

The word *echelon* comes from the Latin word for staircase, which also gives rise to modern English words such as *escalator*.

Definition C5.8 A matrix is in **reduced row echelon form (RREF)** if it is in row echelon form, and:

i. all leading entries are equal to 1; and

ii. each column containing a leading 1 has zeros everywhere else.

Examples C5.9 The final matrices in Examples C5.1 and C5.5 are both in reduced row echelon form, and so is

$$\begin{pmatrix} 1 & 0 & 3 & 0 \\ 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

It is a fact that by repeatedly performing elementary row operations, any matrix can be put into reduced row echelon form. It is also a fact that the RREF is unique. In other words, if you give the same matrix M to two different people and ask them to put it into RREF, they might use different row operations to get there, but they are guaranteed to arrive at the same final answer.

In summary, a systematic way to solve a linear system is to:

- write down the augmented matrix;
- use elementary row operations to put it into reduced row echelon form;
- read off the solutions.

As in Lemma C4.9, there may be no solutions, exactly one solution, or infinitely many solutions.

Remark C5.10 What's the point of non-reduced row echelon form? The answer is practical: it can be a useful intermediate stage. For example, suppose we have used elementary row operations to reduce a system to

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

This is in REF. It is not in RREF, since it fails both conditions in Definition C5.8. But still, we can use it to write down the solution(s) quickly, as follows. This matrix corresponds to the linear system

$$\begin{aligned} x + y + 4z &= 6 \\ y - 5z &= -3 \\ 2z &= 8. \end{aligned}$$

The third equation gives $z = 4$. Substituting this into the second equation gives $y = 17$. Then substituting these into the first equation gives $x = -27$. So this is the unique solution.

C6 How to invert a matrix

. . . or more accurately, 'How to tell whether a matrix is invertible, and how to invert it if it is'.

In Section C3, we found a determinant-based method for telling whether a matrix is invertible and inverting it if it is: a square matrix A is invertible if and only if $\det(A) \neq 0$, and in that case, $A^{-1} = \operatorname{adj}(A)/\det(A)$. However, this method is disastrously inefficient. Calculating a 4×4 determinant is already long and tedious for a human being. Computers are faster, but even so, calculating the determinant of a 100×100 matrix using the definition in Section C3 needs about 10^{158} operations, which even at a trillion trillion trillion operations per second would take more than

1 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000 000

years. And in mathematical models of real-world situations, a 100×100 matrix isn't especially big—systems of thousands of equations in thousands of variables are routine.

(There are faster methods of computing determinants, but still, determinants are hardly ever used in real-life numerical computing.)

In this section, we will use elementary row operations to derive a much more efficient method. We begin with the question of *whether* a matrix is invertible.

Lemma C6.1 *Let A be a matrix and let A' be another matrix obtained from A by a sequence of elementary row operations. Then* $\ker(A) = \ker(A')$.

Proof This follows by induction from Lemma C5.4, taking $\mathbf{b} = \mathbf{0}$ there. □

Lemma C6.2 *The only $n \times n$ invertible matrix in reduced row echelon form is I_n.*

Proof Let R be an $n \times n$ invertible matrix in REF. Since R is invertible, Theorem C2.4 implies that the rows of R are linearly independent, so by Example B3.2(vi), R has no rows consisting entirely of zeros. Since R is in REF, the leading entry of each of the n rows is to the left of the leading entries below it, and since R has only n columns, the leading entries must all be on the main diagonal. Hence R is of the form

$$\begin{pmatrix} * & * & \cdots & * \\ 0 & * & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & * \end{pmatrix}$$

where $*$ denotes any scalar and the diagonal entries R_{ii} are all nonzero.

Now assume that R is in *reduced* REF. Every leading entry is 1, and every column containing a leading entry has zeros everywhere else, so $R = I_n$. □

Theorem C6.3 (Equivalent conditions for invertibility, part 4) *Let A be an $n \times n$ matrix. Then A is invertible if and only if the reduced row echelon form of A is I_n.*

Proof Write R for the RREF of A. By Theorem C2.3, A is invertible if and only if $\ker(A) = \{\mathbf{0}\}$, and R is invertible if and only if $\ker(R) = \{\mathbf{0}\}$. But $\ker(A) = \ker(R)$ by Lemma C6.1, so A is invertible if and only if R is. Lemma C6.2 now completes the proof. □

This gives an efficient practical method for deciding whether a matrix is invertible: use elementary row operations to compute the RREF, then check to see whether the RREF is the identity. If so, the original matrix is invertible; if not, it isn't. We will see an example later, once we have answered the next question: if a matrix is invertible, how do we actually compute the inverse?

Let $A = (a_{ij})$ be an invertible $n \times n$ matrix. By Theorem C2.3(ix), for any vector $\mathbf{b} \in \mathbb{R}^n$, the inhomogeneous linear system $A\mathbf{x} = \mathbf{b}$ has exactly one solution, namely, $\mathbf{x} = A^{-1}\mathbf{b}$. So we can find $A^{-1}\mathbf{b}$ by solving this system. We know how to solve *any* linear system $A\mathbf{x} = \mathbf{b}$, but the fact that A is invertible makes things especially simple: the RREF is I, so the unique solution $\mathbf{x}$ is simply whatever is to the right of the vertical bar at the end of the reduction process. For instance, in Example C5.1, the final matrix in the reduction process was

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

and the unique solution was $\mathbf{x} = \left(\begin{smallmatrix} 1 \\ 1 \\ 2 \end{smallmatrix}\right)$.

These observations give us a way of computing A^{-1}, as follows. By Lemma A4.3(ii), the first column of A^{-1} is $A^{-1}\mathbf{e}_1$. In other words, it is the unique solution $\mathbf{x}_1$ of the equation $A\mathbf{x}_1 = \mathbf{e}_1$. And we have just seen how to compute this! We write down the augmented matrix

$$\left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & 1 \\ a_{21} & \cdots & a_{2n} & 0 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 \end{array}\right),$$

we do elementary row operations until the part to the left of the bar is in RREF (which will be I_n, since A is invertible), and then we read off the vector to the right of the bar. This will be the first column of A^{-1}.

Of course, the same thing works for all the other columns too. And since the same process is involved every time, we might as well do all the columns at once, putting all of $\mathbf{e}_1, \ldots, \mathbf{e}_n$ to the right of the bar. This gives an algorithm for computing the inverse of an invertible matrix, as well as (simultaneously) an algorithm for deciding *whether* a matrix is invertible. It is best illustrated by some examples.

Example C6.4 Is the matrix

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

invertible? If so, what is its inverse?

Place the vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ (that is, the identity matrix) next to A:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 4 & 0 & 5 & 0 & 0 & 1 \end{array}\right).$$

Using elementary row operations on the whole 3×6 matrix, put the left-hand half into reduced row echelon form:

$$\begin{aligned} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 4 & 0 & 5 & 0 & 0 & 1 \end{array}\right) &\xrightarrow[R_3 \to R_3 - 4R_1]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 \\ 0 & 0 & -3 & -4 & 0 & 1 \end{array}\right) \\ &\xrightarrow[\substack{R_2 \to (1/3)R_2 \\ R_3 \to (-1/3)R_3}]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 1 & 4/3 & 0 & -1/3 \end{array}\right) \\ &\xrightarrow[R_1 \to R_1 - 2R_3]{} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -5/3 & 0 & 2/3 \\ 0 & 1 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 1 & 4/3 & 0 & -1/3 \end{array}\right). \end{aligned}$$

The RREF of A is I, so by Theorem C6.3, A is invertible. Moreover, the argument just given shows that A^{-1} is the right-hand half; that is,

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

Example C6.5 Is the matrix

$$A = \begin{pmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}$$

invertible? If so, what is its inverse?

Apply the same method:

$$\left(\begin{array}{ccc|ccc} 9 & 8 & 7 & 1 & 0 & 0 \\ 6 & 5 & 4 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \xrightarrow[R_1 \to (1/9)R_1]{} \left(\begin{array}{ccc|ccc} 1 & 8/9 & 7/9 & 1/9 & 0 & 0 \\ 6 & 5 & 4 & 0 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[\substack{R_2 \to R_2 - 6R_1 \\ R_3 \to R_3 - 3R_1}]{} \left(\begin{array}{ccc|ccc} 1 & 8/9 & 7/9 & 1/9 & 0 & 0 \\ 0 & -1/3 & -2/3 & -2/3 & 1 & 0 \\ 0 & -2/3 & -4/3 & -1/3 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[R_3 \to R_3 - 2R_2]{} \left(\begin{array}{ccc|ccc} 1 & 8/9 & 7/9 & 1/9 & 0 & 0 \\ 0 & -1/3 & -2/3 & -2/3 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{array}\right).$$

We haven't yet reduced the left-hand half to RREF, but it already has a zero row, so the RREF will have a zero row too. Hence A is not invertible, by Theorem C6.3. (The right-hand half plays no part this time; it only would have been useful if A had turned out to be invertible.)

C7 How to calculate everything else

If I gave you a list of vectors in $\mathbb{R}^n$, how could you determine whether they were linearly independent? How could you find a basis of the subspace that they span? Or if I gave you a matrix, how could you calculate a basis of its row space? What about its column space, or its kernel? And how could you compute its rank and nullity? Or, suppose I gave you a subspace of $\mathbb{R}^n$. How could you find an orthonormal basis of it? How could you find a basis of its orthogonal complement?

In this section, we will see how to answer all these questions. To do so, we need the following three useful results.

Lemma C7.1 *Let A be a matrix and let A' be another matrix obtained from A by a sequence of elementary row operations. Then* $\mathrm{row}(A) = \mathrm{row}(A')$ *and* $\mathrm{rank}(A) = \mathrm{rank}(A')$.

Proof The first equation follows from Lemmas C1.5 and C6.1. The second follows from the first by taking the dimension of each side. □

Warning C7.2 The matrices A and A' need not have the same *column* space. For instance, $A = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ can be transformed by an elementary row operation into $A' = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$, but $\mathrm{col}(A) = \mathrm{span}\{\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\} \neq \mathrm{span}\{\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)\} = \mathrm{col}(A')$. There is also a notion of elementary *column* operation, and a matrix has the same column space as anything obtained from it by elementary column operations.

Lemma C7.3 *Let R be a matrix in row echelon form. Then the nonzero rows of R are a basis of* $\mathrm{row}(R)$.

(Strictly speaking, it is actually the *transposes* of the nonzero rows that are a basis of $\mathrm{row}(R)$; see the definition on page 50. But never mind!)

Proof By definition, the rows of R span $\mathrm{row}(R)$. Omitting the zero rows does not alter the span, so the nonzero rows of R also span $\mathrm{row}(R)$. It remains to show that the nonzero rows are linearly independent.

Suppose that the matrix R is $m \times n$, with k nonzero rows. By definition of REF, they are the first k rows. For $i \in \{1, \dots, k\}$, write the ith row as $\mathbf{r}_i \in \mathbb{R}^n$, write the leading entry in the ith row as ℓ_i, and let us suppose that this leading entry is in the p_ith column. By definition of REF,

$$1 \leq p_1 < p_2 < \cdots < p_k \leq n.$$

Now let $c_1, \dots, c_k$ be scalars such that

$$c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + \cdots + c_k\mathbf{r}_k = \mathbf{0}. \tag{C:4}$$

We must prove that $c_1 = \cdots = c_k = 0$.

Apart from the first row, there are no nonzero entries in the p_1th column. So, comparing the p_1th entries on each side of equation (C:4) gives $c_1\ell_1 = 0$. But ℓ_1 is a leading entry, so is not zero, so $c_1 = 0$. Hence (C:4) now gives

$$c_2\mathbf{r}_2 + \cdots + c_k\mathbf{r}_k = \mathbf{0}.$$

If we delete the first row from A then what remains is still in REF, so we can repeat the same argument to get $c_2 = 0$. Continuing in this way gives $c_i = 0$ for all $i \in \{1, 2, \dots, k\}$, as required. □

Proposition C7.4 *The rank of a matrix is equal to the number of nonzero rows in any reduced echelon form.*

Proof This follows from Lemmas C7.1 and C7.3. □

These results tell us that a lot of information about a matrix can be read off from any row echelon form. This is illustrated in the following examples.

Example C7.5 Given a matrix, how can we calculate a basis of its row space?

Take, for instance,

$$A = \begin{pmatrix} 1 & 1 & 3 & 1 & 6 \\ 2 & -1 & 0 & 1 & -1 \\ -3 & 2 & 1 & -2 & 1 \\ 4 & 1 & 6 & 1 & 3 \end{pmatrix},$$

After a sequence of elementary row operations, we find that the RREF of A is

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

By Lemma C7.3, the nonzero rows of R (or really their transposes) form a basis of $\mathrm{row}(R)$, which by Lemma C7.1 is equal to $\mathrm{row}(A)$. So

$$\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 4 \end{pmatrix} \tag{C:5}$$

form a basis of row(A).

Here we calculated the *reduced* row echelon form of A, but any other row echelon form would also have worked, for the same reasons.

Example C7.6 Given a list of vectors, how can we calculate a basis of their span? For instance, take

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 3 \\ 1 \\ 6 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -3 \\ 2 \\ 1 \\ -2 \\ 1 \end{pmatrix}, \quad \mathbf{v}_4 = \begin{pmatrix} 4 \\ 1 \\ 6 \\ 1 \\ 3 \end{pmatrix}.$$

Turn these into row vectors (i.e. take transposes) and put them together as the rows of a matrix. In this example, this gives the matrix A of Example C7.5. Then $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\} = \text{row}(A)$, and we have already found a basis of row(A): the three vectors in (C:5). These, then, are a basis of $\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$.

Example C7.7 Similarly, to calculate a basis of the column space of a matrix, we calculate a basis of the row space of its transpose. For instance, one basis of the column space of

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

is the list of three vectors in (C:5), for exactly the reasons given in the last example.

Example C7.8 Given a matrix, how can we calculate its rank and nullity?

By Proposition C7.4, we can calculate the rank by computing a row echelon form and counting its nonzero rows. So for the matrix A of Example C7.5, we have $\text{rank}(A) = 3$. Then by the rank-nullity theorem,

$$\text{nullity}(A) = (\text{number of columns of } A) - \text{rank}(A) = 5 - 3 = 2.$$

Example C7.9 Given a list of vectors, how can we determine whether they are linearly independent?

Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m \in \mathbb{R}^n$ are linearly independent if and only if $\dim(\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}) = m$ (Proposition B5.8). But $\dim(\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_m\})$ is the rank of the matrix A with rows $\mathbf{v}_1^T, \ldots, \mathbf{v}_m^T$, and we already know how to compute the rank of a matrix. Since A is an $m \times n$ matrix, it has rank m if and only if a row echelon form of A has no zero rows.

Thus, to determine whether $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent:

- form the matrix A with rows $\mathbf{v}_1^T, \ldots, \mathbf{v}_m^T$;
- compute a row echelon form R of A;
- if R has a zero row then $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly dependent; otherwise, they are independent.

For $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4 \in \mathbb{R}^5$ as in Example C7.6, the last row of the reduced row echelon form is zero, so they are linearly dependent.

Example C7.10 Given a matrix A, how can we calculate a basis of $\ker(A)$?

By definition, $\ker(A)$ is the set of solutions $\mathbf{x}$ of the homogeneous linear system $A\mathbf{x} = \mathbf{0}$. We already know how to solve linear systems: put A into reduced row echelon form, then just read off the solutions. For instance, take the matrix A of Example C7.5, whose RREF is

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

So the solutions of $A\mathbf{x} = \mathbf{0}$ are the vectors $\mathbf{x}$ satisfying

$$\begin{aligned} x_1 + x_3 - x_5 &= 0, \\ x_2 + 2x_3 + 3x_5 &= 0, \\ x_4 + 4x_5 &= 0. \end{aligned}$$

The leading entries of the RREF are in columns 1, 2 and 4. Correspondingly, each of these equations contains exactly one of x_1, x_2 and x_4 (each with a coefficient of 1) together with some of the remaining variables. Moving the remaining variables to the right-hand sides, we find that $\mathbf{x}$ is a solution if and only if

$$\begin{aligned} x_1 &= -x_3 + x_5, \\ x_2 &= -2x_3 - 3x_5, \\ x_4 &= -4x_5. \end{aligned}$$

So, writing $s = x_3$ and $t = x_5$,

$$\ker(A) = \left\{ \begin{pmatrix} -s+t \\ -2s-3t \\ s \\ -4t \\ t \end{pmatrix} : s, t \in \mathbb{R} \right\} = \left\{ s \begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{pmatrix} : s, t \in \mathbb{R} \right\}.$$

(In the terminology introduced in Example C4.4, x_3 and x_5 are 'free' variables.) This has a basis

$$\begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{pmatrix}.$$

This confirms our calculation in Example C7.8 that $\ker(A)$ has dimension 2.

Example C7.11 Given a subspace V of $\mathbb{R}^n$, presented as the span of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_m$, how can we find a basis of its orthogonal complement $V^\perp$?

Once more, let A be the $m \times n$ matrix with rows $\mathbf{v}_1^T, \ldots, \mathbf{v}_m^T$. Then $V = \operatorname{row}(A)$, so $V^\perp = \operatorname{row}(A)^\perp = \ker(A)$ by Lemma C1.5. Finding a basis of $V^\perp$ is, therefore, the same as finding a basis of $\ker(A)$, which we just saw how to do in Example C7.10.

The one remaining question is how to find an orthonormal basis of a given subspace of $\mathbb{R}^n$. We showed in Proposition B7.7 that every subspace does have at least one orthonormal basis. If we work carefully through the proofs of that proposition and the results it depends on (especially the proof of Lemma B7.6), we arrive at the following algorithm for actually *constructing* an orthonormal basis.

Let V be a linear subspace of $\mathbb{R}^n$, with basis $\mathbf{y}_1, \dots, \mathbf{y}_m$. We can construct an *orthonormal* basis $\mathbf{v}_1, \dots, \mathbf{v}_m$ of V by putting

$$\begin{aligned}
\mathbf{v}_1 &= \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|}, \\
\mathbf{v}_2 &= \frac{\mathbf{y}_2 - (\mathbf{y}_2 \cdot \mathbf{v}_1)\mathbf{v}_1}{\|\mathbf{y}_2 - (\mathbf{y}_2 \cdot \mathbf{v}_1)\mathbf{v}_1\|}, \\
\mathbf{v}_3 &= \frac{\mathbf{y}_3 - \big[(\mathbf{y}_3 \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{y}_3 \cdot \mathbf{v}_2)\mathbf{v}_2\big]}{\big\|\mathbf{y}_3 - \big[(\mathbf{y}_3 \cdot \mathbf{v}_1)\mathbf{v}_1 + (\mathbf{y}_3 \cdot \mathbf{v}_2)\mathbf{v}_2\big]\big\|}, \\
&\vdots \\
\mathbf{v}_k &= \frac{\mathbf{y}_k - \sum_{i=1}^{k-1} (\mathbf{y}_k \cdot \mathbf{v}_i)\mathbf{v}_i}{\Big\|\mathbf{y}_k - \sum_{i=1}^{k-1} (\mathbf{y}_k \cdot \mathbf{v}_i)\mathbf{v}_i\Big\|} \\
&\vdots
\end{aligned}$$

This procedure is called the **Gram–Schmidt process**.

Example C7.12 Let $V = \operatorname{span}\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3\}$, where

$$\mathbf{y}_1 = \begin{pmatrix} -5 \\ 3 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{y}_2 = \begin{pmatrix} 1 \\ 0 \\ -4 \\ 3 \end{pmatrix}, \quad \mathbf{y}_3 = \begin{pmatrix} 1 \\ -3 \\ 2 \\ 0 \end{pmatrix}.$$

How can we find an orthonormal basis of V?

First check that $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ are linearly independent, as in Example C7.9. It follows that $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ form a basis of V. So, we can now apply the Gram–Schmidt process to obtain an orthonormal basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ of V.

To calculate $\mathbf{v}_1$:

$$\mathbf{v}_1 = \frac{\mathbf{y}_1}{\|\mathbf{y}_1\|} = \frac{1}{6} \begin{pmatrix} -5 \\ 3 \\ 1 \\ 1 \end{pmatrix}.$$

To calculate $\mathbf{v}_2$: we have

$$\mathbf{y}_2 - (\mathbf{y}_2 \cdot \mathbf{v}_1)\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -4 \\ 3 \end{pmatrix} - \left(\frac{1}{6} \times -6\right) \frac{1}{6} \begin{pmatrix} -5 \\ 3 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 \\ 3 \\ -23 \\ 19 \end{pmatrix}$$

and so

$$\mathbf{v}_2 = \frac{\mathbf{y}_2 - (\mathbf{y}_2 \cdot \mathbf{v}_1)\mathbf{v}_1}{\|\mathbf{y}_2 - (\mathbf{y}_2 \cdot \mathbf{v}_1)\mathbf{v}_1\|} = \frac{1}{30} \begin{pmatrix} 1 \\ 3 \\ -23 \\ 19 \end{pmatrix}.$$

To calculate $\mathbf{v}_3$: we have

$$\begin{aligned}\mathbf{y}_3 - \big[(\mathbf{y}_3\cdot\mathbf{v}_1)\mathbf{v}_1 + (\mathbf{y}_3\cdot\mathbf{v}_2)\mathbf{v}_2\big] &= \begin{pmatrix}1\\-3\\2\\0\end{pmatrix} - \left[\Big(\frac{1}{6}\times -12\Big)\frac{1}{6}\begin{pmatrix}-5\\3\\1\\1\end{pmatrix} + \Big(\frac{1}{30}\times -54\Big)\frac{1}{30}\begin{pmatrix}1\\3\\-23\\19\end{pmatrix}\right]\\ &= \frac{1}{150}\begin{pmatrix}-91\\-273\\143\\221\end{pmatrix}\end{aligned}$$

and so

$$\mathbf{v}_3 = \frac{\mathbf{y}_3 - \big[(\mathbf{y}_3\cdot\mathbf{v}_1)\mathbf{v}_1 + (\mathbf{y}_3\cdot\mathbf{v}_2)\mathbf{v}_2\big]}{\big\|\mathbf{y}_3 - \big[(\mathbf{y}_3\cdot\mathbf{v}_1)\mathbf{v}_1 + (\mathbf{y}_3\cdot\mathbf{v}_2)\mathbf{v}_2\big]\big\|} = \frac{1}{390}\begin{pmatrix}-91\\-273\\143\\221\end{pmatrix} = \frac{1}{30}\begin{pmatrix}-7\\-21\\11\\17\end{pmatrix}.$$

These vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are automatically an orthonormal basis of V. (You can also check directly that $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are orthonormal and belong to V, which by Corollary B6.5 implies that they are an orthonormal basis of V.)

As this example shows, the Gram–Schmidt process is rather labour-intensive—and if the numbers aren't as convenient as in this example, there are usually a lot of square roots involved too. But it is clearly a method that could be implemented on a computer, and indeed, most computer algebra packages do include an implementation.

C8 How many solutions does a linear system have?

We already saw an answer in Lemma C4.9: 0, 1, or ∞. But in a sense, this answer is not very satisfactory. After all, if the set of solutions of one linear system is a line and the set of solutions of another is a plane, then both have infinitely many solutions, but we would intuitively say that the second has more. If the line and plane are through the origin, we already have the language to make this precise: the first solution-set has dimension 1, and the second has dimension 2.

So that we can talk about the dimension of the solution-set, let us stick to *homogeneous* linear systems for now. In other words, let us consider equations of the form $A\mathbf{x} = \mathbf{0}$ where A is an $m\times n$ matrix. The set of solutions is $\ker(A) \subseteq \mathbb{R}^n$, and its dimension can be thought of as the 'number of independent solutions'. How many of these would we expect?

Very roughly, we expect lots of solutions when there are many variables and only a few equations, and few solutions when there are few variables constrained by many equations. For n equations in n variables, there is usually exactly one solution.

However, we have to be careful what we mean by the 'number of equations'.

For instance, if we have two equations

$$\begin{aligned} 7x - 8y + 5z &= 0 \\ 70x - 80y + 50z &= 0 \end{aligned}$$

then there might as well be only one equation, since the second is a scalar multiple of the first. Or, less obviously, if we have three equations

$$\begin{aligned} 2x + 3y + 4z &= 0 \\ 20x + 90y - 80z &= 0 \\ 14x + 51y - 32z &= 0 \end{aligned}$$

then there might as well be only two, since each of them is a linear combination of the other two. So, what we really want is the 'number of independent equations'. Each equation corresponds to a row of A, so this is the 'number of independent rows' of A. In precise terms, it is the dimension of the row space of A.

