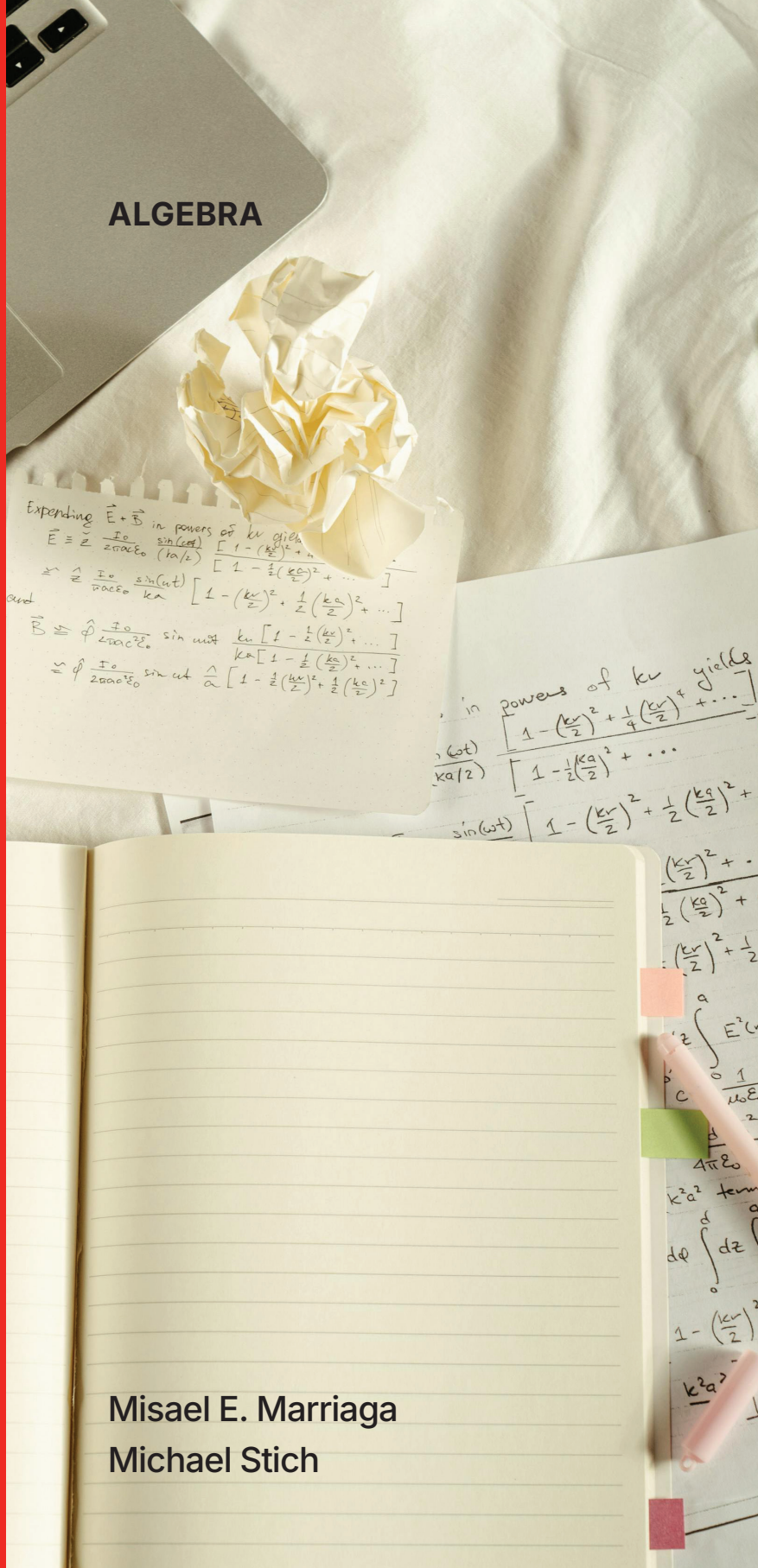
Universidad Rey Juan Carlos

ALGEBRA



Material docente

Misael E. Marriaga
Michael Stich

ACADEMIA ABIERTA

Esta monografía es un proyecto de la editorial Academia Abierta

Universidad
Rey Juan Carlos

Colección
Material docente

Algebra
Lecture notes

ACADEMIA ABIERTA

Autores
Misael E. Marriaga
Michael Stich

Biomedical Engineering
September 05, 2023

Esta monografía ha sido publicada gracias a la financiación de la convocatoria de asignaturas en abierto promovida por la Oficina de Conocimiento y Cultura Libres de la Universidad Rey Juan Carlos.

Cómo citar
Marriaga, M. M., & Stich, M. (2023). *Álgebra. Lectures notes*. Editorial Academia Abierta. Recuperado a partir de https://monografias.urjc.es/index.php/omp-urjc/catalog/book/41

Depósito
Archivo abierto de la URJC (BURJC digital)

DOI: https://doi.org/10.33732/MD-41

Editorial
Academia Abierta. Servicio de Publicaciones de la URJC
C/Tulipán s/n
28933 Móstoles
https://monografias.urjc.es/index.php/omp