To summarize the argument so far: the intuitive idea of the 'number of independent solutions' is made precise as the dimension of the kernel of A, and the intuitive idea of the 'number of independent equations' is made precise as the dimension of the row space of A. In other words, they are the nullity and rank of A, respectively. And the number of variables in the system is simply n, which is the number of columns of A.

Now, the rank-nullity theorem tells us that

$$\operatorname{nullity}(A) = n - \operatorname{rank}(A).$$

So in informal terms, the rank-nullity theorem can be understood as saying that for a homogeneous linear system,

$$\begin{aligned} &\text{number of independent solutions} \\ &\qquad = \text{number of variables} \ - \ \text{number of independent equations}. \end{aligned}$$

I say 'informally' because phrases such as 'number of independent solutions' aren't precisely defined. (However, the only sensible interpretations are the ones given here.)

In particular, if there are more variables than equations then there are certainly more variables than *independent* equations, so there should be at least one nontrivial solution. And indeed, there is:

Proposition C8.1 *A homogeneous linear system with more variables than equations has at least one nontrivial solution.*

Such systems are said to be **underdetermined**.

Proof Write n for the number of variables and m for the number of equations; then $m < n$. Write A for the matrix of coefficients. Then A is an $m \times n$ matrix, so $\operatorname{rank}(A) \leq m$. Hence $\operatorname{rank}(A) < n$, and so $\operatorname{nullity}(A) > 0$ by the rank-nullity theorem. But $\operatorname{nullity}(A)$ is the dimension of the set of solutions of the system, so $\mathbf{0}$ is not the only solution. □

What about inhomogeneous systems? These need not have any solutions at all, even if they have more variables than equations. For instance, the system

$$\begin{aligned} 2x_1 + 3x_2 + 4x_3 &= 1 \\ 20x_1 + 30x_2 + 40x_3 &= 1 \end{aligned}$$

has no solutions.

However, something can be said. Suppose that our inhomogeneous system $A\mathbf{x} = \mathbf{b}$ does have at least one solution: $\mathbf{u}$, say. Then by Lemma C4.5, the set of all solutions consists of the vectors $\mathbf{u} + \mathbf{w}$ where $\mathbf{w}$ is a solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$. So, the set of solutions of the inhomogeneous system is obtained from the set of solutions of the homogeneous system by translating it by the vector $\mathbf{u}$, as in Figure C.2.

Example C8.2 Take an inhomogeneous system $A\mathbf{x} = \mathbf{b}$, and suppose that the solutions of the homogeneous system $A\mathbf{x} = \mathbf{0}$ form a 2-dimensional subspace of $\mathbb{R}^n$—that is, a plane through the origin. Then the set of all solutions of the inhomogeneous system is either empty or a plane (although not one that passes through the origin). Put another way, either the inhomogeneous system has no solutions, or its general solution has two free variables.

Next time: we gain a geometric viewpoint on much of the algebra we have developed.

Summary of Chapter C

This is for you to fill in.

The most important definitions and ideas in this chapter

The most important results in this chapter

Points I didn't understand

Chapter D

Linear transformations

To be read before the lecture of Monday, 5 November 2018

In the last chapter, we showed that matrices are very useful for understanding and solving linear systems. That's an *algebraic* use of matrices. In this chapter, we'll focus on a *geometric* use of matrices. We'll show that they are closely related to linear transformations—things like rotations, reflections and projections, which are often easy to visualize and draw.

A linear transformation is a function of a certain kind. So for this chapter, it's important that you're in control of the concept of function and the associated notation and terminology. *I strongly suggest that you reread the summary of functions on pages 14–16.* There are two mistakes that students make particularly often. The first is forgetting that

Every function comes with a specified domain and codomain.

For instance, the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by $x \mapsto x^2$ is not equal to the function $g\colon \mathbb{R} \to \mathbb{C}$ defined by $x \mapsto x^2$. This is simply by definition of function: f and g have different codomains, so cannot be equal. The second mistake is to confuse the arrows $\to$ and $\mapsto$. The $\to$ symbol goes between sets (as in '$f\colon \mathbb{R} \to \mathbb{R}$') and the $\mapsto$ symbol goes between elements (as in '$x \mapsto x^2$'). Don't get them muddled up! And see pages 14–16 for further explanation.

D1 Definition and examples

Often we are interested in ways of transforming vectors in $\mathbb{R}^n$, such as by rotating or reflecting or stretching them. These are all methods for taking one vector in $\mathbb{R}^n$ and producing another. More generally, we will be interested in ways of taking a vector in $\mathbb{R}^n$ and producing a vector in $\mathbb{R}^m$, where perhaps $m \neq n$. For instance, when the sun is out, every point in 3-dimensional space $\mathbb{R}^3$ casts a shadow on the ground, which is the 2-dimensional space $\mathbb{R}^2$.

Thus, we are interested in functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. Since we're doing *linear* algebra, we're primarily interested in those functions that 'respect the linear structure'. For example, if you rotate or reflect or stretch a plane in $\mathbb{R}^3$, you get another plane. The shadow cast on the ground by a plane is usually a

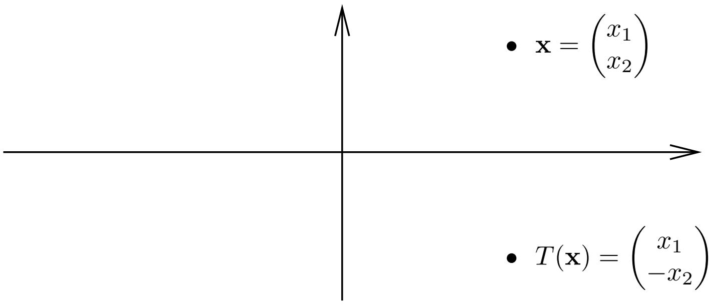

$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$
$T(\mathbf{x}) = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$

Figure D.1: Reflection in the horizontal axis as a linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^2$

plane, and although it is occasionally a line (if the plane aligns exactly with the sun), it is never curved. No bending occurs.

Put another way, a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a function $\mathbb{R}^n \to \mathbb{R}^m$ that 'gets along well with addition and scalar multiplication', in the following exact sense.

Definition D1.1 Let $n, m \geq 0$. A **linear transformation** (or **linear map**, or **linear mapping**) from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a function $T\colon \mathbb{R}^n \to \mathbb{R}^m$ with the following properties:

i. $T(\mathbf{0}) = \mathbf{0}$;

ii. $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;

iii. $T(c\mathbf{x}) = cT(\mathbf{x})$ for all $c \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$.

These might remind you of the three conditions in the definition of subspace (Definition B1.1).

In condition (i), the $\mathbf{0}$ on the left-hand side is the zero vector of $\mathbb{R}^n$, whereas the $\mathbf{0}$ on the right-hand side is the zero vector of $\mathbb{R}^m$. This is the only possible interpretation if the equation is to make sense.

Examples D1.2 i. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be reflection in the x-axis. Thus,

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}$$

for all $x_1, x_2 \in \mathbb{R}$ (Figure D.1). (Strictly speaking, we should have written $T\big(\big(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\big)\big)$ on the left-hand side, but we usually drop one pair of brackets.)

I claim that T is a linear transformation. To prove this, we check the three conditions in Definition D1.1:

- $T(\mathbf{0}) = T\big(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\big) = \big(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\big) = \mathbf{0}$.
- For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$,

$$T(\mathbf{x}+\mathbf{y}) = T \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ -(x_2 + y_2) \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ -y_2 \end{pmatrix} = T(\mathbf{x}) + T(\mathbf{y}).$$

- For $c \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^2$,

$$T(c\mathbf{x}) = T\begin{pmatrix} cx_1 \\ cx_2 \end{pmatrix} = \begin{pmatrix} cx_1 \\ -cx_2 \end{pmatrix} = c\begin{pmatrix} x_1 \\ -x_2 \end{pmatrix} = cT(\mathbf{x}).$$

So T is indeed a linear transformation.

ii. Define a function $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 + 2x_2 \\ x_2 - 4x_1 \\ -x_2 \end{pmatrix}$$

$(x_1, x_2 \in \mathbb{R})$. Again, I claim that T is a linear transformation. Let's check the three conditions:

- $T(\mathbf{0}) = T\left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}\right) = \mathbf{0}$.
- For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$,

$$\begin{aligned} T(\mathbf{x} + \mathbf{y}) &= T\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} \\ &= \begin{pmatrix} 3(x_1 + y_1) + 2(x_2 + y_2) \\ (x_2 + y_2) - 4(x_1 + y_1) \\ -(x_2 + y_2) \end{pmatrix} \\ &= \begin{pmatrix} 3x_1 + 2x_2 \\ x_2 - 4x_1 \\ -x_2 \end{pmatrix} + \begin{pmatrix} 3y_1 + 2y_2 \\ y_2 - 4y_1 \\ -y_2 \end{pmatrix} \\ &= T(\mathbf{x}) + T(\mathbf{y}). \end{aligned}$$

- For $c \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^2$,

$$T(c\mathbf{x}) = T\begin{pmatrix} cx_1 \\ cx_2 \end{pmatrix} = \begin{pmatrix} 3cx_1 + 2cx_2 \\ cx_2 - 4cx_1 \\ -cx_2 \end{pmatrix} = c\begin{pmatrix} 3x_1 + 2x_2 \\ x_2 - 4x_1 \\ -x_2 \end{pmatrix} = cT(\mathbf{x}).$$

This proves that T is a linear transformation, as claimed.

iii. Define a function $T\colon \mathbb{R}^2 \to \mathbb{R}^1 = \mathbb{R}$ by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3x_1 - 2x_2$$

$(x_1, x_2 \in \mathbb{R})$. You can show that T is a linear transformation by checking the three conditions in Definition D1.1, much as in the previous two examples (exercise).

iv. Define functions $S, T\colon \mathbb{R}^2 \to \mathbb{R}^2$ by

$$S\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ x_1 \end{pmatrix}, \qquad T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 \\ x_2 \end{pmatrix}$$

$(x_1, x_2 \in \mathbb{R})$. Then neither S nor T is a linear transformation.

To prove that S is not, it is enough to show that it fails just *one* of the three conditions in Definition D1.1. In fact, it fails the first one, since $S(\mathbf{0}) \neq \mathbf{0}$.

To prove that T is not, note that

$$T\left(2\begin{pmatrix}1\\0\end{pmatrix}\right) = T\begin{pmatrix}2\\0\end{pmatrix} = \begin{pmatrix}4\\0\end{pmatrix}$$

whereas

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

and so $T\bigl(2\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)\bigr) \neq 2T\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$. So, T fails the third condition in Definition D1.1, and is therefore not a linear transformation.

In the first two examples above, checking the three conditions was quite tedious. We will see a *much* quicker way of doing it in Section D3. For instance, this approach will enable us to prove without fuss that rotations in the plane are linear transformations.

Meanwhile, a small saving in labour can be made by taking advantage of the following lemma. It shows that the three conditions in Definition D1.1 can, in fact, be reduced to one:

Lemma D1.3 *Let $n, m \geq 0$, and let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a function. Then T is a linear transformation if and only if for all $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,*

$$T(a\mathbf{x} + b\mathbf{y}) = aT(\mathbf{x}) + bT(\mathbf{y}).$$

Proof First suppose that T is a linear transformation. Then for all $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have

$$T(a\mathbf{x} + b\mathbf{y}) = T(a\mathbf{x}) + T(b\mathbf{y}) = aT(\mathbf{x}) + bT(\mathbf{y})$$

where the first equality follows from condition (ii) of Definition D1.1 and the second follows from condition (iii).

Conversely, suppose that $T(a\mathbf{x} + b\mathbf{y}) = aT(\mathbf{x}) + bT(\mathbf{y})$ for all $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. We show that the three conditions in Definition D1.1 hold.

i. We have $T(1\,\mathbf{0}+1\,\mathbf{0}) = 1T(\mathbf{0})+1T(\mathbf{0})$, or equivalently $T(\mathbf{0}) = T(\mathbf{0})+T(\mathbf{0})$. Subtracting $T(\mathbf{0})$ from each side gives $\mathbf{0} = T(\mathbf{0})$.

ii. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then $T(1\mathbf{x} + 1\mathbf{y}) = 1T(\mathbf{x}) + 1T(\mathbf{y})$, or equivalently $T(\mathbf{x} + \mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$, as required.

iii. Let $c \in \mathbb{R}$ and $\mathbf{x} \in \mathbb{R}^n$. Then $T(c\mathbf{x} + 0\mathbf{x}) = cT(\mathbf{x}) + 0T(\mathbf{x})$, or equivalently $T(c\mathbf{x}) = cT(\mathbf{x})$, as required. □

Examples D1.4 i. Let's use Lemma D1.3 to show that the function $T\colon \mathbb{R}^2 \to \mathbb{R}$ of Example D1.2(iii) is a linear transformation. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and $a, b \in \mathbb{R}$. Then

$$\begin{aligned} T(a\mathbf{x} + b\mathbf{y}) &= T\begin{pmatrix}ax_1 + by_1\\ ax_2 + by_2\end{pmatrix} \\ &= 3(ax_1 + by_1) - 2(ax_2 + by_2) \\ &= a(3x_1 - 2x_2) + b(3y_1 - 2y_2) \\ &= aT(\mathbf{x}) + bT(\mathbf{y}). \end{aligned}$$

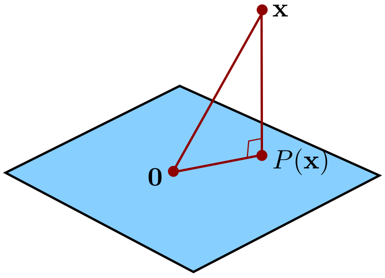


Figure D.2: A point $\mathbf{x}$ in the air and its shadow $P(\mathbf{x})$ on the ground

Hence by Lemma D1.3, T is a linear transformation.

ii. Define a function $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ by

$$T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$$

$(x_1, x_2 \in \mathbb{R})$. Then T is a linear transformation. Again, you can easily check this using Lemma D1.3 (exercise).

In the introduction to this section, I mentioned shadows cast by the sun as an informal example of a linear transformation. When the sun is directly overhead, every point $\mathbf{x}$ in the air is mapped to the closest point $P(\mathbf{x})$ on the ground, much as in Figure D.2 or the similar Figure B.6 (page 66). We now examine this situation more closely.

Let $n \geq 0$ and let V be a subspace of $\mathbb{R}^n$. (In the previous paragraph, n was 3 and V was 2-dimensional.) In Proposition B7.8(ii), we proved that for every $\mathbf{x} \in \mathbb{R}^n$, there are unique vectors $\mathbf{v} \in V$ and $\mathbf{w} \in V^\perp$ such that $\mathbf{x} = \mathbf{v} + \mathbf{w}$. Put another way, for every $\mathbf{x} \in \mathbb{R}^n$, there is a unique vector $\mathbf{v} \in V$ such that $\mathbf{x} - \mathbf{v} \in V^\perp$. So, the following definition is logically valid.

Definition D1.5 Let V be a linear subspace of $\mathbb{R}^n$ and let $\mathbf{x} \in \mathbb{R}^n$. The **orthogonal projection** of $\mathbf{x}$ onto V is the unique vector $P_V(\mathbf{x}) \in V$ such that $\mathbf{x} - P_V(\mathbf{x}) \in V^\perp$.

(Again, see Figures B.6 and D.2.) Another point of view is that $P_V(\mathbf{x})$ is the point of V closest to $\mathbf{x}$. You'll prove this during a workshop.

This describes $P_V(\mathbf{x})$ geometrically. How can we describe it algebraically? The next lemma gives the answer.

Lemma D1.6 *Let V be a linear subspace of $\mathbb{R}^n$. Let $\mathbf{x} \in \mathbb{R}^n$, and write $P_V(\mathbf{x})$ for the orthogonal projection of $\mathbf{x}$ onto V. Then*

$$P_V(\mathbf{x}) = \sum_{i=1}^{m} (\mathbf{x} \cdot \mathbf{v}_i)\mathbf{v}_i$$

for any orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of V.

Proof By definition, $P_V(\mathbf{x})$ is the unique element of V such that $\mathbf{x} - P_V(\mathbf{x}) \in V^\perp$. But by Lemma B7.5, $\sum(\mathbf{x} \cdot \mathbf{v}_i)\mathbf{v}_i$ is an element of V with this property. Hence $P_V(\mathbf{x}) = \sum(\mathbf{x} \cdot \mathbf{v}_i)\mathbf{v}_i$. □

Given a subspace V of $\mathbb{R}^n$, we have shown how each point $\mathbf{x} \in \mathbb{R}^n$ gives rise to another point of $\mathbb{R}^n$, its orthogonal projection $P_V(\mathbf{x})$ onto V. This defines a function $P_V \colon \mathbb{R}^n \to \mathbb{R}^n$. And this function is, in fact, a linear transformation:

Lemma D1.7 *Let V be a subspace of $\mathbb{R}^n$. Define a function $P_V \colon \mathbb{R}^n \to \mathbb{R}^n$ by taking $P_V(\mathbf{x})$ to be the orthogonal projection of $\mathbf{x}$ onto V, for each $\mathbf{x} \in \mathbb{R}^n$. Then P_V is a linear transformation.*

Proof Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$. We will show that $P_V(a\mathbf{x}+b\mathbf{y}) = aP_V(\mathbf{x}) + bP_V(\mathbf{y})$.

By definition, $P_V(a\mathbf{x} + b\mathbf{y})$ is the *unique* point of V such that

$$(a\mathbf{x} + b\mathbf{y}) - P_V(a\mathbf{x} + b\mathbf{y}) \in V^\perp.$$

So, it is enough to prove that $aP_V(\mathbf{x}) + bP_V(\mathbf{y})$ also has this property. That is, it is enough to prove that $aP_V(\mathbf{x}) + bP_V(\mathbf{y})$ is a point of V satisfying

$$(a\mathbf{x} + b\mathbf{y}) - (aP_V(\mathbf{x}) + bP_V(\mathbf{y})) \in V^\perp. \tag{D:1}$$

To see this, first note that $aP_V(\mathbf{x}) + bP_V(\mathbf{y})$ is indeed a point of V, since $P_V(\mathbf{x}), P_V(\mathbf{y}) \in V$ and V is a linear subspace of $\mathbb{R}^n$. Second,

$$(a\mathbf{x} + b\mathbf{y}) - (aP_V(\mathbf{x}) + bP_V(\mathbf{y})) = a(\mathbf{x} - P_V(\mathbf{x})) + b(\mathbf{y} - P_V(\mathbf{y})).$$

Now $\mathbf{x} - P_V(\mathbf{x}) \in V^\perp$ and $\mathbf{y} - P_V(\mathbf{y}) \in V^\perp$ by definition of P_V, so any linear combination of them also belongs to $V^\perp$. Hence (D:1) holds, as required. □

Examples D1.8 Consider orthogonal projection onto a subspace V of $\mathbb{R}^3$. Then $\dim V \in \{0, 1, 2, 3\}$, so there are four cases to consider:

i. Suppose that $\dim V = 0$. Then V is the trivial subspace $\{\mathbf{0}\}$, so $P_V(\mathbf{x}) = \mathbf{0}$ for all $\mathbf{x} \in V$.

ii. Suppose that $\dim V = 1$. Then V is a line through the origin, and P_V maps each point $\mathbf{x} \in \mathbb{R}^3$ to the closest point $P_V(\mathbf{x})$ to $\mathbf{x}$ on the line V. For instance, if $V = \operatorname{span}\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}\right\}$ (the 'x-axis') then $P_V\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}x_1\\0\\0\end{pmatrix}$.

iii. Suppose that $\dim V = 2$. Then V is a plane through the origin, and again, P_V maps $\mathbf{x} \in \mathbb{R}^3$ to the closest point $P_V(\mathbf{x})$ on V. For instance, if $V = \operatorname{span}\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}\right\}$ then $P_V\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}x_1\\x_2\\0\end{pmatrix}$ (Figure D.2).

iv. Suppose that $\dim V = 3$. Then $V = \mathbb{R}^3$, and $P_V(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^3$. That is, P_V is the identity mapping.

D2 The standard matrix of a linear transformation

We've seen various examples of linear transformations, some motivated by geometric considerations (such as orthogonal projections) and some given by an algebraic formula. One might ask: what are *all* the linear transformations $\mathbb{R}^n \to \mathbb{R}^m$? Is there any way of organizing them?

In this section and the next, we'll completely answer these questions. We'll see that linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ are very closely related to $m \times n$ matrices. Any linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ gives rise to an $m \times n$ matrix (called its 'standard matrix'), and any $m \times n$ matrix gives rise to a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$. This back-and-forth process sets up a one-to-one correspondence between linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ and $m \times n$ matrices, as we will see.

We begin with a simple observation. Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Any vector $\mathbf{x} \in \mathbb{R}^n$ can be expressed as a linear combination of the standard basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^n$:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n.$$

Since T is linear, this implies that

$$T(\mathbf{x}) = x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) + \cdots + x_n T(\mathbf{e}_n)$$

for all $\mathbf{x} \in \mathbb{R}^n$ (using Lemma D1.3 and induction). But this tells us something important: *the whole of the linear transformation T is determined by the vectors $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n) \in \mathbb{R}^m$*. In other words, if we have two linear transformations $T, S\colon \mathbb{R}^n \to \mathbb{R}^m$ with $T(\mathbf{e}_j) = S(\mathbf{e}_j)$ for each $j \in \{1, \ldots, n\}$, then $T(\mathbf{x})$ and $S(\mathbf{x})$ must be equal for *all* $\mathbf{x} \in \mathbb{R}^n$.

We have just shown that given vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n \in \mathbb{R}^m$, there is *at most* one linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ such that

$$T(\mathbf{e}_1) = \mathbf{u}_1,\ T(\mathbf{e}_2) = \mathbf{u}_2,\ \ldots,\ T(\mathbf{e}_n) = \mathbf{u}_n.$$

The result we are about to prove states that, in fact, there is *exactly* one. Better still, none of this is specific to the standard basis: the same is true for any basis of $\mathbb{R}^n$ whatsoever.

Proposition D2.1 *Let $n, m \geq 0$, let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of $\mathbb{R}^n$, and let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be any vectors in $\mathbb{R}^m$. Then there is exactly one linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$T(\mathbf{v}_1) = \mathbf{u}_1,\ T(\mathbf{v}_2) = \mathbf{u}_2,\ \ldots,\ T(\mathbf{v}_n) = \mathbf{u}_n.$$

Proof First we show that there is at most one such T, then that there is at least one such T.

At most one Let $T, S\colon \mathbb{R}^n \to \mathbb{R}^m$ be linear transformations such that $T(\mathbf{v}_j) = \mathbf{u}_j$ and $S(\mathbf{v}_j) = \mathbf{u}_j$ for all $j \in \{1, \ldots, n\}$. We must show that $T = S$.

Since T and S are functions $\mathbb{R}^n \to \mathbb{R}^m$, this means showing that $T(\mathbf{x}) = S(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. Let $\mathbf{x} \in \mathbb{R}^n$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ span $\mathbb{R}^n$, we can write

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

for some $c_1, \ldots, c_n \in \mathbb{R}$. So by Lemma D1.3 and induction,

$$T(\mathbf{x}) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_nT(\mathbf{v}_n).$$

But $T(\mathbf{v}_j) = \mathbf{u}_j$ for each j, so

$$T(\mathbf{x}) = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n. \tag{D:2}$$

Similarly,

$$S(\mathbf{x}) = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n.$$

Hence $T(\mathbf{x}) = S(\mathbf{x})$, as required.

At least one We must prove that there exists a linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ such that $T(\mathbf{v}_j) = \mathbf{u}_j$ for all $j \in \{1, \ldots, n\}$. Since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of $\mathbb{R}^n$, each vector $\mathbf{x} \in \mathbb{R}^n$ can be written as

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

for unique scalars $c_1, \ldots, c_n \in \mathbb{R}$. Define

$$T(\mathbf{x}) = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n \in \mathbb{R}^m.$$

(The inspiration for that step was (D:2).) This defines a function $T\colon \mathbb{R}^n \to \mathbb{R}^m$ such that $T(\mathbf{v}_j) = \mathbf{u}_j$ for each $j \in \{1, \ldots, n\}$.

To show that T is linear, we use Lemma D1.3. Let $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Write $\mathbf{x} = \sum_{j=1}^n c_j\mathbf{v}_j$ and $\mathbf{y} = \sum_{j=1}^n d_j\mathbf{v}_j$; then

$$a\mathbf{x} + b\mathbf{y} = \sum_{j=1}^n (ac_j + bd_j)\mathbf{v}_j.$$

Hence

$$\begin{aligned} T(a\mathbf{x} + b\mathbf{y}) &= \sum_{j=1}^n (ac_j + bd_j)\mathbf{u}_j \\ &= a\sum_{j=1}^n c_j\mathbf{u}_j + b\sum_{j=1}^n d_j\mathbf{u}_j \\ &= aT(\mathbf{x}) + bT(\mathbf{y}), \end{aligned}$$

as required. □

Example D2.2 Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be any basis of $\mathbb{R}^n$ and let $c_1, \ldots, c_n$ be any scalars. Then by Proposition D2.1, there is exactly one linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^n$ such that $T(\mathbf{v}_j) = c_j\mathbf{v}_j$ for each $j = 1, \ldots, n$.

Geometrically, this transformation scales $\mathbb{R}^n$ by a factor of c_j in the direction of $\mathbf{v}_j$. For example, if $n = 2$ and $\mathbf{v}_1, \mathbf{v}_2$ is the basis of $\mathbb{R}^2$ shown in Figure B.5 (page 63), and if $c_1 = 2$ and $c_2 = 3$, then T scales by a factor of 2 in the north-east direction and 3 in the south-east direction. (Or it might be the other way round, depending on which of the vectors in Figure B.5 is $\mathbf{v}_1$ and which is $\mathbf{v}_2$.)

If any of the scale factors c_j are negative, then T *reverses* the direction parallel to $\mathbf{v}_j$. For instance, take $n = 2$ again, take the standard basis $\mathbf{e}_1, \mathbf{e}_2$, and take $c_1 = 1$ and $c_2 = -1$. Then T is the unique linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$ such that $T(\mathbf{e}_1) = \mathbf{e}_1$ and $T(\mathbf{e}_2) = -\mathbf{e}_2$. This is exactly the T of Example D1.2(i), namely, reflection in the x-axis. Note that reflecting in $\operatorname{span}\{\mathbf{e}_1\}$ (the x-axis) means reversing the direction parallel to $\mathbf{e}_2$.

Proposition D2.1 implies that if you have a basis of $\mathbb{R}^n$, then knowing what a linear transformation $T \colon \mathbb{R}^n \to \mathbb{R}^m$ does to every basis element tells you what it does to *every* point of $\mathbb{R}^n$. Applied to the standard basis, this means that a linear transformation $T \colon \mathbb{R}^n \to \mathbb{R}^m$ is completely determined by the vectors $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n) \in \mathbb{R}^m$. Since these n m-dimensional vectors completely specify T, it is natural to compile them into an $m \times n$ matrix and give it a name:

Definition D2.3 Let $T \colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The **standard matrix** of T is the $m \times n$ matrix

$$[T] = \big(T(\mathbf{e}_1) | T(\mathbf{e}_2) | \cdots | T(\mathbf{e}_n)\big).$$

That is, it is the $m \times n$ matrix $[T]$ whose jth column is $T(\mathbf{e}_j)$, where $\mathbf{e}_j$ is the jth standard basis vector of $\mathbb{R}^n$.

Examples D2.4 i. Let $T \colon \mathbb{R}^2 \to \mathbb{R}^2$ be reflection in the x-axis, as in Example D1.2(i). Then

$$T(\mathbf{e}_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad T(\mathbf{e}_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

so T has standard matrix

$$[T] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

ii. Take the linear transformation $T \colon \mathbb{R}^2 \to \mathbb{R}^3$ of Example D1.2(ii). Then

$$T(\mathbf{e}_1) = \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix}, \quad T(\mathbf{e}_2) = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix},$$

so the standard matrix of T is the 3×2 matrix

$$[T] = \begin{pmatrix} 3 & 2 \\ -4 & 1 \\ 0 & -1 \end{pmatrix}.$$

iii. The standard matrix of the linear transformation $T \colon \mathbb{R}^2 \to \mathbb{R}$ defined in Example D1.2(iii) is $(3 \quad -2)$, similarly.

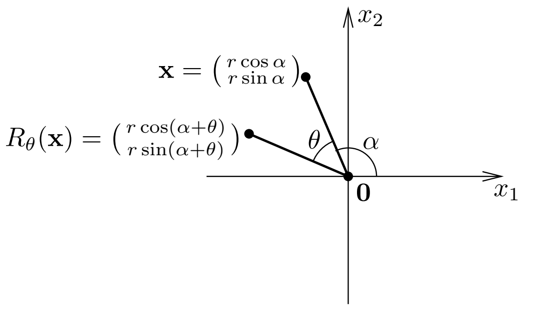

x_2
$\mathbf{x} = \binom{r\cos\alpha}{r\sin\alpha}$
$R_\theta(\mathbf{x}) = \binom{r\cos(\alpha+\theta)}{r\sin(\alpha+\theta)}$
θ
α
$\mathbf{0}$
x_1

Figure D.3: Rotation of the plane by an angle θ about the origin

iv. Let $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ be orthogonal projection onto the plane span$\{\mathbf{e}_1, \mathbf{e}_2\}$, as in Example D1.8(iii). Then

$$T(\mathbf{e}_1) = \mathbf{e}_1, \quad T(\mathbf{e}_2) = \mathbf{e}_2, \quad T(\mathbf{e}_3) = \mathbf{0},$$

so

$$[T] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

D3 Matrices vs. linear transformations

We have just seen that any linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ gives rise to an $m \times n$ matrix. But in the other direction, any $m \times n$ matrix gives rise to a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$, as follows.

Let A be an $m \times n$ matrix. Then any vector $\mathbf{x} \in \mathbb{R}^n$ can be multiplied by A to give a vector $A\mathbf{x} \in \mathbb{R}^m$. This defines a function $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$. In other words, we define

$$L_A(\mathbf{x}) = A\mathbf{x} \in \mathbb{R}^m \tag{D:3}$$

for each $\mathbf{x} \in \mathbb{R}^n$.

Lemma D3.1 *Let A be an $m \times n$ matrix. Then the function $L_A\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation.*

Proof By Lemma D1.3, it is enough to show that $L_A(a\mathbf{x} + b\mathbf{y}) = aL_A(\mathbf{x}) + bL_A(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and scalars a, b. And indeed,

$$L_A(a\mathbf{x} + b\mathbf{y}) = A(a\mathbf{x} + b\mathbf{y}) = aA\mathbf{x} + bA\mathbf{y} = aL_A(\mathbf{x}) + bL_A(\mathbf{y}),$$

using the basic identities for matrix multiplication in equation (A:12) (page 34). □

(Beware: the notation L_A is just something I have made up for this course. Other people use other notation. Poole writes L_A as T_A, for instance.)

Examples D3.2 i. Let $\theta \in \mathbb{R}$. Let $R_\theta\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the function that rotates by an angle θ anticlockwise about $\mathbf{0}$ (Figure D.3). As the figure shows, for a point $\mathbf{x} \in \mathbb{R}^2$ with polar coordinates (r, α), the point $R_\theta(\mathbf{x})$ has polar coordinates $(r, \alpha + \theta)$. Now using standard trigonometric formulas,

$$\begin{aligned} R_\theta(\mathbf{x}) = \begin{pmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{pmatrix} &= \begin{pmatrix} r\cos(\alpha)\cos(\theta) - r\sin(\alpha)\sin(\theta) \\ r\cos(\alpha)\sin(\theta) + r\sin(\alpha)\cos(\theta) \end{pmatrix} \\ &= \begin{pmatrix} x_1\cos\theta - x_2\sin\theta \\ x_1\sin\theta + x_2\cos\theta \end{pmatrix} = A\mathbf{x} \end{aligned}$$

where

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

We have just shown that $R_\theta(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^2$. But this means exactly that $R_\theta = L_A$. So by Lemma D3.1, R_θ is linear.

The linearity of R_θ should also be clear visually. The sum $\mathbf{x} + \mathbf{y}$ of two vectors $\mathbf{x}$ and $\mathbf{y}$ is the fourth corner of the parallelogram whose other three corners are $\mathbf{0}$, $\mathbf{x}$ and $\mathbf{y}$ (Figure A.1, page 22). But if you rotate a parallelogram, the result is still a parallelogram. Hence $R_\theta(\mathbf{x} + \mathbf{y})$ is the fourth corner of the parallelogram whose other three corners are $\mathbf{0}$, $R_\theta(\mathbf{x})$ and $R_\theta(\mathbf{y})$. In other words, $R_\theta(\mathbf{x}+\mathbf{y}) = R_\theta(\mathbf{x}) + R_\theta(\mathbf{y})$. A similar geometric argument shows that $R_\theta(c\mathbf{x}) = cR_\theta(\mathbf{x})$ (exercise).

ii. Take the 3×2 matrix

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

Then L_A is the linear transformation $\mathbb{R}^2 \to \mathbb{R}^3$ defined by $L_A(\mathbf{x}) = A\mathbf{x}$ ($\mathbf{x} \in \mathbb{R}^2$). So,

$$L_A(\mathbf{x}) = \begin{pmatrix} 3 & 2 \\ -4 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 + 2x_2 \\ x_2 - 4x_1 \\ -x_2 \end{pmatrix}$$

for all $\mathbf{x} \in \mathbb{R}^2$.

We first met this linear transformation in Example D1.2(ii), where it was called T. There, we checked laboriously that it was linear by verifying the three conditions directly. But this example demonstrates a quicker way: once we have recognized that $T = L_A$ for some matrix A, Lemma D3.1 implies immediately that T is linear.

Examples D2.4(ii) and D3.2(ii) both involve the linear transformation

$$\begin{array}{cccc} T\colon & \mathbb{R}^2 & \to & \mathbb{R}^3 \\ & \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} & \mapsto & \begin{pmatrix} 3x_1 + 2x_2 \\ x_2 - 4x_1 \\ -x_2 \end{pmatrix} \end{array} \tag{D:4}$$

and the matrix

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

Example D2.4(ii) states that $A = [T]$ (that is, A is the standard matrix of T), whereas Example D3.2(ii) states that $T = L_A$. Is it coincidence that *both* statements are true? For instance, if we start with any linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$, take its standard matrix $[T]$, and then take the linear transformation $L_{[T]}\colon \mathbb{R}^n \to \mathbb{R}^m$, is it always equal to the linear transformation T that we started with? Or did they just happen to be equal for this particular linear transformation T?

There was no coincidence; it is indeed true for all linear transformations:

Lemma D3.3 *Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $L_{[T]} = T$. Thus, $[T]\mathbf{x} = T(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$.*

Proof First we check that the linear transformations T and $L_{[T]}$ have the same domain and codomain. (If you don't see why this is necessary, see page 15.) By definition, $[T]$ is an $m \times n$ matrix, so $L_{[T]}$ is a map $\mathbb{R}^n \to \mathbb{R}^m$. So is T. Hence T and $L_{[T]}$ do indeed have the same domain and codomain.

For each $j \in \{1, \ldots, n\}$, we have $L_{[T]}(\mathbf{e}_j) = [T]\mathbf{e}_j$ (by definition of $L_{[T]}$). But $[T]\mathbf{e}_j$ is the jth column of the matrix $[T]$ (by Lemma A4.3(ii)), which by Definition D2.3 is $T(\mathbf{e}_j)$.

We have just shown that $L_{[T]}$ and T are linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ satisfying $L_{[T]}(\mathbf{e}_j) = T(\mathbf{e}_j)$ for each $j \in \{1, \ldots, n\}$. But Proposition D2.1 implies that there is only *one* linear transformation satisfying these equations. Hence $L_{[T]} = T$, as required.

Finally, for each $\mathbf{x} \in \mathbb{R}^n$ we have $L_{[T]}(\mathbf{x}) = T(\mathbf{x})$; but by definition of $L_{[T]}$, this means that $[T]\mathbf{x} = T(\mathbf{x})$. □

Remark D3.4 Students often find proofs like this hard at first. The key is to take it slowly, referring carefully to the definitions. You should constantly ask yourself 'what type of thing is this?' For example, when you see the expression $[T]\mathbf{e}_j$, you should say to yourself: $[T]$ is an $m \times n$ matrix, and $\mathbf{e}_j$ is an n-dimensional vector, so the matrix product $[T]\mathbf{e}_j$ makes sense and is an m-dimensional vector. Or when you see $T(\mathbf{e}_j)$, you should think: T is a function $\mathbb{R}^n \to \mathbb{R}^m$, and $\mathbf{e}_j \in \mathbb{R}^n$, so T can be applied to $\mathbf{e}_j$ to give an element $T(\mathbf{e}_j)$ of $\mathbb{R}^m$. Take your time, and breathe.

Lemma D3.3 says that if we start with a linear transformation T, turn it into a matrix $[T]$, then turn *that* into a linear transformation $L_{[T]}$, we get back exactly what we started with. It's natural to ask what happens the other way round: if we take a matrix A, turn it into a linear transformation L_A, then turn *that* into a matrix $[L_A]$, is the result just A again? Yes:

Lemma D3.5 *Let A be an $m \times n$ matrix. Then $[L_A] = A$.*

Proof First we check that the matrices $[L_A]$ and A have the same number of rows and columns. By definition, L_A is a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$, so $[L_A]$ is an $m \times n$ matrix. So is A. Hence $[L_A]$ and A do indeed have the same number of rows and columns.

Let $j \in \{1, \ldots, n\}$. The jth column of $[L_A]$ is $L_A(\mathbf{e}_j)$ (by Definition D2.3). But $L_A(\mathbf{e}_j) = A\mathbf{e}_j$ (by definition of L_A), and $A\mathbf{e}_j$ is the jth column of A (by Lemma A4.3(ii)). Hence $[L_A]$ and A have the same jth column. This holds for all j, so $[L_A] = A$. □

Example D3.6 Again, consider the 3×2 matrix A of (D:5). Example D3.2(ii) shows that L_A is the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ of (D:4). Example D2.4(ii) shows that $[L_A]$, the standard matrix of L_A, is equal to A. This is exactly what Lemma D3.5 predicts.

Putting together the last two lemmas gives a complete description of all linear transformations in terms of matrices:

Theorem D3.7 *Let $m, n \geq 0$. There is a one-to-one correspondence between*

$$\{\textit{linear transformations } \mathbb{R}^n \to \mathbb{R}^m\} \qquad \textit{and} \qquad \{m \times n \textit{ matrices}\},$$

with a linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ corresponding to its standard matrix $[T]$, and an $m \times n$ matrix A corresponding to the linear transformation L_A.

Proof The statement of the theorem means three things: (i) that if T is a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ then $[T]$ is an $m \times n$ matrix; (ii) that if A is an $m \times n$ matrix then L_A is a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$; and (iii) that the two processes $T \mapsto [T]$ and $A \mapsto L_A$ are inverse to one another (that is, $L_{[T]} = T$ for all T and $[L_A] = A$ for all A). We have already proved every part of this. □

For example, under this one-to-one correspondence, the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^3$ of (D:4) corresponds to the 3×2 matrix A of (D:5).

Corollary D3.8 *Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there is exactly one $m \times n$ matrix A such that $T = L_A$.*

Proof This follows immediately from Theorem D3.7. Alternatively, we can deduce it from the previous lemmas. To prove there is at least one such A, put $A = [T]$: then $L_A = L_{[T]} = T$ by Lemma D3.3. To prove there is only one such A, let B be an $m \times n$ matrix such that $T = L_B$: then $B = [L_B] = [T]$ by Lemma D3.5. □

This corollary implies that *every* linear transformation can be expressed as multiplication by a matrix.

Example D3.9 Let V be a linear subspace of $\mathbb{R}^n$. By Lemma B7.5, orthogonal projection onto V is a linear transformation $P_V\colon \mathbb{R}^n \to \mathbb{R}^n$. So, there is one (and only one) $n \times n$ matrix A such that for all vectors $\mathbf{x}$,

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

This matrix A is simply the standard matrix of P_V.

Warning D3.10 We have proved that there is a very close relationship between linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ and $m \times n$ matrices...

$$\boxed{\textit{But linear transformations and matrices are not the same!}}$$

A linear transformation is a *function* with certain properties. A matrix is a *grid of numbers*. So, when T is a linear transformation and A is a matrix, it never makes sense to write '$T = A$', any more than it makes sense to write '$\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = 5$'. In both cases, the left- and right-hand sides are simply different types of thing. Many mistakes that students make come from treating linear transformations as matrices, or vice versa.

D4 Composing and inverting linear transformations

If doing a linear transformation is in any way interesting or useful, then doing one transformation after another is also likely to be interesting or useful. For instance, reflecting in any line through the origin defines a linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$, so we might ask: what is the combined effect of reflecting in one line and then another? (I'll leave this particular question for you to think about.)

In this section, we first look at what happens when you compose linear transformations. Then we consider inverses of linear transformations: when they exist, and what they are.

First recall from pages 14–16 the notion of the *composite* $g \circ f$ (or gf) of two functions $A \xrightarrow{f} B \xrightarrow{g} C$. It is only defined if the codomain of f is equal to the domain of B! Also recall from there the notion of the *identity* function 1_A on a set A. It has the properties that $1_A \circ f = f$ for any function $f\colon Z \to A$, and similarly $g \circ 1_A = g$ for any function $g\colon A \to B$.

Lemma D4.1 *i. Let $p, n, m \geq 0$, and let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ and $U\colon \mathbb{R}^p \to \mathbb{R}^n$ be linear transformations. Then the composite $T \circ U\colon \mathbb{R}^p \to \mathbb{R}^m$ is also a linear transformation.*

ii. Let $n \geq 0$. Then the identity function $1_{\mathbb{R}^n}\colon \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation.

Proof We use Lemma D1.3. For (i), let $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$. Then

$$\begin{aligned}
(T \circ U)(a\mathbf{x} + b\mathbf{y}) &= T(U(a\mathbf{x} + b\mathbf{y})) && \text{by definition of } T \circ U \\
&= T(aU(\mathbf{x}) + bU(\mathbf{y})) && \text{by linearity of } U \\
&= aT(U(\mathbf{x})) + bT(U(\mathbf{y})) && \text{by linearity of } T \\
&= a(T \circ U)(\mathbf{x}) + b(T \circ U)(\mathbf{y}) && \text{by definition of } T \circ U.
\end{aligned}$$

For (ii), let $a, b \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then

$$1_{\mathbb{R}^n}(a\mathbf{x} + b\mathbf{y}) = a\mathbf{x} + b\mathbf{y} = a1_{\mathbb{R}^n}(\mathbf{x}) + b1_{\mathbb{R}^n}(\mathbf{y}). \qquad \square$$

Example D4.2 In Example D3.2(i), we constructed the linear transformation $R_{\pi/2}\colon \mathbb{R}^2 \to \mathbb{R}^2$ that rotates the plane about $\mathbf{0}$ by an angle of $\pi/2$, and in Lemma D1.7, we showed that orthogonal projection P_V of $\mathbb{R}^2$ onto the x-axis $V = \operatorname{span}\{\mathbf{e}_1\}$ is also a linear transformation $\mathbb{R}^2 \to \mathbb{R}^2$. So there is a composite linear transformation $P_V \circ R_{\pi/2}\colon \mathbb{R}^2 \to \mathbb{R}^2$:

$$\begin{tikzcd}
\mathbb{R}^2 \arrow[r, "R_{\pi/2}"] \arrow[rr, bend right, "P_V \circ R_{\pi/2}"'] & \mathbb{R}^2 \arrow[r, "P_V"] & \mathbb{R}^2.
\end{tikzcd}$$

This composite transformation $P_V \circ R_{\pi/2}$ takes a point in the plane, rotates it by $\pi/2$, then projects it onto the x-axis. For instance, $R_{\pi/2}(\mathbf{e}_1) = \mathbf{e}_2$, so

$$(P_V \circ R_{\pi/2})(\mathbf{e}_1) = P_V(R_{\pi/2}(\mathbf{e}_1)) = P_V(\mathbf{e}_2) = \mathbf{0}.$$

Similarly, you can show that $(P_V \circ R_{\pi/2})(\mathbf{e}_2) = -\mathbf{e}_1$.

Given two linear transformations $\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$, you can compose them to get $T \circ U$, then take its standard matrix $[T \circ U]$. Alternatively, you can multiply together the individual standard matrices of T and U to get another matrix, $[T][U]$. We now show that the two end results, $[T \circ U]$ and $[T][U]$, are actually the same. In short:

> *Multiplying matrices corresponds to composing linear transformations.*

(By 'corresponds', I'm referring to the one-to-one correspondence of Theorem D3.7.) In the same spirit, identity matrices correspond to identity linear transformations. All this and more is in the following result.

Lemma D4.3 *i. Let $n, m, p \geq 0$, let A be an $m \times n$ matrix, and let B be an $n \times p$ matrix. Then $L_{AB} = L_A \circ L_B$.*

ii. Let $n \geq 0$. Then $L_{I_n} = 1_{\mathbb{R}^n}$.

iii. Let $n, m, p \geq 0$, and let $T \colon \mathbb{R}^n \to \mathbb{R}^m$ and $U \colon \mathbb{R}^p \to \mathbb{R}^n$ be linear transformations. Then $[T \circ U] = [T][U]$.

iv. Let $n \geq 0$. Then $[1_{\mathbb{R}^n}] = I_n$.

Proof For (i), first note that L_A, L_B and $L_A \circ L_B$ have domains and codomains as shown:

$$\begin{tikzcd} \mathbb{R}^p \arrow[r, "L_B"] \arrow[rr, bend right, "L_A \circ L_B"'] & \mathbb{R}^n \arrow[r, "L_A"] & \mathbb{R}^m. \end{tikzcd}$$

Also AB is an $m \times p$ matrix, so L_{AB} is a linear transformation $\mathbb{R}^p \to \mathbb{R}^m$. Hence $L_A \circ L_B$ and L_{AB} are functions with the same domain ($\mathbb{R}^p$) and the same codomain ($\mathbb{R}^m$). (Again, see page 15 if you don't understand why this step is necessary.) Now for all $\mathbf{x} \in \mathbb{R}^p$, directly from the definitions,

$$(L_A \circ L_B)(\mathbf{x}) = L_A(L_B(\mathbf{x})) = L_A(B\mathbf{x}) = AB\mathbf{x} = L_{AB}(\mathbf{x}).$$

Hence $L_A \circ L_B = L_{AB}$.

For (ii), first note that I_n is an $n \times n$ matrix, so L_{I_n} is a function $\mathbb{R}^n \to \mathbb{R}^n$. So the functions L_{I_n} and $1_{\mathbb{R}^n}$ have the same domain ($\mathbb{R}^n$) and the same codomain (also $\mathbb{R}^n$). Now for all $\mathbf{x} \in \mathbb{R}^n$, directly from the definitions,

$$L_{I_n}(\mathbf{x}) = I_n\mathbf{x} = \mathbf{x} = 1_{\mathbb{R}^n}(\mathbf{x}).$$

Hence $L_{I_n} = 1_{\mathbb{R}^n}$.

We will deduce (iii) from (i) and (iv) from (ii). For (iii), write $A = [T]$ and $B = [U]$. Then $T = L_A$ and $U = L_B$ by Lemma D3.3, so

$$[T \circ U] = [L_A \circ L_B] = [L_{AB}] = AB = [T][U],$$

where the second equality follows from (i) and the third from Lemma D3.5. Finally, for (iv), we have

$$[1_{\mathbb{R}^n}] = [L_{I_n}] = I_n,$$

where the first equality follows from (ii) and the second from Lemma D3.5. □

Example D4.4 Is there a 2×2 matrix A such that $A \neq I$ but $A^{13} = I$? Lemma D4.3 makes this easy to answer. Put $A = [R_{2\pi/13}]$, where the linear transformation $R_{2\pi/13} : \mathbb{R}^2 \to \mathbb{R}^2$ is as defined in Example D3.2(i). We have $R_{2\pi/13} \neq 1_{\mathbb{R}^2}$, and different linear transformations have different standard matrices (by Theorem D3.7), so $A \neq I$. But $R_{2\pi/13}^{13} = 1_{\mathbb{R}^2}$ (where $R_{2\pi/13}^{13}$ means the 13-fold composite $R_{2\pi/13} \circ R_{2\pi/13} \circ \cdots \circ R_{2\pi/13}$), so

$$A^{13} = [R_{2\pi/13}]^{13} = [R_{2\pi/13}^{13}] = [1_{\mathbb{R}^2}] = I$$

by Lemma D4.3 and induction. So A has the properties required.

Now we turn to inverses. Recall the terminology of bijective and inverse functions (pages 14–16).

Lemma D4.5 *Let $n, m \geq 0$. Let $T \colon \mathbb{R}^n \to \mathbb{R}^m$ be a bijective linear transformation. Then the inverse function $T^{-1} \colon \mathbb{R}^m \to \mathbb{R}^n$ is also a linear transformation.*

Proof We use Lemma D1.3. Let $a, b \in \mathbb{R}$ and $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$. By linearity of T,

$$T\big(aT^{-1}(\mathbf{v}) + bT^{-1}(\mathbf{w})\big) = aT(T^{-1}(\mathbf{v})) + bT(T^{-1}(\mathbf{w})) = a\mathbf{v} + b\mathbf{w}.$$

So $T\big(aT^{-1}(\mathbf{v}) + bT^{-1}(\mathbf{w})\big) = a\mathbf{v} + b\mathbf{w}$. Applying T^{-1} to both sides gives $aT^{-1}(\mathbf{v}) + bT^{-1}(\mathbf{w}) = T^{-1}(a\mathbf{v} + b\mathbf{w})$, as required. □

A linear transformation is said to be **invertible** if it has an inverse linear transformation. By Lemma D4.5, a linear transformation is invertible if and only if it is bijective. When T is invertible, we can ask: what is the standard matrix of T^{-1} in terms of that of T?

Lemma D4.6 *A linear transformation T is invertible if and only if its standard matrix is invertible. In that case, $[T^{-1}] = [T]^{-1}$.*

Proof Let $T \colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

First suppose that T is invertible. We have $T^{-1} \circ T = 1_{\mathbb{R}^n}$, so by Lemma D4.3, taking the standard matrix of each side gives $[T^{-1}][T] = I_n$. Similarly, $[T][T^{-1}] = I_m$. It follows that $[T]$ is invertible with inverse $[T^{-1}]$.

Conversely, suppose that the matrix $[T]$ is invertible. We have $[T]^{-1}[T] = I_n$, so by Lemma D4.3, $L_{[T]^{-1}} \circ L_{[T]} = 1_{\mathbb{R}^n}$. But then by Lemma D3.3, $L_{[T]^{-1}} \circ T = 1_{\mathbb{R}^n}$. Similarly, $T \circ L_{[T]^{-1}} = 1_{\mathbb{R}^m}$. So T is invertible, with inverse $L_{[T]^{-1}}$. □

We already proved that any invertible matrix has the same number of rows and columns. This implies that the domain and codomain of an invertible linear transformation have the same dimension:

Proposition D4.7 *Let $n, m \geq 0$. If there exists an invertible linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$, then $n = m$.*

Proof If such a linear transformation T exists then by Lemma D4.6, there is an $m \times n$ invertible matrix $[T]$. But invertible matrices are square (Theorem C2.1), so $m = n$. □

For example, there is no bijective linear transformation $\mathbb{R}^2 \to \mathbb{R}^3$, even though there do exist non-bijective linear transformations $\mathbb{R}^2 \to \mathbb{R}^3$ (and, for that matter, bijective nonlinear functions $\mathbb{R}^2 \to \mathbb{R}^3$).

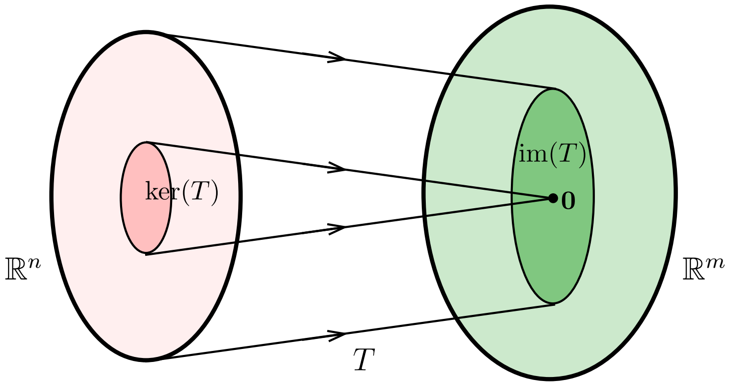


Figure D.4: The kernel and image of a linear transformation

D5 The rank of a linear transformation

We now know about the correspondence between linear transformations and matrices (Theorem D3.7), and we have a definition of the rank of a matrix (Section C1). We *could* put these two things together to get a definition of the rank of a linear transformation; that is, we could define the rank of a linear transformation to be the rank of its standard matrix.

However, despite the correspondence between linear transformations and matrices, each type of thing has its own 'feel' and brings separate insights. (Matrices are more algebraic, linear transformations more geometric.) So, we will define the rank of a linear transformation without reference to matrices, later *proving* a relationship with the rank of a matrix. And we will do similar things for the concepts of nullity and kernel.

Definition D5.1 Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

i. The **image** of T is

$$\operatorname{im}(T) = \{\mathbf{y} \in \mathbb{R}^m : \mathbf{y} = T(\mathbf{x}) \text{ for some } \mathbf{x} \in \mathbb{R}^n\}.$$

ii. The **kernel** of T is

$$\ker(T) = \{\mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = \mathbf{0}\}.$$

See Figure D.4.

Example D5.2 Let V be the subspace $\operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2\}$ of $\mathbb{R}^3$, and consider $P_V\colon \mathbb{R}^3 \to \mathbb{R}^3$, orthogonal projection onto V. We met this transformation before in Example D1.8(iii), and gave a formula for it.

The image of P_V is the set of points $\mathbf{y} \in \mathbb{R}^3$ that can be expressed as $P_V(\mathbf{x})$ for some $\mathbf{x} \in \mathbb{R}^3$. In fact, $\operatorname{im}(P_V) = V$. For on the one hand, $P_V(\mathbf{x}) \in V$ for all $\mathbf{x}$, giving $\operatorname{im}(P_V) \subseteq V$, and on the other hand, any element $\mathbf{x}$ of V is equal to $P_V(\mathbf{x})$, giving $V \subseteq \operatorname{im}(P_V)$.

The kernel of P_V is $V^\perp = \operatorname{span}(\mathbf{e}_3)$, since for $\mathbf{x} \in \mathbb{R}^3$, we have $P_V(\mathbf{x}) = \mathbf{0} \iff x_1 = x_2 = 0$.

Remark D5.3 Some people (including Poole) call the image of a linear transformation its **range**. But other people use 'range' to mean codomain, so it is less ambiguous to avoid the word altogether.

Lemma D5.4 *Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then*

$$\operatorname{im}(T) = \operatorname{col}([T]), \qquad \ker(T) = \ker([T]).$$

In other words, the image of a linear transformation is equal to the column space of its standard matrix, and the kernel of a linear transformation is equal to the kernel of its standard matrix.

Proof For both equations we use Lemma D3.3, which tells us that $T(\mathbf{x}) = [T]\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Then $\operatorname{im}(T) = \operatorname{col}([T])$ using Lemma B2.7 (applied with $A = [T]$), and $\ker(T) = \ker([T])$ immediately. □

Lemma D5.5 *Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $\operatorname{im}(T)$ is a linear subspace of $\mathbb{R}^m$ and $\ker(T)$ is a linear subspace of $\mathbb{R}^n$.*

Proof This follows from Lemma D5.4 and the facts that the column space and kernel of an $m \times n$ matrix are subspaces of $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively. □

Since the image and kernel of a linear transformation are subspaces, it makes sense to talk about their dimensions.

Definition D5.6 Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

i. The **rank** of T is $\operatorname{rank}(T) = \dim(\operatorname{im}(T))$.

ii. The **nullity** of T is $\operatorname{nullity}(T) = \dim(\ker(T))$.

Lemma D5.4 immediately implies:

Lemma D5.7 *Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $\operatorname{rank}(T) = \operatorname{rank}([T])$ and $\operatorname{nullity}(T) = \operatorname{nullity}([T])$.* □

So, as promised, the definitions of rank and nullity for linear transformations fit very well with the definitions for matrices.

Example D5.8 Consider the linear transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^4$ defined by

$$T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ 0 \\ 0 \end{pmatrix}.$$

Then

$$\operatorname{im}(T) = \left\{ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \in \mathbb{R}^4 : y_3 = y_4 = 0 \right\}$$

and

$$\ker(T) = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 : x_2 = x_3 = 0 \right\}.$$

So $\operatorname{im}(T)$ is 2-dimensional, giving $\operatorname{rank}(T) = \dim(\operatorname{im}(T)) = 2$, and $\ker(T)$ is 1-dimensional, giving $\operatorname{nullity}(T) = \dim(\ker(T)) = 1$. Another way to compute the rank and nullity is to observe that

$$[T] = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

then calculate the rank and nullity of the matrix $[T]$ using the methods of Chapter C, and finally use the fact that T has the same rank and nullity as $[T]$.

Theorem D5.9 (Rank–nullity for linear transformations) *For any linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$,*

$$\operatorname{rank}(T) + \operatorname{nullity}(T) = n.$$

Proof Since the rank and nullity of T are equal to the rank and nullity of its standard matrix $[T]$ (which has n columns), this follows from the rank-nullity theorem for matrices (Section C1). □

Remark D5.10 The rank-nullity theorem for linear transformations can be understood as follows (Figure D.4). Think of T as some kind of process. The number n is the dimension of the domain, which can loosely be thought of as 'how much stuff you start with'. The nullity of T is the dimension of the set of vectors that are mapped to $\mathbf{0}$, and can therefore be viewed as 'how much stuff you lose'. The rank of T is the dimension of the image, which is 'how much stuff you end up with'. So Theorem D5.9, rearranged as $\operatorname{rank}(T) = n - \operatorname{nullity}(T)$, states that

$$\begin{aligned}&\text{how much stuff you end up with} \\ &\qquad = \text{how much stuff you start with} \; - \; \text{how much stuff you lose.}\end{aligned}$$

This interpretation of the rank-nullity theorem is closely related to our earlier interpretation of it in terms of linear systems (page 103).

Remark D5.11 If you do Fundamentals of Pure Mathematics next semester, you will probably meet the first isomorphism theorem for groups, which implies that for a homomorphism $\theta\colon G \to H$ of finite groups, $|\operatorname{im}(\theta)| = |G|/|\ker(\theta)|$ (where the bars denote the number of elements). This is closely related to the rank-nullity theorem.

Example D5.12 The linear transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^4$ of Example D5.8 has rank 2 and nullity 1. Theorem D5.9 predicts that the domain of T has dimension $2 + 1 = 3$, which indeed it does.

In Section C2, we found many conditions on a matrix equivalent to it being invertible. Using the version of the rank-nullity theorem that we just proved, we can do something similar for linear transformations.

First, we look at some conditions on a linear transformation that are equivalent to it being injective (one-to-one), and also some conditions equivalent to it being surjective (onto).

Lemma D5.13 *Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then:*

i. T is injective $\iff \ker(T) = \{\mathbf{0}\} \iff \operatorname{nullity}(T) = 0$.

ii. T is surjective $\iff \operatorname{im}(T) = \mathbb{R}^m \iff \operatorname{rank}(T) = m$.

Proof For (i): suppose that T is injective. Certainly $\mathbf{0} \in \ker(T)$. On the other hand, if $\mathbf{x} \in \ker(T)$ then $T(\mathbf{x}) = \mathbf{0} = T(\mathbf{0})$, so by injectivity, $\mathbf{x} = \mathbf{0}$. Hence $\ker(T) = \{\mathbf{0}\}$.

Conversely, suppose that $\ker(T) = \{\mathbf{0}\}$. To prove that T is injective, let $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ with $T(\mathbf{x}) = T(\mathbf{x}')$; we must show that $\mathbf{x} = \mathbf{x}'$. By linearity,

$$T(\mathbf{x} - \mathbf{x}') = T(\mathbf{x}) - T(\mathbf{x}') = \mathbf{0},$$

so $\mathbf{x} - \mathbf{x}' \in \ker(T)$. But $\ker(T) = \{\mathbf{0}\}$, so $\mathbf{x} - \mathbf{x}' = \mathbf{0}$, so $\mathbf{x} = \mathbf{x}'$, as required.

We have shown that T is injective if and only if $\ker(T) = \{\mathbf{0}\}$. By Lemma B5.10, this is equivalent to the condition that $\operatorname{nullity}(T) = 0$.

For (ii), it is immediate from the definitions that T is surjective if and only if $\operatorname{im}(T) = \mathbb{R}^m$. By Lemma B5.10, this is equivalent to the condition that $\operatorname{rank}(T) = m$. □

We are often interested in linear transformations with the same domain and codomain. So, it is worth giving them a special name.

Definition D5.14 A **linear operator** on $\mathbb{R}^n$ is a linear transformation $\mathbb{R}^n \to \mathbb{R}^n$.

When we take $m = n$ in the last lemma, we obtain an important result about linear operators on $\mathbb{R}^n$.

Theorem D5.15 *Let T be a linear operator on $\mathbb{R}^n$. Then*

$$T \textit{ is injective} \iff T \textit{ is bijective} \iff T \textit{ is surjective.}$$

Proof By Lemma D5.13 and the rank-nullity theorem for linear transformations (Theorem D5.9),

$$T \text{ is injective} \iff \operatorname{nullity}(T) = 0 \iff \operatorname{rank}(T) = n \iff T \text{ is surjective.}$$

So if T is either injective or surjective then it is both, that is, bijective. Conversely, if T is bijective then by definition it is both injective and surjective. □

Warning D5.16 This is only true for linear transformations whose domain and codomain have the same dimension! For transformations $\mathbb{R}^n \to \mathbb{R}^m$ with $m \neq n$, it is certainly possible to be injective but not surjective (as in Example D1.4(ii)) or surjective but not injective (as in Example D1.2(iii)).

Remark D5.17 Theorem D5.15 strongly resembles Proposition B5.4. In fact, it can be deduced from that proposition (exercise). It also resembles the fact that for a finite set A, a function $f\colon A \to A$ is injective if and only if it bijective if and only if it is surjective.

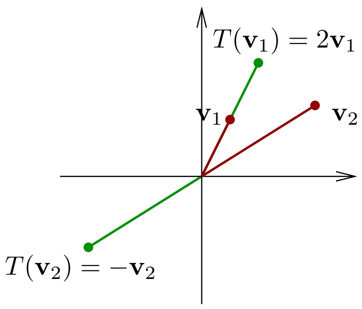

$T(\mathbf{v}_1) = 2\mathbf{v}_1$
$\mathbf{v}_1$
$\mathbf{v}_2$
$T(\mathbf{v}_2) = -\mathbf{v}_2$

Figure D.5: A linear operator T on $\mathbb{R}^2$ and its effect on non-standard basis vectors $\mathbf{v}_1$ and $\mathbf{v}_2$

D6 Change of basis

In Proposition D2.1, we saw that a linear transformation $T\colon \mathbb{R}^n \to \mathbb{R}^m$ can be described just by saying what it does to each element of a basis of $\mathbb{R}^n$. To keep things simple, let's stick to the case $n = m$, that is, linear operators on $\mathbb{R}^n$ (Definition D5.14). Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of $\mathbb{R}^n$. Proposition D2.1 tells us that we can define a linear operator T on $\mathbb{R}^n$ just by specifying vectors $T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n) \in \mathbb{R}^n$.

For example, take any basis $\mathbf{v}_1, \mathbf{v}_2$ of $\mathbb{R}^2$ (Figure D.5). Then by Proposition D2.1, there is a unique linear operator T on $\mathbb{R}^2$ such that $T(\mathbf{v}_1) = 2\mathbf{v}_1$ and $T(\mathbf{v}_2) = -\mathbf{v}_2$.

The description of T in terms of *this* basis $\mathbf{v}_1, \mathbf{v}_2$ is quite simple and convenient. But the description of T in terms of the standard basis will probably be more complicated. Maybe, for instance, $T(\mathbf{e}_1) = -1.85\mathbf{e}_1 + \frac{\pi}{4}\mathbf{e}_2$ and $T(\mathbf{e}_2) = -\sqrt{2/11}\mathbf{e}_1 + 0.79\mathbf{e}_2$. (I've just made these numbers up; the actual numbers will depend on what $\mathbf{v}_1$ and $\mathbf{v}_2$ are.) This shows that sometimes it's easier to use a non-standard basis of $\mathbb{R}^n$.

In this section, we will see how to do this. We will see that it's possible to talk about the matrix of a linear operator on $\mathbb{R}^n$ with respect to *any* basis of $\mathbb{R}^n$, not necessarily the standard one. And we will see how to convert back and forth between matrices with respect to the standard basis and matrices with respect to non-standard bases.

We begin by defining the matrix of a linear operator with respect to a basis. To warm up to the definition, we'll first think about the *standard* matrix.

Example D6.1 Let

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

and consider the linear operator $T = L_A$ on $\mathbb{R}^3$ given by $T(\mathbf{x}) = A\mathbf{x}$ ($\mathbf{x} \in \mathbb{R}^3$). Then, for instance,

$$T(\mathbf{e}_1) = \begin{pmatrix} 3 \\ -4 \\ -2 \end{pmatrix} = 3\mathbf{e}_1 - 4\mathbf{e}_2 - 2\mathbf{e}_3.$$

The coefficients here are the entries of the first column of A. This is no coincidence: A is the the standard matrix $[T]$ of T (by the correspondence in Theorem D3.7), and by definition, the jth column of $[T]$ is $T(\mathbf{e}_j)$. The same goes for $\mathbf{e}_2$ and $\mathbf{e}_3$: so

$$T(\mathbf{e}_j) = \sum_{i=1}^{3} A_{ij}\mathbf{e}_i$$

for each $j = 1, 2, 3$, where A_{ij} means the (i, j)-entry of A.

Of course, there was nothing special about this particular example. In general, for a linear operator T on $\mathbb{R}^n$ with standard matrix $A = [T]$,

$$T(\mathbf{e}_j) = \sum_{i=1}^{n} A_{ij}\mathbf{e}_i \tag{D:6}$$

for all $j \in \{1, \ldots, n\}$, simply because $T(\mathbf{e}_j)$ is the jth column of A.

All of that was about the *standard* basis. But equation (D:6) suggests how to define the matrix of an operator with respect to *any* basis:

Definition D6.2 Let T be a linear operator on $\mathbb{R}^n$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of $\mathbb{R}^n$. The **matrix of T with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_n$** is the $n \times n$ matrix $B = (B_{ij})$ defined by

$$T(\mathbf{v}_j) = \sum_{i=1}^{n} B_{ij}\mathbf{v}_i$$

for each $j \in \{1, \ldots, n\}$.

To understand the words ‘defined by’ here, recall from Lemma B4.3 that any element of $\mathbb{R}^n$ can be expressed *uniquely* as a linear combination of the basis vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. Here we are expressing $T(\mathbf{v}_j)$ as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$, and calling the coefficients $B_{1j}, \ldots, B_{nj}$.

Examples D6.3 i. The matrix of a linear operator with respect to the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is simply its standard matrix, by equation (D:6). Example D6.1 gives a particular example of this.

ii. Let $\mathbf{v}_1, \mathbf{v}_2$ be any basis of $\mathbb{R}^2$, and let T be the unique linear operator on $\mathbb{R}^2$ such that $T(\mathbf{v}_1) = 2\mathbf{v}_1$ and $T(\mathbf{v}_2) = -\mathbf{v}_2$ (as in Figure D.5). What is the matrix of T with respect to the basis $\mathbf{v}_1, \mathbf{v}_2$?

Call this matrix B. We have

$$T(\mathbf{v}_1) = 2\mathbf{v}_1 + 0\mathbf{v}_2,$$

so the first column of B is $\left(\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right)$. Similarly,

$$T(\mathbf{v}_2) = 0\mathbf{v}_1 + (-1)\mathbf{v}_2,$$

so the second column of B is $\left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right)$. Hence

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

iii. More generally, take any basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of $\mathbb{R}^n$, take any scalars $c_1, \ldots, c_n$, and consider the linear operator T on $\mathbb{R}^n$ defined by

$$T(\mathbf{v}_1) = c_1\mathbf{v}_1,\ T(\mathbf{v}_2) = c_2\mathbf{v}_2,\ \ldots,\ T(\mathbf{v}_n) = c_n\mathbf{v}_n$$

(as in Example D2.2). Then the matrix of T with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is

$$\begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_n \end{pmatrix}.$$

This is an example of a **diagonal** matrix, that is, a square matrix M such that $M_{ij} = 0$ for all i, j such that $i \neq j$. Diagonal matrices are extremely easy to work with. So, given a linear operator T on $\mathbb{R}^n$, an important question is whether we can find a basis of $\mathbb{R}^n$ such that the matrix of T with respect to that basis is diagonal. If we can, everything becomes much simpler. We'll explore this question in the final chapter of the course.

Let's come back to the operator T of Figure D.5. We know that whatever $\mathbf{v}_1$ and $\mathbf{v}_2$ may be, the matrix of T with respect to the basis $\mathbf{v}_1, \mathbf{v}_2$ is $\left(\begin{smallmatrix} 2 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. But what is its *standard* matrix? The answer depends on what the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ actually are, and is given by the main result of this section:

Theorem D6.4 (Change of basis) *Let T be a linear operator on $\mathbb{R}^n$ and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of $\mathbb{R}^n$. Write:*

- *A for the standard matrix of T;*
- *B for the matrix of T with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_n$;*
- *P for the matrix $(\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_n)$.*

Then P is invertible, $A = PBP^{-1}$, and $B = P^{-1}AP$.

This matrix P is called a **change of basis** matrix.

Proof The columns of the $n \times n$ matrix P form a basis of $\mathbb{R}^n$, so by Theorem C2.3, P is invertible. Both equations will follow if we can prove that $AP = PB$.

We prove this by evaluating $T(\mathbf{v}_j)$ in two ways, for each $j \in \{1, \ldots n\}$. On the one hand,

$$\begin{aligned} T(\mathbf{v}_j) &= \sum_{i=1}^{n} B_{ij}\mathbf{v}_i && \text{by definition of } B \\ &= \sum_{i=1}^{n} B_{ij} \sum_{i'=1}^{n} P_{i'i}\mathbf{e}_{i'} && \text{by definition of } P \\ &= \sum_{i'=1}^{n} \left(\sum_{i=1}^{n} P_{i'i}B_{ij} \right) \mathbf{e}_{i'} && \text{by changing the order of summation} \\ &= \sum_{i'=1}^{n} (PB)_{i'j}\mathbf{e}_{i'} && \text{by definition of matrix multiplication.} \end{aligned}$$

On the other hand,

$$
\begin{aligned}
T(\mathbf{v}_j) &= T\left(\sum_{j'=1}^{n} P_{j'j}\mathbf{e}_{j'}\right) && \text{by definition of } P \\
&= \sum_{j'=1}^{n} P_{j'j} T(\mathbf{e}_{j'}) && \text{by linearity of } T \\
&= \sum_{j'=1}^{n} P_{j'j} \sum_{i'=1}^{n} A_{i'j'}\mathbf{e}_{i'} && \text{by definition of } A \\
&= \sum_{i'=1}^{n} \left(\sum_{j'=1}^{n} A_{i'j'}P_{j'j}\right)\mathbf{e}_{i'} && \text{by changing the order of summation} \\
&= \sum_{i'=1}^{n} (AP)_{i'j}\mathbf{e}_{i'} && \text{by definition of matrix multiplication.}
\end{aligned}
$$

Comparing the two expressions for $T(\mathbf{v}_j)$ gives

$$\sum_{i'=1}^{n} (PB)_{i'j}\mathbf{e}_{i'} = \sum_{i'=1}^{n} (AP)_{i'j}\mathbf{e}_{i'}$$

for all j. It follows that $(PB)_{i'j} = (AP)_{i'j}$ for all i' and j; hence $PB = AP$, as required. □

Example D6.5 Let's use this result to find the standard matrix of the operator T of Figure D.5, taking (for instance) $\mathbf{v}_1 = \binom{3}{1}$ and $\mathbf{v}_2 = \binom{7}{4}$. We have already seen in Example D6.3(ii) that the matrix of T with respect to the basis $\mathbf{v}_1, \mathbf{v}_2$ is

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

The change of basis matrix P is

$$P = (\mathbf{v}_1|\mathbf{v}_2) = \begin{pmatrix} 3 & 7 \\ 1 & 4 \end{pmatrix}.$$

Theorem D6.4 implies that the standard matrix $A = [T]$ is given by

$$A = PBP^{-1} = \begin{pmatrix} 3 & 7 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 3 & 7 \\ 1 & 4 \end{pmatrix}^{-1} = \frac{1}{5}\begin{pmatrix} 31 & -63 \\ 12 & -26 \end{pmatrix}.$$

You can check this answer by calculating directly that when you multiply this matrix by $\mathbf{v}_1$, you get $2\mathbf{v}_1$, and when you multiply it by $\mathbf{v}_2$, you get $-\mathbf{v}_2$.

Remark D6.6 You can think of a basis as a kind of language, and a change of basis matrix as a kind of dictionary, as follows.

Suppose I give you the Catalan sentence

Veig el vostre veí

and ask you to transform it from the present to the past tense. Unless you speak Catalan, you're stuck; but suppose I also allow you to use a piece of software that translates between Catalan and English. Then you can solve the problem like this. First use the software to translate the original sentence from Catalan to English. It outputs

I see your neighbour.

Then transform it into the past tense:

I saw your neighbour.

Then use the translation software again to put it back into Catalan, giving the final result:

Vaig veure el vostre veí.

And that solves the problem. In a diagram:

$$\begin{tikzcd}
\textit{Veig el vostre veí} \arrow[r, mapsto, "\text{translate}"] \arrow[d, dotted, mapsto, "\text{transform}"'] & \textit{I see your neighbour} \arrow[d, mapsto, "\text{transform}"] \\
\textit{Vaig veure el vostre veí} & \textit{I saw your neighbour} \arrow[l, mapsto, "\text{translate}"]
\end{tikzcd}$$

Think of the standard basis as like English, the other basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ as like Catalan, the change of basis matrix P as like translation from Catalan into English, and the linear transformation as like transforming from the present to the past tense:

$$\begin{tikzcd}
\mathbf{x} \arrow[r, mapsto, "P"] \arrow[d, dotted, mapsto, "B"'] & P\mathbf{x} \arrow[d, mapsto, "A"] \\
B\mathbf{x} = P^{-1}AP\mathbf{x} & AP\mathbf{x} \arrow[l, mapsto, "P^{-1}"]
\end{tikzcd}$$

Thus, the equation $B = P^{-1}AP$ is like the three-step method for transforming a Catalan sentence from the present to the past tense.

Theorem D6.4 tells us that matrices of the same linear operator with respect to different bases are related in a certain way. It is useful to have some terminology for this.

Definition D6.7 Let A and B be $n \times n$ matrices. We say that A is **similar** to B (or 'A and B are similar'), and write $A \sim B$, if there exists an invertible $n \times n$ matrix P such that $A = PBP^{-1}$.

Examples D6.8 i. The matrices

$$A = \frac{1}{5}\begin{pmatrix} 31 & -63 \\ 12 & -26 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$$

are similar, because we showed in Example D6.5 that $A = PBP^{-1}$ for a certain invertible matrix P.

ii. The zero matrix $0 = \left(\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is not similar to any matrix apart from itself, since $P0P^{-1} = 0$ for every invertible P.

The next lemma says that the word 'similar' behaves in a sensible way: if A is similar to B then B is similar to A, and so on.

Lemma D6.9 *i. Let A be an $n \times n$ matrix. Then $A \sim A$.*

ii. Let A and B be $n \times n$ matrices. If $A \sim B$ then $B \sim A$.

iii. Let A, B and C be $n \times n$ matrices. If $A \sim B$ and $B \sim C$, then $A \sim C$.

In the jargon that you probably either met recently or will meet soon, these three conditions say that similarity is an **equivalence relation** on the set of all $n \times n$ matrices.

Proof For (i), we have $A = IAI^{-1}$, so $A \sim A$.

For (ii), suppose that $A \sim B$; then we can choose an invertible matrix P such that $A = PBP^{-1}$. Put $Q = P^{-1}$. Then Q is invertible and $B = P^{-1}AP = QAQ^{-1}$, so $B \sim A$.

For (iii), suppose that $A \sim B$ and $B \sim C$. Then we can choose invertible matrices P and Q such that $A = PBP^{-1}$ and $B = QCQ^{-1}$. By Lemma A5.5, PQ is invertible with inverse $Q^{-1}P^{-1}$, giving

$$A = P(QCQ^{-1})P^{-1} = (PQ)C(PQ)^{-1}.$$

Hence $A \sim C$. □

Proposition D6.10 *Let T be a linear operator on $\mathbb{R}^n$, and let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and $\mathbf{v}'_1, \ldots, \mathbf{v}'_n$ be bases of $\mathbb{R}^n$. Then the matrix of T with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is similar to the matrix of T with respect to $\mathbf{v}'_1, \ldots, \mathbf{v}'_n$.*

Proof Write B for the matrix of T with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_n$ and B' for the matrix of T with respect to $\mathbf{v}'_1, \ldots, \mathbf{v}'_n$. Write A for the standard matrix of T. Theorem D6.4 implies that both $A \sim B$ and $A \sim B'$. Then $B \sim A$ by Lemma D6.9(ii), so $B \sim B'$ by Lemma D6.9(iii). □

Looking at matrices of the same linear transformation with respect to different bases is something like looking at a sculpture from different angles. It's the same sculpture, but you see different aspects of it from different angles, and understanding what you're seeing may be easier from some angles than others. In the same way, the linear transformation T of Example D6.5 is much easier to understand (and has a much simpler matrix) from the point of view of the non-standard basis $\binom{3}{1}, \binom{7}{4}$ than from the point of view of the standard basis.

D7 The determinant of a linear operator

We have seen that linear transformations $\mathbb{R}^n \to \mathbb{R}^m$ correspond to $m \times n$ matrices. What about the special case $m = n$? For linear transformations, taking $m = n$ means considering only linear *operators*. For matrices, taking $m = n$ means considering only *square* matrices. So, transformations correspond to matrices and operators correspond to square matrices.

Some of the things we do with matrices require them to be square. Taking the determinant is one of those things. We will see that there is a corresponding definition of the determinant of a linear operator. To get to the definition, we need a lemma:

Lemma D7.1 *Similar matrices have the same determinant.*

Proof Let A and B be similar matrices. Then $A = PBP^{-1}$ for some invertible matrix P. Recall from Proposition C3.4(vi) that $\det(XY) = \det(X)\det(Y)$ for any $n \times n$ matrices X and Y, and from Corollary C3.6 that if X is an invertible matrix then $\det(X^{-1}) = 1/\det X$. It follows that

$$\det(A) = \det(P)\det(B)\det(P^{-1}) = \det(P)\det(B)/\det(P) = \det(B). \qquad \square$$

Let T be a linear operator on $\mathbb{R}^n$. By Proposition D6.10 and the lemma we just proved, the matrix of T with respect to *any* basis always has the same determinant, *no matter which basis we choose*. We can therefore define the **determinant** $\det(T)$ to be the determinant of any of these matrices.

Example D7.2 Consider again the linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ of Example D6.5. The standard matrix of T is $\frac{1}{5}\left(\begin{smallmatrix} 31 & -63 \\ 12 & -26 \end{smallmatrix}\right)$, and its matrix with respect to a certain non-standard basis is $\left(\begin{smallmatrix} 2 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Our results guarantee (and you can check directly) that these two matrices have the same determinant. Of course, it's easier to calculate the second determinant, which is -2. So by definition, $\det(T) = -2$.

The concept of linear operator sheds light on the idea of determinant. I introduced determinants in terms of area and volume (Section C3). We saw that for a matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, the area of the parallelogram G with edges parallel to $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$ is $|\det(A)|$. This parallelogram G is given by

$$G = \left\{x_1\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right) + x_2\left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right) : 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\right\}.$$

On the other hand, the unit square S is given by

$$S = \{x_1\mathbf{e}_1 + x_2\mathbf{e}_2 : 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\}.$$

It follows that the linear transformation $L_A\colon \mathbb{R}^2 \to \mathbb{R}^2$ maps the unit square S onto the parallelogram G: for

$$\begin{aligned} \{L_A(\mathbf{x}) : \mathbf{x} \in S\} &= \{A(x_1\mathbf{e}_1 + x_2\mathbf{e}_2) : 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\} \\ &= \{x_1(A\mathbf{e}_1) + x_2(A\mathbf{e}_2) : 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\} \\ &= G \end{aligned}$$

where in the last step we used the fact that $A\mathbf{e}_1$ and $A\mathbf{e}_2$ are the columns of A.

Here's the point: the unit square S has area 1, and the parallelogram G has area $|\det(A)|$, which is equal to $|\det(L_A)|$ (by definition of the latter). So, applying L_A to S multiplies its area by a factor of $|\det(L_A)|$.

In fact, it can be shown that for *any* sensible subset H of $\mathbb{R}^2$, and for any linear transformation $T\colon \mathbb{R}^2 \to \mathbb{R}^2$, the area of $\{T(\mathbf{x}) : \mathbf{x} \in H\}$ is the area of H multiplied by $|\det(T)|$. ('Sensible' just means that it is possible to measure the area of H.) And similar results hold in $\mathbb{R}^3$ for volume instead of area—and even in higher dimensions, once a higher-dimensional notion of 'volume' has been defined. This principle is crucial to computing multi-dimensional integrals, as you may have discovered already in SVCDE.

In a slogan:

Determinant is volume scale factor.

More accurately, the *absolute value* of determinant is volume scale factor. In other words, applying a linear transformation T to a shape multiplies its volume by $|\det(T)|$.

In questions about volume scale factors, it's useful to know that in mathematics, the word **circle** always refers to a hollow figure (such as $\{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$) and the word **disc** is used for a circle with the inside filled in (such as $\{\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$). Similarly, a **sphere** is always hollow and the solid, filled-in shape is called a **ball**.

Finally, a linear operator is invertible unless its volume scale factor is zero:

Lemma D7.3 *A linear operator T is invertible if and only if* $\det(T) \neq 0$.

Proof T is invertible if and only if its standard matrix $[T]$ is invertible (by Lemma D4.6), if and only if $\det([T]) \neq 0$ (by Theorem C3.9). But $\det(T) = \det([T])$ by definition. □

Intuitively, then, an operator is invertible unless it crushes something down to nothing. This fits with the fact that an operator is invertible unless its kernel is nontrivial (a consequence of Lemma D5.13 and Theorem D5.15).

Next time: we learn that many features of a linear operator are encapsulated by certain scalars called its 'eigenvalues'.

Summary of Chapter D

This is for you to fill in.

The most important definitions and ideas in this chapter

The most important results in this chapter

Points I didn't understand

Chapter E

Eigenvalues and eigenvectors

To be read before the lecture of Monday, 19 November 2018

The last chapter was about linear transformations $T\colon \mathbb{R}^n \to \mathbb{R}^m$, where the domain $\mathbb{R}^n$ and codomain $\mathbb{R}^m$ could have different dimensions. But there are especially interesting things to say about linear *operators*, which are linear transformations whose domain and codomain are the same (Definition D5.14). In terms of matrices, these are the square ones.

Here are some of the things you can do with a linear operator that you can't do with a linear transformation in general:

- Iterate it—that is, do it repeatedly. If T is a linear operator on $\mathbb{R}^n$ then the composites $T^2 = T \circ T$ ('do T twice'), $T^3 = T \circ T \circ T$, etc., all make sense. These composites only exist because T has the same domain and codomain.
- Look for an inverse. (When $n \neq m$, we *can* ask whether a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ has an inverse, but Proposition D4.7 says that the answer is always no.)
- Take its determinant (as defined in Section D7).
- Ask which points $\mathbf{x}$ of $\mathbb{R}^n$ are 'fixed' by T, in the sense that $T(\mathbf{x}) = \mathbf{x}$.
- Ask whether it is 'diagonalizable'—that is, whether its matrix with respect to some basis of $\mathbb{R}^n$ is diagonal. We investigate this in Sections E3 and E6.

We'll see that when we study linear operators (or square matrices), a crucial role is played by the so-called eigenvalues and eigenvectors. They enable us to solve certain problems very easily. For instance, let

$$A = \begin{pmatrix} 14 & -14 & -16 \\ -14 & 23 & -2 \\ -16 & -2 & 8 \end{pmatrix}.$$

Can you find a formula for A^r, for an arbitrary positive integer r? Eigenvalues and eigenvectors will make this problem easy (Example E6.8). Or, consider the following problem (which we'll come back to):

> Wolves can be classified as either young or adult (capable of reproducing). In a certain wolf community, the following happens every year: 60% of the young wolves die, 10% of the adults die, 20% of the young become adults, and new young are born at a rate of 3.9 individuals per adult. What is the ratio of young wolves to adult wolves in the long run?

On the face of it, this has nothing to do with linear algebra. Yet the easiest way of solving it also uses eigenvalues and eigenvectors.

E1 Definitions and examples

Here we meet the definitions of eigenvalue and eigenvector. We approach them via some examples.

Examples E1.1 i. Consider reflection of the plane in the x-axis, which is a linear operator T on $\mathbb{R}^2$ (Figure D.1). When we are using this operator, are any points of $\mathbb{R}^2$ more special or interesting than the others?

Obviously a special role is played by the x-axis: it's the axis of reflection! More precisely, the x-axis is exactly the set of vectors $\mathbf{x}$ such that $T(\mathbf{x}) = \mathbf{x}$. A point $\mathbf{x}$ such that $T(\mathbf{x}) = \mathbf{x}$ is called a **fixed point** of T; so in this case, the fixed points of T are exactly the points on the x-axis.

A little less obviously, the y-axis also plays a special role for T, since it has the special property that when it is reflected in the x-axis, it is mapped onto itself—but flipped around. This isn't true for any other line. We can make this precise by observing that the y-axis is exactly the set of vectors $\mathbf{x}$ such that $T(\mathbf{x}) = -\mathbf{x}$.

So in summary, for this particular linear operator T on $\mathbb{R}^2$, there are two especially interesting types of vector: those satisfying $T(\mathbf{x}) = \mathbf{x}$, and those satisfying $T(\mathbf{x}) = -\mathbf{x}$.

ii. Now let V be any subspace of $\mathbb{R}^n$ and consider P_V, the operator on $\mathbb{R}^n$ of orthogonal projection onto V (Definition D1.5). Think of the case $n = 3$ and $V = \text{span}\{\mathbf{e}_1, \mathbf{e}_2\}$ if you like.

Again, let us ask: when we are applying this operator P_V, which vectors are particularly special or interesting?

The subspace V is certainly interesting, since that's what we're projecting onto. The definition of P_V immediately implies that a vector $\mathbf{x} \in \mathbb{R}^n$ belongs to V if and only if $P_V(\mathbf{x}) = \mathbf{x}$.

On the other hand, the definition of P_V also involves the orthogonal complement $V^\perp$ of V. And again, it's an easy consequence of the definition of P_V that a vector $\mathbf{x} \in \mathbb{R}^n$ belongs to $V^\perp$ if and only if $P_V(\mathbf{x}) = \mathbf{0}$.

So in summary, for this particular linear operator P_V on $\mathbb{R}^n$, there are two especially interesting types of vector: those satisfying $P_V(\mathbf{x}) = \mathbf{x}$, and those satisfying $P_V(\mathbf{x}) = \mathbf{0}$.

iii. Generally, for any linear operator T on $\mathbb{R}^n$, we've seen that an important role is played by the kernel of T, which is the set of vectors $\mathbf{x}$ satisfying $T(\mathbf{x}) = \mathbf{0}$.

Bringing these three examples together, we see that for a linear operator T on $\mathbb{R}^n$, a special role is played by the vectors $\mathbf{x}$ satisfying $T(\mathbf{x}) = 1\mathbf{x}$ or $T(\mathbf{x}) = -1\mathbf{x}$ or $T(\mathbf{x}) = 0\mathbf{x}$. In other examples, a special role is played by the vectors $\mathbf{x}$ satisfying $T(\mathbf{x}) = \lambda\mathbf{x}$ for other scalars λ. This leads to the following definition.

Definition E1.2 Let T be a linear operator on $\mathbb{R}^n$.

i. An **eigenvalue** of T is a real number λ such that $T(\mathbf{x}) = \lambda\mathbf{x}$ for some vector $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{R}^n$.

ii. An **eigenvector** of T with eigenvalue λ is a vector $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{R}^n$ such that $T(\mathbf{x}) = \lambda\mathbf{x}$.

In other words, an eigenvector of T is a nonzero vector $\mathbf{x}$ such that $T(\mathbf{x})$ is a scalar multiple of $\mathbf{x}$. Figure D.5 (page 126) shows an eigenvector $\mathbf{v}_1$ with eigenvalue 2 and an eigenvector $\mathbf{v}_2$ with eigenvalue -1.

Warning E1.3 By definition, eigenvectors are nonzero. The zero vector $\mathbf{0}$ is not counted as an eigenvector. Why not? Because by definition of linearity, $T(\mathbf{0}) = \lambda\mathbf{0}$ for *all* real λ. If we allowed $\mathbf{0}$ as an eigenvector, then every real number would be an eigenvalue. This would not be a very useful definition!

Examples E1.4 i. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be reflection in the x-axis. Then $T\binom{x}{0} = \binom{x}{0}$ for all $x \in \mathbb{R}$. Hence 1 is an eigenvalue of T, and $\binom{x}{0}$ is an eigenvector of T with eigenvalue 1 for all real $x \neq 0$.

Also, $T\binom{0}{y} = -\binom{0}{y}$ for all $x \in \mathbb{R}$. Hence -1 is also an eigenvalue of T, and $\binom{0}{y}$ is an eigenvector of T with eigenvalue -1 for all real $y \neq 0$.

In fact, ±1 are the *only* eigenvalues of T. Later, we'll prove a general result that allows us to conclude this immediately. For now, you should be able to see intuitively that if $\mathbf{x}$ is a vector such that $T(\mathbf{x})$ is a scalar multiple of $\mathbf{x}$ then $\mathbf{x}$ is either on the x-axis or on the y-axis.

ii. Similarly, let V be any 2-dimensional subspace of $\mathbb{R}^3$ (plane through $\mathbf{0}$), and let $T\colon \mathbb{R}^3 \to \mathbb{R}^3$ be reflection in V. Then every nonzero point on the plane V is an eigenvector with eigenvalue 1, and every nonzero point on the line $V^\perp$ through the origin perpendicular to V is an eigenvector with eigenvalue -1. These are the only two eigenvalues of T.

iii. Let V be any subspace of $\mathbb{R}^n$, and consider orthogonal projection $P_V : \mathbb{R}^n \to \mathbb{R}^n$. Then for the reasons explained in Example E1.1(ii), every nonzero vector in V is an eigenvector with eigenvalue 1, and every nonzero vector in $V^\perp$ is an eigenvector with eigenvalue 0. In fact, it can be shown that there are no other eigenvalues or eigenvectors.

iv. Consider the linear operator $R_{\pi/3}$ on $\mathbb{R}^2$ (rotation by $\pi/3$ about $\mathbf{0}$). There are *no* vectors $\mathbf{x} \in \mathbb{R}^2$ such that $R_{\pi/3}(\mathbf{x})$ is a scalar multiple of $\mathbf{x}$, apart from $\mathbf{0}$. Hence $R_{\pi/3}$ has no eigenvalues or eigenvectors at all.

The same is true of R_θ for any other value of θ except integer multiples of π. (Exercise: why did I exclude those values?)

v. Define a linear operator T on $\mathbb{R}^n$ by $T(\mathbf{x}) = -7\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. Then every nonzero vector is an eigenvector with eigenvalue -7.

Remark E1.5 Eigen*vectors* are not allowed to be zero, but eigen*values* can be. In fact, eigenvectors with eigenvalue 0 are very useful: for a linear operator T on $\mathbb{R}^n$, the eigenvectors of T with eigenvalue 0 are exactly the nonzero elements of the kernel of T. This is because $T(\mathbf{x}) = 0\mathbf{x} \iff T(\mathbf{x}) = \mathbf{0}$.

Proposition E1.6 *Let T be an operator on $\mathbb{R}^n$. Then 0 is an eigenvalue of T if and only if T is not invertible.*

Proof By Remark E1.5, 0 is an eigenvalue of T if and only if $\ker(T) \neq \{\mathbf{0}\}$. By Lemma D5.13(i), this is equivalent to T not being injective, which by Theorem D5.15 is equivalent to T not being invertible. □

Geometrically, an eigenvector of an operator T is *roughly speaking* a line unmoved by T. For instance, when $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ is reflection in the x-axis, both the x-axis and the y-axis are unmoved by T. But I say 'roughly speaking' for three important reasons.

First, by 'unmoved' I mean that the line is mapped into itself by T, not that individual points on it are fixed by T. If $\mathbf{x}$ is an eigenvector with eigenvalue λ then the line $\operatorname{span}\{\mathbf{x}\}$ is mapped into itself by T, scaling by a factor of λ, since

$$T(c\mathbf{x}) = cT(\mathbf{x}) = c\lambda\mathbf{x} = \lambda(c\mathbf{x})$$

for all $c \in \mathbb{R}$. When $\lambda < 0$, this means that the line is flipped around. For instance, this is the case for the y-axis in the reflection example above.

Second, things look a bit different when $\lambda = 0$. In that case, the line spanned by an eigenvector $\mathbf{x}$ is still mapped into itself by T, but everything on the line is mapped to $\mathbf{0}$.

Finally, if $\mathbf{x}$ is an eigenvector with eigenvalue λ then so is $c\mathbf{x}$ for each real $c \neq 0$, since $T(c\mathbf{x}) = \lambda(c\mathbf{x})$. So, it's not quite right to say that an eigenvector is a line unmoved by T, since different eigenvectors can all span the same line (e.g. $\mathbf{x}$, $3\mathbf{x}$ and $-2\mathbf{x}$). But it gives some useful intuition.

A major theme of the last chapter was that we can translate back and forth between linear transformations and matrices (Theorem D3.7). For linear operators, this means the following. In one direction, every linear operator T on $\mathbb{R}^n$ has a standard matrix $[T]$ (an $n \times n$ matrix). In the other direction, every $n \times n$ matrix A gives rise to a linear operator L_A on $\mathbb{R}^n$, defined by $L_A(\mathbf{x}) = A\mathbf{x}$ ($\mathbf{x} \in \mathbb{R}^n$).

We now take the definitions of eigenvalue and eigenvector for linear operators and translate them into the language of matrices. An eigenvalue or eigenvector of a matrix A is exactly an eigenvalue or eigenvector of the linear operator L_A. Explicitly, this means the following:

Definition E1.7 Let A be an $n \times n$ real matrix.

i. An **eigenvalue** of A is a real number λ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some vector $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{R}^n$.

ii. An **eigenvector** of A with eigenvalue λ is a vector $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{R}^n$ such that $A\mathbf{x} = \lambda\mathbf{x}$.

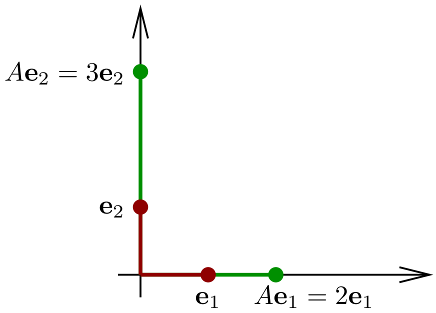

$A\mathbf{e}_2 = 3\mathbf{e}_2$
$\mathbf{e}_2$
$\mathbf{e}_1$
$A\mathbf{e}_1 = 2\mathbf{e}_1$

Figure E.1: Eigenvectors and eigenvalues of the 2×2 matrix in Example E1.8(ii)

Examples E1.8 i. Let $A = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$. Then

$$A\begin{pmatrix}x\\0\end{pmatrix} = \begin{pmatrix}x\\0\end{pmatrix}, \qquad A\begin{pmatrix}0\\y\end{pmatrix} = -\begin{pmatrix}0\\y\end{pmatrix}$$

$(x, y \in \mathbb{R})$. Hence $\left(\begin{smallmatrix}x\\0\end{smallmatrix}\right)$ is an eigenvector of T with eigenvalue 1 for any real $x \neq 0$, and $\left(\begin{smallmatrix}0\\y\end{smallmatrix}\right)$ is an eigenvector of T with eigenvalue -1 for any real $y \neq 0$.

ii. Let A be a diagonal matrix, that is, a matrix of the form

$$A = \begin{pmatrix} c_1 & 0 & \cdots & 0\\ 0 & c_2 & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & c_n\end{pmatrix}$$

where $c_1, \dots, c_n \in \mathbb{R}$. Then $A\mathbf{e}_j = c_j\mathbf{e}_j$ for each $j \in \{1, \dots, n\}$ (since by Lemma A4.3(ii), $A\mathbf{e}_j$ is the jth column of A). Hence $\mathbf{e}_j$ is an eigenvector of A, with eigenvalue c_j, for each j. Figure E.1 shows the case $A = \left(\begin{smallmatrix}2 & 0\\ 0 & 3\end{smallmatrix}\right)$.

Recall from Lemma D3.3 that when $T\colon \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, $T(\mathbf{x}) = [T]\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. It follows that the eigenvalues and eigenvectors of a linear operator T are exactly the eigenvalues and eigenvectors of its standard matrix $[T]$. So when reasoning with eigenvalues and eigenvectors, it makes little difference whether we work with linear operators or square matrices.

Example E1.9 The linear operator T on $\mathbb{R}^2$ defined by reflection in the x-axis has standard matrix $[T] = \left(\begin{smallmatrix}1 & 0\\ 0 & -1\end{smallmatrix}\right)$ (Example D2.4(i)). We already saw in Example E1.4(i) that 1 and -1 are eigenvalues of the operator T, so they are also eigenvalues of the matrix $[T]$; moreover, the eigenvectors are the same. In fact, $[T]$ is exactly the matrix A of Example E1.8(i), so we have just confirmed what we found there.

It is useful to think about the set of all the eigenvectors of T that share a particular eigenvalue:

Definition E1.10 Let T be a linear operator on $\mathbb{R}^n$, and let $\lambda \in \mathbb{R}$. The **λ-eigenspace** of T is

$$E_\lambda(T) = \{\mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = \lambda\mathbf{x}\}.$$

Similarly, for an $n \times n$ real matrix A, the **λ-eigenspace** of A is

$$E_\lambda(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \lambda\mathbf{x}\}.$$

In other words, the λ-eigenspace of a matrix or operator is the set of all eigenvectors with eigenvalue λ, together with the zero vector $\mathbf{0}$.

Example E1.11 Let V be a plane through $\mathbf{0}$ in $\mathbb{R}^3$ and let $T \colon \mathbb{R}^3 \to \mathbb{R}^3$ be reflection in V, as in Example E1.4(ii). Then the 1-eigenspace of T is V, and the (-1)-eigenspace of T is the line $V^\perp$. For $\lambda \neq \pm 1$, the λ-eigenspace of T is $\{\mathbf{0}\}$, because ± 1 are the only eigenvalues. (We will see how to prove this in the next section.)

The definitions of eigenspace for operators and matrices are really the same, in the sense that the λ-eigenspace of an operator is equal to the λ-eigenspace of its standard matrix. The next few results could equally well be presented in terms of matrices or linear operators. We will mostly present them in the language of matrices.

First, eigenspaces are closely related to kernels. Evidently $E_0(A) = \ker(A)$. More generally:

Lemma E1.12 *Let A be an $n \times n$ matrix, and let λ be a scalar. Then $E_\lambda(A) = \ker(A - \lambda I)$.*

Proof For $\mathbf{x} \in \mathbb{R}^n$, we have

$$\begin{aligned} \mathbf{x} \in E_\lambda(A) &\iff A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0} \\ &\iff (A - \lambda I)\mathbf{x} = \mathbf{0} \\ &\iff \mathbf{x} \in \ker(A - \lambda I). \end{aligned} \qquad \square$$

We have already seen that if $\mathbf{x}$ is an eigenvector for some matrix or operator, with eigenvalue λ, then so is every scalar multiple of $\mathbf{x}$ (apart from $\mathbf{0}$). So every eigenspace is closed under scalar multiplication. Better still:

Lemma E1.13 *Every eigenspace of an $n \times n$ square matrix, or of a linear operator on $\mathbb{R}^n$, is a linear subspace of $\mathbb{R}^n$.*

Proof Let A be an $n \times n$ matrix and $\lambda \in \mathbb{R}$. Then $E_\lambda(A) = \ker(A - \lambda I)$, which by Lemma B1.4 is a linear subspace of $\mathbb{R}^n$. This proves the result on matrices. The result follows for operators T, since $E_\lambda(T) = E_\lambda([T])$. □

We defined the λ-eigenspace for *all* scalars λ, even if λ is not an eigenvalue. If λ is not an eigenvalue then $E_\lambda(A)$ is the trivial subspace $\{\mathbf{0}\}$. The eigenvalues of A are exactly those scalars λ such that $E_\lambda(A)$ is nontrivial:

Proposition E1.14 *Let A be an $n \times n$ square matrix and $\lambda \in \mathbb{R}$. The following are equivalent:*

i. λ is an eigenvalue of A;

ii. the λ-eigenspace $E_\lambda(A)$ is nontrivial (that is, not equal to $\{\mathbf{0}\}$);

iii. $A - \lambda I$ is not invertible;

iv. $\det(A - \lambda I) = 0$.

Proof (i)⇔(ii): by definition of eigenvalue, λ is an eigenvalue of A if and only if $E_\lambda(A)$ contains some nonzero vector, or in other words, is nontrivial.

(ii)⇔(iii): $E_\lambda(A) = \ker(A - \lambda I)$, so (ii) states that $\ker(A - \lambda I)$ is nontrivial, which by Theorem C2.3 is equivalent to $A - \lambda I$ not being invertible.

(iii)⇔(iv) follows immediately from Theorem C3.9. □

Corollary E1.15 (Equivalent conditions for invertibility, part 5) *Let A be a square matrix. Then A is invertible if and only if 0 is not an eigenvalue of A.*

Proof Take $\lambda = 0$ in (i)⇔(iii) of Proposition E1.14. □

Remark E1.16 Proposition E1.6 and Corollary E1.15 say essentially the same thing: an operator or square matrix is invertible if and only if it does not have 0 as an eigenvalue. Either of these results could be deduced from the other by using the correspondence between operators and square matrices.

E2 The characteristic polynomial

Imagine that someone hands you a specific square matrix. How would you find its eigenvalues and eigenvectors? For some of the examples in the previous sections, we had a simple geometric description of the corresponding linear operator, so we were able to use our intuition to guess. But what if we don't have a geometric description?

This section provides a method for computing eigenvalues and eigenvectors of any square matrix. The key is Proposition E1.14, where we saw that λ is an eigenvalue for a square matrix A if and only if $\det(A - \lambda I) = 0$.

Lemma E2.1 *Let A be an $n \times n$ matrix. Then $\det(A - \lambda I)$ is a polynomial in λ of degree n, with leading coefficient $(-1)^n$.*

Proof Omitted (but not hard). □

This polynomial has a name:

Definition E2.2 Let A be a square matrix. The **characteristic polynomial** of A is the polynomial $\chi_A(\lambda) = \det(A - \lambda I)$.

Proposition E1.14 immediately implies:

Proposition E2.3 *Let A be a real square matrix. Then the eigenvalues of A are exactly the real roots of its characteristic polynomial χ_A.* □

Examples E2.4 i. Once again, consider reflection of $\mathbb{R}^2$ in the x-axis (Examples E1.1(i), E1.4(i) and E1.9). Its standard matrix A is $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, and

$$\chi_A(\lambda) = \det\begin{pmatrix} 1-\lambda & 0 \\ 0 & -1-\lambda \end{pmatrix} = (1-\lambda)(-1-\lambda) = (\lambda - 1)(\lambda + 1).$$

So the roots of χ_A are ± 1. By Proposition E2.3, the eigenvalues of A are exactly 1 and -1; there are no others. (Previously, we had concluded this by a geometric argument.)

We can also confirm our previous statements about the eigenspaces of A. The 1-eigenspace is

$$E_1(A) = \ker(A - 1I) = \ker \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\}.$$

So as expected, the 1-eigenspace is exactly the x-axis, and the eigenvectors with eigenvalue 1 are all the points $\left(\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right)$ with $x \neq 0$. A similar calculation tells us that $E_{-1}(A)$ is exactly the y-axis.

ii. In Example E1.4(iv), we argued geometrically that the rotation operator $R_{\pi/3}$ on $\mathbb{R}^2$ has no eigenvalues. Here is algebraic confirmation. The standard matrix of $R_{\pi/3}$ is

$$\begin{pmatrix} \cos \pi/3 & -\sin \pi/3 \\ \sin \pi/3 & \cos \pi/3 \end{pmatrix} = \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}$$

(by Example D3.2(i)), which has characteristic polynomial

$$(1/2 - \lambda)(1/2 - \lambda) + (\sqrt{3}/2)^2 = \lambda^2 - \lambda + 1.$$

This quadratic has discriminant $(-1)^2 - 4 \times 1 \times 1 = -3 < 0$, so it has no real roots. Hence $R_{\pi/3}$ has no eigenvalues.

iii. Let

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

Then

$$\chi_A(\lambda) = \det \begin{pmatrix} 1-\lambda & 2 & 1 \\ 0 & -1-\lambda & 0 \\ -2 & -2 & 4-\lambda \end{pmatrix},$$

which after some calculation gives

$$\chi_A(\lambda) = -(\lambda - 2)(\lambda - 3)(\lambda + 1).$$

The roots of χ_A are 2, 3 and -1, so these are the eigenvalues of A.

To calculate the eigenspaces, we have to take each eigenvalue λ in turn and compute $E_\lambda(A) = \ker(A - \lambda I)$ using Gaussian elimination, as in Example C7.10.

For instance, we calculate the 2-eigenspace as follows. We have

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

Computing in the usual way, we find that $A - 2I$ has RREF

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

and so

$$E_2(A) = \ker(A - 2I) = \left\{ \begin{pmatrix} t \\ 0 \\ t \end{pmatrix} : t \in \mathbb{R} \right\} = \operatorname{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

So the eigenvectors of A with eigenvalue 2 are the vectors $\left(\begin{smallmatrix} t \\ 0 \\ t \end{smallmatrix}\right)$ with $t \neq 0$, and a basis of the 2-eigenspace $E_2(A)$ is $\left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right)$.

Applying a similar method to each of the other two eigenvalues, we eventually find that $E_3(A) = \operatorname{span}\left\{\left(\begin{smallmatrix} 1 \\ 0 \\ 2 \end{smallmatrix}\right)\right\}$ and $E_{-1}(A) = \operatorname{span}\left\{\left(\begin{smallmatrix} 1 \\ -1 \\ 0 \end{smallmatrix}\right)\right\}$.

(When you do calculations like this, it's wise to check that your claimed eigenvectors really are eigenvectors! For instance, you should check that A multiplied by $\left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right)$ really is equal to $2\left(\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right)$.)

iv. Let

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

A calculation similar to the one in (iii) shows that

$$\chi_A(\lambda) = -(\lambda + 2)^2(\lambda - 1),$$

so the eigenvalues of A are -2 and 1.

Let us calculate the (-2)-eigenspace. We have

$$A - (-2)I = \begin{pmatrix} -3 & 3 & -3 \\ -6 & 6 & -6 \\ 0 & 0 & 0 \end{pmatrix},$$

which is a matrix of rank 1 and nullity 2. So this time, the eigenspace is 2-dimensional. One basis of it is

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

In this case, it was easy to spot a basis by inspection. In general, we can calculate a basis of $E_\lambda(A) = \ker(A - \lambda I)$ by the method in Example C7.10.

v. The characteristic polynomial of the diagonal matrix

$$A = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_n \end{pmatrix}$$

is

$$\det \begin{pmatrix} c_1 - \lambda & 0 & \cdots & 0 \\ 0 & c_2 - \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_n - \lambda \end{pmatrix}$$

which by Example C3.1(v) gives

$$\chi_A(\lambda) = (c_1 - \lambda)(c_2 - \lambda) \cdots (c_n - \lambda).$$

In Example E1.8(ii), we found that $c_1, c_2, \ldots, c_n$ are eigenvalues of A. We now know that they are the *only* eigenvalues of A, since they are the only roots of χ_A.

Remark E2.5 The method illustrated in these examples can in principle be used to compute the eigenvalues and eigenspaces of any square matrix. However, I want to be honest about its limitations.

Certainly there's no problem with this method for 2×2 matrices: it's fast and easy. But even for 3×3 matrices, we run into problems. If I choose a 3×3 matrix at random and ask you to find its eigenvalues using the method above, it's likely you won't be able to. After all, its characteristic polynomial will be some cubic (that is, some polynomial of degree three) with random coefficients, so how are you going to find its roots? Probably the only method you've seen for solving cubics is to guess a root by trying small integers and, assuming you find one, reducing it to a quadratic. But the chances of finding a root by trial and error when the matrix is chosen at random are very slim indeed.

(There is in fact a formula for the roots of a cubic polynomial, like the quadratic formula but much more complicated. There's even one for polynomials of degree four. But it's a theorem that no such formula exists for polynomials of degree five or higher, as you'll learn if you take Galois Theory in Year 4.)

It's important to realize that we're nice to you. When we give you a matrix in an assignment or exam and ask you to compute its eigenvalues, we always choose examples where it's possible for you to do it, and the answers are almost always small integers. But don't let this mislead you into thinking that the method is practical in general. It's emphatically not!

Computers are perfectly capable of finding the roots of any polynomial to any number of decimal places you want. But even on a computer, the characteristic polynomial method for finding eigenvalues is wildly impractical. Computing determinants by the method that we've learned takes an extremely long time: about $n!$ operations for an $n \times n$ matrix, and $n!$ grows very fast with n. For instance, we saw at the start of Section C6 that even for a 100×100 matrix (which is only a modest size), even on a supercomputer, it would take a ridiculous number of years—far longer than the age of the universe. Since the characteristic polynomial is a determinant, this method for computing eigenvalues is a disaster.

On the other hand, modern computer algebra packages can compute the eigenvalues of a randomly-generated 100×100 matrix in a fraction of a second. That's because they *don't* use the characteristic polynomial to do it. If you want to know how they actually do it, take Numerical Linear Algebra and Applications next year!

In summary, the characteristic polynomial is highly impractical as a method for computation. It is, however, very useful for *theoretical* purposes, as we are about to see.

The definition of eigenvalue gives no immediate clue as to how many eigenvalues a matrix could have. For instance, could there be a matrix A such that *every* real number is an eigenvalue of A? The answer is, in fact, no. As the examples above suggest, a matrix or linear operator can have only finitely many eigenvalues. More exactly:

Corollary E2.6 *An $n \times n$ square matrix or linear operator on $\mathbb{R}^n$ has at most n eigenvalues. In particular, it has only finitely many eigenvalues.*

Proof The eigenvalues of an operator are the same as those of its standard matrix, so it is enough to prove the result for matrices. By Lemma E2.1, the characteristic polynomial of an $n \times n$ matrix is a nonzero polynomial of degree n. But a nonzero polynomial of degree n has at most n roots, so the result follows from Proposition E2.3. □

Example E2.7 In Example E1.4(i), we considered the operator T on $\mathbb{R}^2$ that reflects the plane in the x-axis, and we showed that 1 and -1 are eigenvalues of T. Corollary E2.6 immediately implies that these are the *only* eigenvalues.

Although it was not obvious from the definition of eigenvalue that there would only be finitely many, perhaps it should not be *too* surprising. Take a square matrix A. For each scalar λ, we obtain another square matrix $A - \lambda I$. Now, in Workshop 6, you may have discovered that a square matrix 'chosen at random' is 'usually' invertible (in some sense that we didn't attempt to formulate exactly). So, as λ varies, we would expect most of these matrices $A - \lambda I$ to be invertible. In other words, we would expect λ to be an eigenvalue for only a few exceptional values of λ. Corollary E2.6 makes this vague statement precise.

In any case, every square matrix or linear operator has attached to it a finite set of scalars: its set of eigenvalues, which is called the **spectrum** of the matrix or operator. This is related to usages of the word *spectrum* in the physical sciences. When you hear people talk about mass spectroscopy or the emission spectrum of hydrogen, there is a connection with eigenvalues.

Remark E2.8 (Non-examinable.) In some branches of mathematics, it is important to consider eigenvalues and eigenvectors for *differential* operators. This course is not the place to give the definitions, but it is easy to see the resemblance between differential equations such as

$$f' = \lambda f$$

and the equation

$$T(\mathbf{x}) = \lambda \mathbf{x}$$

in the definition of eigenvalue. If we write f' as $D(f)$ then the first equation becomes $D(f) = \lambda f$. This D is an example of a 'differential operator', and it is linear in the sense that $D(af + bg) = (af + bg)' = af' + bg' = aD(f) + bD(g)$. The right way to make the connection precise is to use the language of vector spaces, but that also lies beyond this course.

The characteristic polynomial of a matrix tells us *what the eigenvalues are*. But it also gives us information about *how big the eigenspaces are*. It is only partial information, as we will see; but it is useful all the same.

Definition E2.9 Let A be a square matrix and let λ be an eigenvalue of A. The **geometric multiplicity** of λ is $\dim(E_\lambda(A))$, the dimension of the λ-eigenspace.

By Proposition E1.14(ii), the geometric multiplicity of an eigenvalue is always at least 1.

Examples E2.10 i. Let

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}.$$

The eigenvalues are 3 and 5, and

$$\begin{aligned} E_3(A) &= \ker(A - 3I) = \left\{ \left(\begin{smallmatrix} x \\ y \\ 0 \end{smallmatrix} \right) : x, y \in \mathbb{R} \right\} = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2\}, \\ E_5(A) &= \ker(A - 5I) = \left\{ \left(\begin{smallmatrix} 0 \\ 0 \\ z \end{smallmatrix} \right) : z \in \mathbb{R} \right\} = \operatorname{span}\{\mathbf{e}_3\}. \end{aligned}$$

Hence 3 and 5 have geometric multiplicities 2 and 1, respectively.

ii. Let

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

Then $\chi_A(\lambda) = \lambda^2$, so 0 is the only eigenvalue of A. We have

$$E_0(A) = \ker(A) = \left\{ \left(\begin{smallmatrix} x \\ 0 \end{smallmatrix} \right) : x \in \mathbb{R} \right\} = \operatorname{span}\{\mathbf{e}_1\},$$

so 0 has geometric multiplicity 1.

Let λ_0 be an eigenvalue of a matrix A. Then $\chi_A(\lambda_0) = 0$, so the polynomial $\chi_A(\lambda)$ has $(\lambda - \lambda_0)$ as a factor. Now when $\chi_A(\lambda)$ is written in fully factorized form, it may contain as a factor not just $(\lambda - \lambda_0)$ but a higher power, say $(\lambda - \lambda_0)^k$. This number k is called the **algebraic multiplicity** of λ_0. It is always at least 1.

Example E2.11 Let

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

Then

$$\chi_A(\lambda) = -(\lambda + 3)^2(\lambda - 8)(\lambda^2 + 1).$$

Since $\lambda^2 + 1$ has no real roots, $\chi_A(\lambda)$ factorizes no further (over $\mathbb{R}$), so the only eigenvalues of A are -3 and 8. The algebraic multiplicity of -3 is 2, and the algebraic multiplicity of 8 is 1.

Now the crucial fact is:

Theorem E2.12 *Let A be a square matrix and let λ be an eigenvalue of A. Then the geometric multiplicity of λ is less than or equal to its algebraic multiplicity.*

Proof Omitted. (You can find a proof in Lemma 4.26 of Poole, for instance.) □

Examples E2.13 i. The diagonal matrix of Example E2.10(i) has characteristic polynomial $-(\lambda-3)^2(\lambda-5)$, so the algebraic multiplicities of the eigenvalues 3 and 5 are 2 and 1, respectively. These are equal to their geometric multiplicities.

ii. On the other hand, in the 2×2 matrix A of Example E2.10(ii), the unique eigenvalue 0 has algebraic multiplicity 2 (since $\chi_A(\lambda) = (\lambda - 0)^2$) but geometric multiplicity only 1.

iii. Theorem E2.12 tells us that once we have calculated the characteristic polynomial of a matrix, we immediately have an upper bound on the dimension of every eigenspace, without actually having to calculate the eigenspaces.

For instance, suppose that a certain 6×6 real matrix has characteristic polynomial $(\lambda-3)(\lambda+2)^3(\lambda^2+5)$. Then the eigenvalues are 3 and -2.

The algebraic multiplicity of 3 is 1, so its geometric multiplicity is ≤ 1 by Theorem E2.12. But as 3 is an eigenvalue, its geometric multiplicity is at least 1; hence it is exactly 1.

The algebraic multiplicity of -2 is 3, so by Theorem E2.12, its geometric multiplicity is ≤ 3. Thus, its geometric multiplicity is 1, 2 or 3. With the information we are given, nothing more can be said; there are in fact examples showing that all three possibilities occur.

We have been telling the story of the characteristic polynomial in terms of matrices, but it can also be told for linear operators, as follows.

Lemma E2.14 *Similar matrices have the same characteristic polynomial, and therefore the same eigenvalues.*

Proof Let A and B be similar $n\times n$ matrices. Then $A = PBP^{-1}$ for some invertible matrix P. Hence

$$\chi_A(\lambda) = \det(PBP^{-1} - \lambda I) = \det(P(B-\lambda I)P^{-1}) = \det(B - \lambda I) = \chi_B(\lambda),$$

where the third equality holds because similar matrices have the same determinant (Lemma D7.1). □

Now let T be a linear operator T on $\mathbb{R}^n$. Proposition D6.10 tells us that the matrix of T with respect to any basis of $\mathbb{R}^n$ is similar to the matrix of T with respect to any other basis. So by Lemma E2.14, all these matrices have the same characteristic polynomial. We may therefore define the **characteristic polynomial** χ_T to be the characteristic polynomial of the matrix of T with respect to any basis of $\mathbb{R}^n$; it makes no difference which basis we choose.

In particular, χ_T is equal to the characteristic polynomial $\chi_{[T]}$ of its standard matrix. But the standard basis is not always the most convenient one, as the following example shows.

Example E2.15 Let $\mathbf{v}_1 = \binom{3}{1}$ and $\mathbf{v}_2 = \binom{7}{4}$, and let T be the unique linear operator on $\mathbb{R}^2$ such that $T(\mathbf{v}_1) = 2\mathbf{v}_1$ and $T(\mathbf{v}_2) = -\mathbf{v}_2$, as in Example D6.5. The matrix of T with respect to the basis $\mathbf{v}_1, \mathbf{v}_2$ is

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

so the characteristic polynomial χ_T of T is given by

$$\chi_T(\lambda) = \chi_B(\lambda) = (\lambda - 2)(\lambda + 1).$$

We calculated in Example D6.5 that the standard matrix of T is

$$A = \frac{1}{5}\begin{pmatrix} 31 & -63 \\ 12 & -26 \end{pmatrix}.$$

Since A and B are similar, $\chi_A(\lambda) = \chi_B(\lambda)$. In this example, it is much easier to calculate the characteristic polynomial of the non-standard matrix B than of the standard matrix A.

E3 Diagonalizable matrices

Diagonal matrices are fantastically easy to work with. Take a diagonal matrix

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix},$$

which for convenience, I will write as

$$\mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).$$

(Here $\lambda_1, \ldots, \lambda_n$ are any scalars.) Then it is very easy to calculate all the powers of A: quite simply,

$$A^r = \mathrm{diag}(\lambda_1^r, \lambda_2^r, \ldots, \lambda_n^r)$$

for any integer $r \geq 0$. It is very easy to tell whether A is invertible: it is if every λ_i is nonzero, and otherwise it's not. If A *is* invertible then

$$A^{-1} = \mathrm{diag}(1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_n)$$

(which in fact is just the formula above for A^r with $r = -1$). It is very easy to calculate the determinant:

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n.$$

Also very easy are the rank and nullity: the rank is the number of values of i such that $\lambda_i \neq 0$, and the nullity is the number of values of i such that $\lambda_i = 0$. So, the algebra of diagonal matrices is incredibly straightforward.

The geometry is very easy too. Let T be the linear operator on $\mathbb{R}^n$ with standard matrix A (that is, $T = L_A$). Then T simply scales by a factor of λ_1 in the $\mathbf{e}_1$ direction, λ_2 in the $\mathbf{e}_2$ direction, and so on. (We met this kind of operator in Example D2.2.) Figure E.2 shows the case $A = \left(\begin{smallmatrix} 2 & 0 \\ 0 & 3 \end{smallmatrix}\right)$.

So, life would be a breeze if every matrix was diagonal. Of course, that's not true! Most matrices aren't diagonal. And similarly, given a linear operator T on $\mathbb{R}^n$, the standard matrix of T is not usually diagonal. But many linear operators do have the property that their matrix is diagonal *if you choose the*

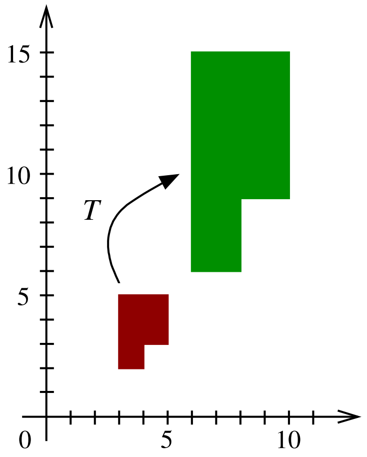


Figure E.2: The linear operator T on $\mathbb{R}^2$ with standard matrix $\text{diag}(2, 3)$

right basis. How do we know which basis is right? As we'll see, the answer involves eigenvectors.

The matrices of an operator T with respect to different bases are all similar, so if the matrix of T with respect to one well-chosen basis is diagonal then all the other matrices are *similar* to a diagonal matrix. There is a word for this.

Definition E3.1 A square matrix is **diagonalizable** if it is similar to some diagonal matrix.

In other words, an $n \times n$ matrix A is diagonalizable if and only if there exists an invertible $n \times n$ matrix P such that $P^{-1}AP$ is diagonal.

Example E3.2 Certainly every diagonal matrix is diagonalizable. But some non-diagonal matrices are diagonalizable too. For instance, we showed in Example D6.5 that $\frac{1}{5}\left(\begin{smallmatrix} 31 & -63 \\ 12 & -26 \end{smallmatrix}\right)$ is similar to the diagonal matrix $\text{diag}(2, -1)$, so it is diagonalizable.

We're interested in the equation $P^{-1}AP = D$, where P is invertible and D is diagonal. Here's how to understand this equation:

Proposition E3.3 *Let A, P and D be $n \times n$ matrices, with P invertible and D diagonal. Write $P = (\mathbf{v}_1 | \mathbf{v}_2 | \cdots | \mathbf{v}_n)$ and $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Then*

$$P^{-1}AP = D \iff A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1, \ A\mathbf{v}_2 = \lambda_2 \mathbf{v}_2, \ \ldots, \ A\mathbf{v}_n = \lambda_n \mathbf{v}_n.$$

Proof We have $P^{-1}AP = D$ if and only if $AP = PD$. For $i \in \{1, \ldots, n\}$, the ith column of AP is $A\mathbf{v}_i$ (by Lemma A4.3(iii)) and the ith column of PD is $\lambda_i \mathbf{v}_i$. Hence $AP = PD$ if and only if $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for all $i \in \{1, \ldots, n\}$. □

So, the columns of a diagonalizing matrix P are eigenvectors, and the entries of the diagonal matrix are eigenvalues.

The central idea in this section is that diagonalizability is intimately connected to the existence of enough eigenvectors:

Proposition E3.4 *Let T be a linear operator on $\mathbb{R}^n$. The following are equivalent:*

i. the matrix of T with respect to some basis of $\mathbb{R}^n$ is diagonal;

ii. there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of T;

iii. there exist n linearly independent eigenvectors of T;

iv. the standard matrix of T is diagonalizable.

Proof We have (ii)⇔(iii) by Proposition B5.4. Now we prove that (i)⇒(iv)⇒(ii)⇒(i).

(i)⇒(iv) follows from Proposition D6.10.

For (iv)⇒(ii), suppose that $[T]$ is diagonalizable. Then $P^{-1}[T]P = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ for some invertible P and scalars λ_i. Since P is invertible, its columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are a basis of $\mathbb{R}^n$. By Proposition E3.3, $[T]\mathbf{v}_i = \lambda_i \mathbf{v}_i$ for all $i \in \{1, \ldots, n\}$. On the other hand, $T(\mathbf{x}) = [T]\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$. So for each i, we have $T(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$. Moreover, $\mathbf{v}_i \neq \mathbf{0}$ (by Example B3.2(vi)). Hence $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are eigenvectors.

(ii)⇒(i): if $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of eigenvectors with respective eigenvalues $\lambda_1, \ldots, \lambda_n$, then the matrix of T with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is $\operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.□

A linear operator T is said to be **diagonalizable** if it satisfies any of the four equivalent conditions of Proposition E3.4. (If it satisfies one, it satisfies them all.)

Examples E3.5 i. The operator T on $\mathbb{R}^2$ shown in Figure D.5 is diagonalizable, since $\mathbf{v}_1, \mathbf{v}_2$ is a basis of $\mathbb{R}^2$ consisting of eigenvectors of T.

ii. Let T be the operator on $\mathbb{R}^2$ that reflects in the x-axis. Then T has two linearly independent eigenvectors, $\mathbf{e}_1$ and $\mathbf{e}_2$, so it is diagonalizable.

iii. Let V be a line through the origin in $\mathbb{R}^3$ and consider P_V, orthogonal projection onto V. Choose a nonzero point $\mathbf{v}_1$ on V and a basis $\mathbf{v}_2, \mathbf{v}_3$ for the plane $V^{\perp}$. Then $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is a basis of $\mathbb{R}^3$. Now $P_V(\mathbf{v}_1) = \mathbf{v}_1$ and $P_V(\mathbf{v}_2) = P_V(\mathbf{v}_3) = \mathbf{0}$, so the matrix of P_V with respect to the basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is $\operatorname{diag}(1, 0, 0)$. Hence P_V is diagonalizable.

Proposition E3.4 tells us what it means geometrically for an operator T to be diagonalizable. It says that T is diagonalizable if and only if it is a scaling operator of the type described in Example D2.2. Figures E.1 and E.2 show an example where the basis concerned happens to be the standard basis (and in particular, orthogonal). Figure E.3 shows another example on $\mathbb{R}^2$ with a non-orthogonal basis $\mathbf{v}_1, \mathbf{v}_2$ of eigenvectors.

Translating from linear operators into square matrices, Proposition E3.4 implies:

Proposition E3.6 *Let A be an $n \times n$ matrix. The following are equivalent:*

i. there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A;

ii. there exist n linearly independent eigenvectors of A;

iii. A is diagonalizable.

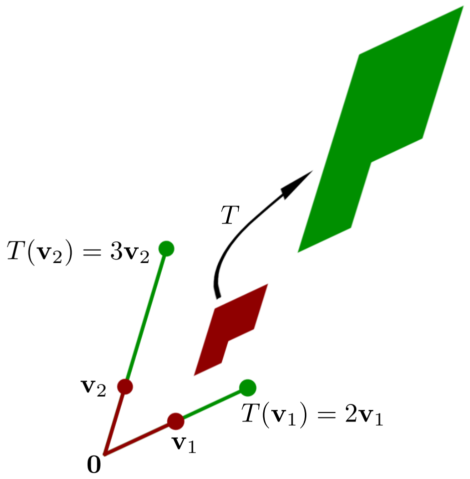

T
$T(\mathbf{v}_2) = 3\mathbf{v}_2$
$\mathbf{v}_2$
$T(\mathbf{v}_1) = 2\mathbf{v}_1$
$\mathbf{v}_1$
$\mathbf{0}$

Figure E.3: A diagonalizable operator T on $\mathbb{R}^2$ with non-orthogonal eigenvectors

Proof Put $T = L_A$. Then $[T] = A$ (by Theorem D3.7), and an eigenvector of A is exactly an eigenvector of T. The result follows from Proposition E3.4. □

If $\mathbf{v}$ is an eigenvector of T with eigenvalue λ then so is every nonzero scalar multiple of $\mathbf{v}$. It follows that if $\mathbf{v}_1$ and $\mathbf{v}_2$ are two eigenvectors with different eigenvalues then neither is a scalar multiple of the other; that is, $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent. The next result shows that the same is true for any number of eigenvectors, not just two.

Proposition E3.7 *Let A be a square matrix. Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be eigenvectors of A, with eigenvalues $\lambda_1, \ldots, \lambda_m$ respectively. If $\lambda_1, \ldots, \lambda_m$ are pairwise distinct (that is, $\lambda_i \neq \lambda_j$ for all $i \neq j$) then $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent.*

Briefly put: eigenvectors with distinct eigenvalues are linearly independent.

Proof We prove this by induction on m. The result holds when $m = 0$, by Example B3.2(v). Suppose that $m \geq 1$, assume the result for $m - 1$, and let $c_1, \ldots, c_m$ be scalars such that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}. \qquad \text{(E:1)}$$

We have to prove that $c_1 = c_2 = \cdots = c_m = 0$.

For all $i \in \{1, \ldots, m\}$, we have

$$(A - \lambda_m I)\mathbf{v}_i = A\mathbf{v}_i - \lambda_m\mathbf{v}_i = (\lambda_i - \lambda_m)\mathbf{v}_i.$$

So multiplying each side of (E:1) by $A - \lambda_m I$ gives

$$c_1(\lambda_1 - \lambda_m)\mathbf{v}_1 + c_2(\lambda_2 - \lambda_m)\mathbf{v}_2 + \cdots + c_{m-1}(\lambda_{m-1} - \lambda_m)\mathbf{v}_{m-1} = \mathbf{0}.$$

By inductive hypothesis, $\mathbf{v}_1, \ldots, \mathbf{v}_{m-1}$ are linearly independent, so

$$c_1(\lambda_1 - \lambda_m) = c_2(\lambda_2 - \lambda_m) = \cdots = c_{m-1}(\lambda_{m-1} - \lambda_m) = 0.$$

But $\lambda_1, \ldots, \lambda_m$ are pairwise distinct, so $c_1 = c_2 = \cdots = c_{m-1} = 0$. Now (E:1) gives $c_m\mathbf{v}_m = \mathbf{0}$, and since the eigenvector $\mathbf{v}_m$ is by definition nonzero, $c_m = 0$ too. So $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are linearly independent, completing the induction. □

Remark E3.8 Since any linearly independent set in $\mathbb{R}^n$ has at most n elements, Proposition E3.7 provides an alternative proof of the fact that an $n \times n$ matrix has at most n eigenvalues (Corollary E2.6). This alternative proof has the advantage of not relying on determinants.

Theorem E3.9 *Let A be an $n \times n$ matrix with n distinct eigenvalues. Then A is diagonalizable.*

Proof By hypothesis, we can choose n eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ with pairwise distinct eigenvalues. By Proposition E3.7, $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. Hence by Proposition E3.6, A is diagonalizable. □

Warning E3.10 The converse is false. For instance, the 2×2 identity matrix is certainly diagonalizable (it's diagonal!) but its only eigenvalue is 1.

Examples E3.11 i. Let

$$A = \begin{pmatrix} 5 & -1 & -2 \\ 0 & -6 & 3 \\ 0 & 0 & 3 \end{pmatrix}.$$

Then

$$\chi_A(\lambda) = \det \begin{pmatrix} 5-\lambda & -1 & -2 \\ 0 & -6-\lambda & 3 \\ 0 & 0 & 3-\lambda \end{pmatrix} = -(\lambda - 5)(\lambda + 6)(\lambda - 3).$$

(This determinant is most efficiently calculated by first expanding down the first column of $A - \lambda I$, then expanding down the second column of the remaining 2×2 matrix, as in Proposition C3.2 and Example C3.3.) So A has eigenvalues 5, -6 and 3. Since A is a 3×3 matrix with 3 distinct eigenvalues, it is diagonalizable.

Suppose we wish to find an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$. Proposition E3.3 tells us how: the diagonal matrix D has the eigenvalues down the diagonal, and the columns of P are any eigenvectors with those eigenvalues. In this example, we calculate that

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in \ker(A - 5I), \quad \begin{pmatrix} 1 \\ 11 \\ 0 \end{pmatrix} \in \ker(A + 6I), \quad \begin{pmatrix} 7 \\ 2 \\ 6 \end{pmatrix} \in \ker(A - 3I),$$

by the usual method for solving linear systems. Hence if we put

$$P = \begin{pmatrix} 1 & 1 & 7 \\ 0 & 11 & 2 \\ 0 & 0 & 6 \end{pmatrix}, \quad D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

then $P^{-1}AP = D$.

Geometrically, this means that the linear operator $L_A \colon \mathbf{x} \mapsto A\mathbf{x}$ has the effect of scaling by a factor of 5 in the direction of $\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right)$, a factor of -6 in the direction of $\left(\begin{smallmatrix} 1 \\ 11 \\ 0 \end{smallmatrix}\right)$, and a factor of 3 in the direction of $\left(\begin{smallmatrix} 7 \\ 2 \\ 6 \end{smallmatrix}\right)$.

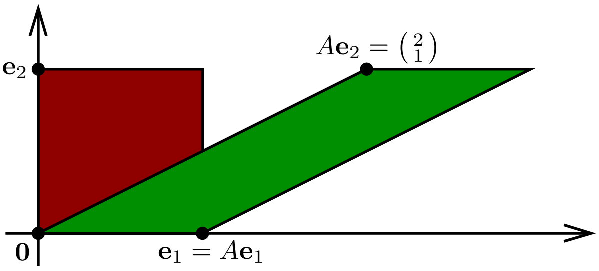

$A\mathbf{e}_2 = \left(\begin{smallmatrix}2\\1\end{smallmatrix}\right)$
$\mathbf{e}_2$
$\mathbf{0}$
$\mathbf{e}_1 = A\mathbf{e}_1$

Figure E.4: A shearing operator on the plane

ii. Let

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

After some calculation, we find that

$$\chi_A(\lambda) = -\lambda^3 + 15\lambda^2 + 18\lambda.$$

Clearly 0 is a root of χ_A, and the quadratic formula gives $(15 \pm \sqrt{297})/2$ as the other two roots. So, these are the eigenvalues of A. The square roots would make it somewhat painful to calculate eigenvectors, but even without doing so, Theorem E3.9 immediately implies that there is a basis of $\mathbb{R}^3$ consisting of eigenvectors of A. In other words, A is diagonalizable.

For all we know so far, *every* square matrix could be diagonalizable. This is not, in fact, true.

Examples E3.12 i. We saw in Example E2.10(ii) that the only eigenvalue of the matrix $A = \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right)$ is 0, and the 0-eigenspace is span$\{\mathbf{e}_1\}$. So there is no basis of $\mathbb{R}^2$ consisting of eigenvectors of A, that is, A is not diagonalizable.

Another way to see this is that if A were diagonalizable then A would be similar to a diagonal matrix whose diagonal entries are eigenvalues of A (by Proposition E3.3). In this case, this would mean that A is similar to the zero matrix. However, the only matrix similar to the zero matrix is itself (as in Example D6.8(ii)), and $A \neq 0$.

ii. Take the 5×5 matrix A of Example E2.11. This has two eigenvalues, with algebraic multiplicities 2 and 1. By Theorem E2.12, the dimensions of the eigenspaces are at most 2 and 1 (respectively). Since $2 + 1 < 5$, it is not possible to find 5 linearly independent eigenvectors of A. Hence A is not diagonalizable.

E4 Orthogonal matrices

In general, applying a linear operator to $\mathbb{R}^n$ distorts the geometry: both lengths and angles can change. For instance, let $A = \left(\begin{smallmatrix}1&2\\0&1\end{smallmatrix}\right)$ (Figure E.4). Then $\mathbf{e}_2$ has length 1, but $A\mathbf{e}_2 = \left(\begin{smallmatrix}2\\1\end{smallmatrix}\right)$ has length $\sqrt{5}$. So the linear operator L_A changes lengths. It also changes angles. For instance, $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal, but $A\mathbf{e}_1$ and $A\mathbf{e}_2$ are not.

From a geometric point of view, it is natural to pay special attention to operators that *do* preserve lengths and angles. Those operators (or their standard matrices) are called 'orthogonal'.

Definition E4.1 A real square matrix A is **orthogonal** if A is invertible and $A^{-1} = A^T$.

So, A is orthogonal if and only if $A^TA = I$ and $AA^T = I$. In fact, either one of $A^TA = I$ and $AA^T = I$ implies the other (for *square* matrices A), by Corollary C2.5.

This doesn't appear to have anything to do with preservation of length and angle! But we will eventually show that it really does. (We start from the definition above, rather than an equivalent condition on preservation of length and angle, because it is easier to work with algebraically.)

Examples E4.2 i. We saw in Example D3.2(i) that rotation by an angle of θ, as a linear operator R_θ on $\mathbb{R}^2$, has standard matrix

$$[R_\theta] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Our geometric intuition says that this matrix *ought* to be orthogonal, because rotating the plane doesn't change lengths or angles. And in fact it is, since

$$[R_\theta]^T[R_\theta] = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

ii. Let $A = cI$, where c is a scalar. Then the linear operator L_A is simply $\mathbf{x} \mapsto c\mathbf{x}$. This doesn't change angles, but it does change lengths (unless $c = \pm 1$), so we would not expect A to be orthogonal. And indeed,

$$A^TA = (cI)(cI) = c^2I,$$

so A is orthogonal if and only if $c = \pm 1$.

If the definition of orthogonal matrix succeeds in capturing the notion of 'keeps lengths and angles the same', then the product of two orthogonal matrices ought to be orthogonal. After all, if each stage of a two-stage process preserves lengths and angles, then the composite process ought to preserve them too. For the same kind of reason, the inverse (or equivalently transpose) of an orthogonal matrix ought to be orthogonal, and the identity matrix should be orthogonal too. This is the intuition behind the following lemma.

Lemma E4.3 *Let A and B be orthogonal $n \times n$ matrices.*

i. If A and B are orthogonal then AB is orthogonal.

ii. The identity matrix I is orthogonal.

iii. If A is orthogonal then so is A^T.

Proof For (i), if A and B are orthogonal then

$$(AB)^T(AB) = B^TA^TAB = B^TIB = B^TB = I,$$

so AB is orthogonal. For (ii), $I^TI = I^2 = I$. For (iii), just note that $(A^T)^T = A$. □

Orthogonality is an important concept, and like many important concepts, there are several equivalent ways of looking at it.

Proposition E4.4 *Let A be a real square matrix. The following are equivalent:*

i. A is orthogonal;

ii. the columns of A are orthonormal;

iii. the rows of A are orthonormal.

Proof We prove that (i)⇔(ii). Since the rows of A are the columns of A^T, the equivalence (i)⇔(iii) will then follow from Lemma E4.3(iii).

We use the convention that the (p, q)-entry of a matrix M is written as M_{pq}. Write $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ for the columns of A. For $i, k \in \{1, \dots, n\}$, we have

$$(A^TA)_{ik} = \sum_{j=1}^{n}(A^T)_{ij}A_{jk} = \sum_{j=1}^{n}A_{ji}A_{jk} = \mathbf{v}_i \cdot \mathbf{v}_k.$$

Now

$$\begin{aligned}
A \text{ is orthogonal} &\iff A^TA = I \\
&\iff (A^TA)_{ik} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases} \quad \text{for all } i, k \in \{1, \dots, n\} \\
&\iff \mathbf{v}_i \cdot \mathbf{v}_k = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{if } i \neq k \end{cases} \quad \text{for all } i, k \in \{1, \dots, n\} \\
&\iff \mathbf{v}_1, \dots, \mathbf{v}_n \text{ are orthonormal,}
\end{aligned}$$

where the last step is by Lemma B6.3. □

In the light of this proposition, orthogonal matrices should really be called *orthonormal* matrices. But unfortunately, they're not.

Remark E4.5 Proposition E4.4 implies that if the rows of a square matrix are orthonormal then so are the columns. Proving this from scratch is really quite hard. Even for 2×2 matrices, it's not easy. If you think it is, try it!

Examples E4.6 i. In the rotation matrix $[R_\theta]$ of Example E4.2(i), both columns have length 1 (because $\cos^2\theta + \sin^2\theta = 1$) and the dot product of the two columns is 0. Hence the columns are orthonormal. This gives a slightly easier proof that $[R_\theta]$ is orthogonal.

ii. By Proposition E4.4, the columns of an orthogonal matrix all have length 1, which implies in particular that all the entries of an orthogonal matrix are between -1 and 1. So if you are given a matrix that contains even one entry with absolute value > 1, you can immediately conclude that it is not orthogonal.

iii. Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of $\mathbb{R}^n$, and recall from Theorem D6.4 that the corresponding change of basis matrix is $(\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_n)$. By Proposition E4.4, the basis is orthonormal if and only if its change of basis matrix is orthogonal.

The next few results fulfil the promise made at the start of this section: that the orthogonal matrices are those that preserve angle and length.

Proposition E4.7 *Let A be a real square matrix. Then A is orthogonal if and only if*

$$(A\mathbf{x}) \cdot (A\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Proof Suppose A is orthogonal. Then for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, using Lemma A5.7,

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

Conversely, suppose that $(A\mathbf{x}) \cdot (A\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Let $i, j \in \{1, \ldots, n\}$. Then

$$(A\mathbf{e}_i) \cdot (A\mathbf{e}_j) = \mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

But $A\mathbf{e}_i$ is the ith column of A (by Lemma A4.3(ii)), so the columns of A are orthonormal. Hence by Proposition E4.4, A is orthogonal. □

The formula '$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\| \, \|\mathbf{y}\| \cos\theta$' for the dot product of vectors $\mathbf{x}$ and $\mathbf{y}$ (where θ is the angle between them) demonstrates that the dot product combines aspects of length and angle. Length can be expressed in terms of the dot product, via the formula $\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$. Less obvious is the converse, that the dot product can be expressed in terms of length alone. This is the main content of the next lemma, which is related to the fact that the lengths of the sides of a triangle determine its angles.

Lemma E4.8 (Polarization identity) *Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then*

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{4}\Big(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2\Big).$$

Proof We have

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 &= (\mathbf{x} + \mathbf{y}) \cdot (\mathbf{x} + \mathbf{y}) - (\mathbf{x} - \mathbf{y}) \cdot (\mathbf{x} - \mathbf{y}) \\ &= (\mathbf{x} \cdot \mathbf{x} + 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y}) - (\mathbf{x} \cdot \mathbf{x} - 2\mathbf{x} \cdot \mathbf{y} + \mathbf{y} \cdot \mathbf{y}) \\ &= 4\mathbf{x} \cdot \mathbf{y}. \end{aligned}$$ □

We can now show that the orthogonal matrices are exactly those that preserve length, in the following sense:

Proposition E4.9 *Let A be a real square matrix. Then A is orthogonal if and only if*

$$\|A\mathbf{x}\| = \|\mathbf{x}\|$$

for all $\mathbf{x} \in \mathbb{R}^n$.

Proof Suppose that A is orthogonal. Then by Proposition E4.7,

$$\|A\mathbf{x}\| = \sqrt{(A\mathbf{x}) \cdot (A\mathbf{x})} = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \|\mathbf{x}\|$$

for all $\mathbf{x} \in \mathbb{R}^n$.

Conversely, suppose $\|A\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$. We show that $(A\mathbf{x}) \cdot (A\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$; then Proposition E4.7 will imply that A is orthogonal. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then

$$\begin{aligned}
(A\mathbf{x}) \cdot (A\mathbf{y}) &= \frac{1}{4}\Big(\|A\mathbf{x} + A\mathbf{y}\|^2 - \|A\mathbf{x} - A\mathbf{y}\|^2\Big) && \text{by the polarization identity} \\
&= \frac{1}{4}\Big(\|A(\mathbf{x} + \mathbf{y})\|^2 - \|A(\mathbf{x} - \mathbf{y})\|^2\Big) && \text{by linearity} \\
&= \frac{1}{4}\Big(\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2\Big) && \text{by hypothesis on } A \\
&= \mathbf{x} \cdot \mathbf{y} && \text{by the polarization identity,}
\end{aligned}$$

as required. □

Warning E4.10 When A is an orthogonal matrix, the angle between $A\mathbf{x}$ and $A\mathbf{y}$ is equal to the angle between $\mathbf{x}$ and $\mathbf{y}$ (assuming $\mathbf{x} \neq \mathbf{0} \neq \mathbf{y}$). This follows from the definition of angle (Definition A3.4) together with Propositions E4.7 and E4.9. So, if A is orthogonal then the operator L_A corresponding to A preserves angles. However, the converse is false. For example, if $A = 2I$ then $L_A(\mathbf{x}) = 2\mathbf{x}$ for all $\mathbf{x}$, so L_A preserves angles even though A is not orthogonal.

We have been discussing orthogonality of matrices. But what should it mean for an *operator* to be orthogonal?

Lemma E4.11 *Let T be a linear operator on $\mathbb{R}^n$. The following are equivalent:*

i. the standard matrix of T is orthogonal;

ii. $T(\mathbf{x}) \cdot T(\mathbf{y}) = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$;

iii. $\|T(\mathbf{x})\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$.

Proof Write A for the standard matrix of T. Then $T = L_A$, so the result follows from Propositions E4.7 and E4.9. □

A linear operator T is **orthogonal** if it satisfies any of the equivalent conditions of Lemma E4.11.

Examples E4.12 i. Which linear operators on $\mathbb{R}^2$ are orthogonal?

We have already seen that when T is rotation of $\mathbb{R}^2$ by any angle, its standard matrix $[T]$ is orthogonal. So by definition, T itself is orthogonal. The same is true for reflection of $\mathbb{R}^2$ in any line through the origin. Indeed, as you discovered in Workshop 8, if F_θ denotes reflection in the line through $\mathbf{0}$ at angle θ from the positive x-axis then

$$[F_\theta] = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}.$$

It is easy to see that the columns are orthonormal, so the operator F_θ is orthogonal.

These results confirm our geometric intuition that rotation and reflection preserve length. In fact, rotations and reflections are the *only* 2×2 orthogonal matrices. Proving this is a pleasant exercise.

ii. The orthogonal operators on $\mathbb{R}^3$ can also be completely classified. It turns out that they are the following:

   - rotation by any angle about any axis through the origin;
   - reflection in any plane through the origin;
   - 'reflection in the origin', that is, $\mathbf{x} \mapsto -\mathbf{x}$.

An important difference between rotations and reflections is that rotations have determinant 1 and reflections have determinant -1. There are no other possibilities for the determinant, because of the following result.

Lemma E4.13 *Every orthogonal matrix or operator has determinant* ± 1.

Proof Let A be an orthogonal matrix. Then

$$1 = \det I = \det(A^T A) = \det(A^T) \det A = \det(A)^2$$

by parts (vi) and (v) of Proposition C3.4. Hence $\det A = \pm 1$.

Now let T be an orthogonal operator. By definition, $[T]$ is orthogonal, so $\det [T] = \pm 1$. But $\det T = \det [T]$, so $\det T = \pm 1$. □

This is geometrically plausible. An orthogonal operator T preserves lengths and angles, so it also preserves area (in the case of $\mathbb{R}^2$), volume (in the case of $\mathbb{R}^3$), and so in in higher dimensions. We saw in Section D7 that applying T multiplies volumes by a factor of $|\det(T)|$. So, $|\det(T)|$ ought to be 1, giving $\det(T) = \pm 1$. Lemma E4.13 proves that this is true.

E5 A little linear algebra over $\mathbb{C}$

This course has been entirely based on the real numbers. We have been working with vectors in $\mathbb{R}^n$, linear subspaces of $\mathbb{R}^n$, matrices with entries in $\mathbb{R}$, linear transformations between $\mathbb{R}^n$ and $\mathbb{R}^m$, and linear systems with both coefficients and solutions in $\mathbb{R}$.

However, in much of what we've done, the only features of $\mathbb{R}$ that we have really needed are that you can add, subtract, multiply and divide real numbers (as long as you don't try to divide by 0). A set equipped with operations of addition, subtraction, multiplication and division, obeying the familiar laws, is called a 'field'. (You will learn the precise definition in a later course.) Other examples of fields are $\mathbb{C}$ and $\mathbb{Q}$. A less obvious example is the set of integers modulo p, for any prime p; it is a highly significant fact that these can be added, subtracted, multiplied and divided like real or complex numbers.

So the question arises: how much of the linear algebra that we have done in this course still works if we replace $\mathbb{R}$ by some other field?

Before I give the answer, let me explain why I am asking at this particular moment. Way back in Section A6, I mentioned that some facts about the real numbers are most easily proved by using the complex numbers. And it turns out that some facts about linear algebra over $\mathbb{R}$ are most easily proved by using linear algebra over $\mathbb{C}$. In the final section of the course, we will meet one such fact. For this reason, we are about to need some complex linear algebra.

The answer to the question is: most of the linear algebra that we have done in this course still works with other fields in place of $\mathbb{R}$. In particular, most of it works when $\mathbb{R}$ is replaced by $\mathbb{C}$. Let us now look a little more closely at linear algebra over $\mathbb{C}$.

The definitions of linear transformation, linear operator, eigenvector, eigenvalue, determinant and characteristic polynomial over $\mathbb{C}$ are all exactly the same as over $\mathbb{R}$, simply changing $\mathbb{R}$ to $\mathbb{C}$ throughout. For instance, let A be an $n \times n$ complex matrix (that is, matrix with entries in $\mathbb{C}$), and let $\lambda \in \mathbb{C}$. An **eigenvector** of A, with **eigenvalue** λ, is a nonzero vector $\mathbf{z} \in \mathbb{C}^n$ such that $A\mathbf{z} = \lambda\mathbf{z}$. Just as for $\mathbb{R}$ (Proposition E2.3), a scalar $\lambda \in \mathbb{C}$ is an eigenvalue of A if and only if it is a root of the characteristic polynomial χ_A.

Warning E5.1 The terminology concerning real and complex eigenvalues is slightly delicate.

- For linear operators on $\mathbb{C}^n$, the eigenvalues are complex numbers (which may or may not be real), and the eigenvectors are in $\mathbb{C}^n$.

- For linear operators on $\mathbb{R}^n$, the eigenvalues are by definition *real* numbers, and the eigenvectors are by definition in $\mathbb{R}^n$.

- For complex matrices, the eigenvalues are complex numbers (which may or may not be real), and the eigenvectors are in $\mathbb{C}^n$.

- But for *real matrices*, the situation is less clear-cut. For example, what does it mean to say 'an eigenvalue of $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$'? If we are treating it as a *real* matrix, it has no eigenvalues. But if we are treating it as a *complex* matrix, it has eigenvalues $\pm i$.

  So in the context of real matrices, it is helpful to speak of 'real eigenvalues' and 'complex eigenvalues', for clarity. For example, the matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has complex eigenvalues $\pm i$, but no real eigenvalues at all.

We have seen that some real matrices have no real eigenvalues. A crucial advantage of $\mathbb{C}$ over $\mathbb{R}$ is that *every complex square matrix has an eigenvalue:*

Theorem E5.2 *Let $n \geq 1$, and let A be an $n \times n$ complex matrix. Then A has at least one eigenvalue.*

Proof The characteristic polynomial χ_A is a polynomial over $\mathbb{C}$ of degree $n \geq 1$, so by the fundamental theorem of algebra (Theorem A6.1), has at least one root in $\mathbb{C}$. Any such root is an eigenvalue of A. □

Soon, we will use this result about complex matrices to prove a result about real matrices; but first we need to consider complex conjugates.

Any $m \times n$ complex matrix $A = (A_{ij})$ has a **complex conjugate** $\overline{A}$. By definition, this is the $m \times n$ complex matrix whose (i, j)-entry is $\overline{A_{ij}}$. Complex conjugation of matrices preserves addition and multiplication, just as in equations (A:14) (page 41):

Lemma E5.3 *Let A and B be $m \times n$ matrices over $\mathbb{C}$, and let C be an $n \times p$ matrix over $\mathbb{C}$. Then*

$$\overline{A+B} = \overline{A} + \overline{B}, \qquad \overline{BC} = \overline{B}\,\overline{C}.$$

Proof We prove the second equation; the first is no harder and is left as an exercise. We use the convention that the (p, q)-entry of a matrix M is written as M_{pq}.

First, both $\overline{BC}$ and $\overline{B}\,\overline{C}$ are $m \times p$ matrices. Now let $1 \leq i \leq m$ and $1 \leq k \leq p$. We have

$$\begin{aligned}
(\overline{BC})_{ik} &= \overline{(BC)_{ik}} && \text{by definition of } \overline{BC} \\
&= \overline{\sum_{j=1}^{n} B_{ij}C_{jk}} && \text{by definition of matrix multiplication} \\
&= \sum_{j=1}^{n} \overline{B_{ij}}\,\overline{C_{jk}} && \text{by equations (A:14)} \\
&= \sum_{j=1}^{n} \overline{B}_{ij}\,\overline{C}_{jk} && \text{by definition of } \overline{B} \text{ and of } \overline{C} \\
&= (\overline{B}\,\overline{C})_{ik} && \text{by definition of matrix multiplication.}
\end{aligned}$$ □

The **length** of a complex vector $\mathbf{z} \in \mathbb{C}^n$ is

$$\|\mathbf{z}\| = \sqrt{\sum_{i=1}^{n} |z_i|^2}.$$

If $\mathbf{z}$ happens to be real then this agrees with the familiar definition of length in $\mathbb{R}^n$. (But note that in general, z_i^2 is not real, and not equal to $|z_i|^2$.)

Since a complex vector $\mathbf{z}$ can be regarded as an $n \times 1$ complex matrix, we already have a definition of the conjugate of $\mathbf{z}$: it is the vector $\overline{\mathbf{z}} \in \mathbb{C}^n$ with ith entry $\overline{z_i}$.

Lemma E5.4 $\overline{\mathbf{z}}^T\mathbf{z} = \|\mathbf{z}\|^2$ *for all* $\mathbf{z} \in \mathbb{C}^n$.

Proof First, $\overline{\mathbf{z}}^T$ is a $1 \times n$ matrix over $\mathbb{C}$ and $\mathbf{z}$ is an $n \times 1$ matrix over $\mathbb{C}$, so $\overline{\mathbf{z}}^T\mathbf{z}$ is a 1×1 matrix over $\mathbb{C}$, that is, a complex number. Now using the fact that $\overline{w}w = |w|^2$ for all $w \in \mathbb{C}$ (equation (A:15), page 41), we have

$$\overline{\mathbf{z}}^T\mathbf{z} = \sum_{i=1}^{n} \overline{z}_i z_i = \sum_{i=1}^{n} |z_i|^2 = \|\mathbf{z}\|^2.$$ □

When $\mathbf{z}$ is real, this lemma simply says that $\mathbf{z}^T\mathbf{z} = \|\mathbf{z}\|^2$. Since $\mathbf{x}^T\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (by Lemma A5.7), this is the familiar statement that $\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2$ for $\mathbf{x} \in \mathbb{R}^n$.

For real matrices A, we are often interested in the condition that A is symmetric: $A^T = A$. For complex matrices A, symmetry turns out to be less interesting than the condition that $\overline{A}^T = A$. (Such matrices are called 'Hermitian'.) Our next result concerns such matrices.

Proposition E5.5 *Let A be an $n \times n$ complex matrix such that $\overline{A}^T = A$. Then every eigenvalue of A is real.*

Proof Let $\lambda \in \mathbb{C}$ be an eigenvalue of A. Choose an eigenvector $\mathbf{z} \in \mathbb{C}^n$ with eigenvalue λ.

We will evaluate $\overline{\mathbf{z}}^T A\mathbf{z}$ in two ways. Note that it is a 1×1 matrix over $\mathbb{C}$, that is, a complex number. On the one hand,

$$\overline{\mathbf{z}}^T A\mathbf{z} = \overline{\mathbf{z}}^T \lambda\mathbf{z} = \lambda\overline{\mathbf{z}}^T\mathbf{z} = \lambda\|\mathbf{z}\|^2$$

where the last step is by Lemma E5.4. On the other hand,

$$\begin{aligned}
\overline{\mathbf{z}}^T A\mathbf{z} &= \overline{\mathbf{z}}^T\overline{A}^T\mathbf{z} && \text{by hypothesis}\\
&= (\overline{A}\,\overline{\mathbf{z}})^T\mathbf{z} && \text{by Lemma A5.8(iii)}\\
&= (\overline{A\mathbf{z}})^T\mathbf{z} && \text{by Lemma E5.3}\\
&= (\overline{\lambda\mathbf{z}})^T\mathbf{z} && \text{since } A\mathbf{z} = \lambda\mathbf{z}\\
&= \overline{\lambda}\,\overline{\mathbf{z}}^T\mathbf{z} && \text{directly from the definitions}\\
&= \overline{\lambda}\|\mathbf{z}\|^2 && \text{by Lemma E5.4.}
\end{aligned}$$

Putting together the two expressions for $\overline{\mathbf{z}}^T A\mathbf{z}$ gives $\lambda\|\mathbf{z}\|^2 = \overline{\lambda}\|\mathbf{z}\|^2$. But $\mathbf{z} \neq \mathbf{0}$ by definition of eigenvector, so $z_i \neq 0$ for some i, so $\|\mathbf{z}\|^2 = \sum_{i=1}^n |z_i|^2 > 0$. Hence $\lambda = \overline{\lambda}$, or equivalently, λ is real. □

We now use what we know about *complex* matrices to prove a result that is purely about *real* matrices.

Theorem E5.6 *Every real symmetric matrix has at least one real eigenvalue.*

Proof Let A be a real symmetric $n \times n$ matrix (where $n \geq 1$). By Theorem E5.2, A has at least one complex eigenvalue λ. Now $A = \overline{A} = \overline{A}^T$ since A is real and symmetric, so λ is real by Proposition E5.5. □

The significance of this result will be revealed in the next and final section.

E6 Symmetric matrices

This whole course has been about the interplay of geometry and algebra. We finish with a beautiful result stating that a certain geometric condition on linear operators is equivalent to a certain algebraic condition on matrices.

By definition, a linear operator T is diagonalizable if and only if its matrix with respect to some basis of $\mathbb{R}^n$ is diagonal. A typical diagonalizable operator is shown in Figure E.3. But we can be more demanding and ask whether the matrix of T with respect to some *orthonormal* basis of $\mathbb{R}^n$ is diagonal. Such an operator T is called **orthogonally diagonalizable**. Figures E.2 and E.5 both show examples.

Why should we care about *orthogonal* diagonalizability? It is because the orthonormal bases are the most convenient ones, as we saw in Section B6. They share many of the properties of the standard basis, which is the one we're most

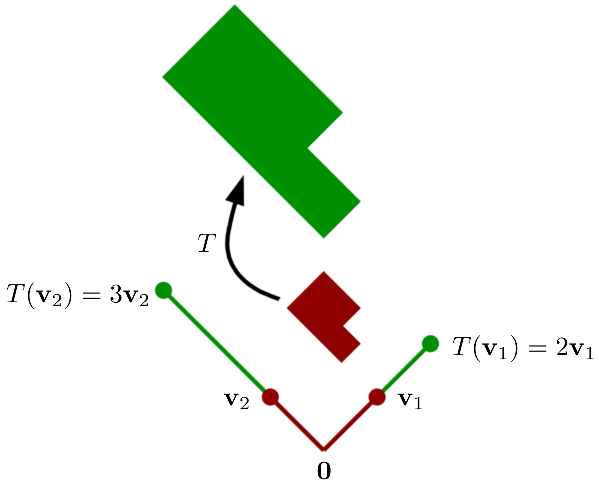

T
$T(\mathbf{v}_2) = 3\mathbf{v}_2$
$T(\mathbf{v}_1) = 2\mathbf{v}_1$
$\mathbf{v}_2$
$\mathbf{v}_1$
$\mathbf{0}$

Figure E.5: An orthogonally diagonalizable operator T on $\mathbb{R}^2$

familiar with. For instance, Lemma B6.6 shows how easy it is to express any given vector as a linear combination of orthonormal basis vectors.

So if we're dealing with an operator T, and we're able to find an orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ such that the matrix of T with respect to $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is orthonormal, then the situation is almost as easy as if the standard matrix of T were diagonal. And as we saw at the start of Section E3, that situation is very easy indeed.

But given an operator, how can you tell whether it is orthogonally diagonalizable? We will prove the major result—perhaps the highlight of the course—that an operator is orthogonally diagonalizable if and only if its standard matrix is symmetric.

The proof will take us most of the rest of this section. We will mostly work with matrices rather than operators, so we will need the following definition.

Definition E6.1 A real square matrix A is **orthogonally diagonalizable** if there exists an orthogonal matrix P such that $P^{-1}AP$ is diagonal.

Since the inverse of an orthogonal matrix is the same as its transpose, we could equivalently have written $P^T AP$ instead of $P^{-1}AP$.

Just as ordinary diagonalizability is equivalent to the existence of a basis of eigenvectors, *orthogonal* diagonalizability is equivalent to the existence of an *orthonormal* basis of eigenvectors:

Proposition E6.2 *Let T be a linear operator on $\mathbb{R}^n$. The following are equivalent:*

i. T is orthogonally diagonalizable;

ii. there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of T;

iii. there exist n orthonormal eigenvectors of T;

iv. the standard matrix of T is orthogonally diagonalizable.

Proof This is almost identical to the proof of Proposition E3.4, inserting the word 'orthogonal' or 'orthonormal' in appropriate places. □

In particular, the operator T is orthogonally diagonalizable if and only if the matrix $[T]$ is orthogonally diagonalizable. So, it makes little difference whether we work with operators or matrices. From now on, we stick with matrices.

Proposition E6.3 *Let A be an $n \times n$ matrix. The following are equivalent:*

i. there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A;

ii. there exist n orthonormal eigenvectors of A;

iii. A is orthogonally diagonalizable.

Proof This follows from Proposition E6.2 in exactly the same way that Proposition E3.6 followed from Proposition E3.4. □

As in the case of ordinary diagonalizability, if $P^{-1}AP = D$ with P orthogonal and D diagonal, then the columns of P form an orthonormal basis of eigenvectors of A and the diagonal entries of D are the corresponding eigenvalues.

Example E6.4 Let

$$A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}.$$

Then $\chi_A(\lambda) = (\lambda - 4)(\lambda - 6)$, an eigenvector with eigenvalue 4 is $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$, and an eigenvector with eigenvalue 6 is $\left(\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}\right)$. These eigenvectors are orthogonal but not orthonormal. We scale them to make them orthonormal, and define P to be the matrix with these orthonormal eigenvectors as its columns:

$$P = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}.$$

Then P is orthogonal, $P^TAP = \text{diag}(4, 6)$, and A is orthogonally diagonalizable.

Notice that in this example, the orthogonally diagonalizable matrix A was symmetric. In fact:

Lemma E6.5 *Every orthogonally diagonalizable matrix is symmetric.*

Proof Let A be an orthogonally diagonalizable matrix. Then $P^{-1}AP = D$ for some orthogonal matrix P and diagonal matrix D. We have $A = PDP^{-1} = PDP^T$, and D is symmetric, so

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A.$$

Hence $A^T = A$, that is, A is symmetric. □

We now prove the converse, which is much harder. It is one of several related results called a 'spectral theorem' (because it has something to do with the spectrum).

Theorem E6.6 (Spectral theorem) *A real square matrix is orthogonally diagonalizable if and only if it is symmetric.*

Proof We have just proved 'only if'. For 'if', we use induction on the size of the matrix. It is clear for 1×1 matrices. Now let $n \geq 2$, let A be an $n \times n$ symmetric matrix, and suppose inductively that $(n-1) \times (n-1)$ symmetric matrices are orthogonally diagonalizable.

By Theorem E5.6 (which we proved using linear algebra over $\mathbb{C}$!), A has at least one real eigenvalue. Thus, $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ for some $\lambda_1 \in \mathbb{R}$ and vector $\mathbf{v}_1 \neq \mathbf{0}$, which we may assume to have length 1. By the orthonormal extension lemma (Lemma B7.6), we can extend $\mathbf{v}_1$ to an orthonormal basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of $\mathbb{R}^n$.

(This is *not* going to be an orthonormal basis of eigenvectors of A, unless we're very lucky! So we have more work to do.)

Put $Q = (\mathbf{v}_1|\mathbf{v}_2|\cdots|\mathbf{v}_n)$. Then Q has orthonormal columns, so it is orthogonal. Now consider the matrix $Q^T A Q = Q^{-1} A Q$. By Lemma A4.3(ii), $Q\mathbf{e}_1 = \mathbf{v}_1$, so the equation $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$ becomes $AQ\mathbf{e}_1 = \lambda_1 Q\mathbf{e}_1$, or equivalently $Q^{-1}AQ\mathbf{e}_1 = \lambda_1 \mathbf{e}_1$. Again by Lemma A4.3(ii), this means that the first column of $Q^T A Q$ is $\lambda_1 \mathbf{e}_1$. Moreover, $Q^T A Q$ is symmetric, since

$$(Q^T A Q)^T = Q^T A^T (Q^T)^T = Q^T A Q$$

(using the symmetry of A). So the first row of $Q^T A Q$ is $\lambda_1 \mathbf{e}_1^T$. Hence

$$Q^T A Q = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & a'_{11} & \cdots & a'_{1,n-1} \\ \vdots & \vdots & & \vdots \\ 0 & a'_{n-1,1} & \cdots & a'_{n-1,n-1} \end{pmatrix} \tag{E:2}$$

for some real numbers a'_{ij}. Write A' for the $(n-1) \times (n-1)$ matrix (a'_{ij}). Since $Q^T A Q$ is symmetric, so is A'. By inductive hypothesis, A' is orthogonally diagonalizable, so $Q'^T A' Q' = \operatorname{diag}(\lambda_2, \ldots, \lambda_n)$ for some $(n-1) \times (n-1)$ orthogonal matrix Q' and $\lambda_2, \ldots, \lambda_n \in \mathbb{R}$.

For convenience, let us write equation (E:2) as

$$Q^T A Q = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right).$$

Define an $n \times n$ matrix R by

$$R = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q' & \\ 0 & & & \end{array}\right).$$

(The idea now is that Q' diagonalizes A', and $Q^T A Q$ is A' in the last $n-1$ coordinates, so R will diagonalize $Q^T A Q$. This will imply that QR diagonalizes A, and we will also be able to show that QR is orthogonal.)

We have

$$\begin{aligned}
R^T(Q^TAQ)R &= \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q' & \\ 0 & & & \end{array}\right)^T \left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A' & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q' & \\ 0 & & & \end{array}\right) \\
&= \left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q'^TA'Q' & \\ 0 & & & \end{array}\right) \\
&= \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n).
\end{aligned}$$

Put $P = QR$; then P^TAP is diagonal.

It remains to prove that P is orthogonal. Indeed, Q is orthogonal (as noted above), and R is orthogonal since

$$R^TR = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q' & \\ 0 & & & \end{array}\right)^T \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q' & \\ 0 & & & \end{array}\right) = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q'^TQ' & \\ 0 & & & \end{array}\right) = I_n.$$

So P is the product of two orthogonal matrices, and is therefore orthogonal by Lemma E4.3. This completes the induction. □

Remark E6.7 This proof is long and may seem rather technical. However, if we had the language of vector spaces available, we could rephrase the proof in such a way that it only took a few lines and was geometrically intuitive. You will learn about vector spaces in Honours Algebra next year.

We'll finish with an example that looks very hard, but which the theory we've developed makes easy.

Example E6.8 Let

$$A = \begin{pmatrix} 14 & -14 & -16 \\ -14 & 23 & -2 \\ -16 & -2 & 8 \end{pmatrix}.$$

What is A^{99}?

We observed at the beginning of Section E3 that powers of diagonal matrices are very easy to calculate. Our matrix A is not diagonal, but it is symmetric, so it must be orthogonally diagonal*izable*.

We calculate that A has characteristic polynomial

$$\chi_A(\lambda) = -(\lambda + 9)(\lambda - 36)(\lambda - 18),$$

so its eigenvalues are -9, 36 and 18. Using row reduction in the usual way, we find an eigenvector $\left(\begin{smallmatrix} 2 \\ 1 \\ 2 \end{smallmatrix}\right)$ with eigenvalue -9, and normalizing so that it has length 1 gives the eigenvector $\left(\begin{smallmatrix} 2/3 \\ 1/3 \\ 2/3 \end{smallmatrix}\right)$. Similarly, we find unit-length eigenvectors

$\begin{pmatrix} -2/3 \\ 2/3 \\ 1/3 \end{pmatrix}$ and $\begin{pmatrix} 1/3 \\ 2/3 \\ -2/3 \end{pmatrix}$ with eigenvalues 36 and 18, respectively. These unit-length eigenvectors are orthonormal, so if we put

$$P = \begin{pmatrix} 2/3 & -2/3 & 1/3 \\ 1/3 & 2/3 & 2/3 \\ 2/3 & 1/3 & -2/3 \end{pmatrix}$$

then P is orthogonal and $P^T AP = D$, where $D = \operatorname{diag}(-9, 36, 18)$.

It follows that $A = PDP^T = PDP^{-1}$. Hence

$$\begin{aligned} A^{99} &= (PDP^{-1})(PDP^{-1}) \cdots (PDP^{-1}) \\ &= PD^{99}P^{-1} \\ &= \begin{pmatrix} 2/3 & -2/3 & 1/3 \\ 1/3 & 2/3 & 2/3 \\ 2/3 & 1/3 & -2/3 \end{pmatrix} \begin{pmatrix} (-9)^{99} & 0 & 0 \\ 0 & 36^{99} & 0 \\ 0 & 0 & 18^{99} \end{pmatrix} \begin{pmatrix} 2/3 & 1/3 & 2/3 \\ -2/3 & 2/3 & 1/3 \\ 1/3 & 2/3 & -2/3 \end{pmatrix} \end{aligned}$$

by cancelling and using the fact that $P^{-1} = P^T$. Multiplying this out, we conclude that A^{99} is equal to

$$\frac{1}{9} \begin{pmatrix} 4 \cdot (-9)^{99} + 4 \cdot 36^{99} + 18^{99} & 2 \cdot (-9)^{99} - 4 \cdot 36^{99} + 2 \cdot 18^{99} & 4 \cdot (-9)^{99} - 2 \cdot 36^{99} - 2 \cdot 18^{99} \\ 2 \cdot (-9)^{99} - 4 \cdot 36^{99} + 2 \cdot 18^{99} & (-9)^{99} + 4 \cdot 36^{99} + 4 \cdot 18^{99} & 2 \cdot (-9)^{99} + 2 \cdot 36^{99} - 4 \cdot 18^{99} \\ 4 \cdot (-9)^{99} - 2 \cdot 36^{99} - 2 \cdot 18^{99} & 2 \cdot (-9)^{99} + 2 \cdot 36^{99} - 4 \cdot 18^{99} & 4 \cdot (-9)^{99} + 36^{99} + 4 \cdot 18^{99} \end{pmatrix}.$$

Of course, what we've really just done is found a formula for A^r for all positive integers r.

Remark E6.9 We can use the same technique to find the powers of any diagonalizable matrix. In this example, our matrix was *orthogonally* diagonalizable (that is, symmetric), which made it very easy to find the inverse of P: it's just the transpose. But we already have a general algorithm for calculating inverses (Section C6), so we could have found P^{-1} anyway, even if P hadn't been orthogonal. You will get some practice on this in the remaining workshops.

Could you have found the 99th power of this matrix without all the theory we developed during this course? Maybe not!

Summary of Chapter E

This is for you to fill in.

The most important definitions and ideas in this chapter

The most important results in this chapter

Points I didn't understand

Summary of the whole course

This is for you to fill in.

The most important definitions and ideas in the course

The most important results in the course

Points I didn't understand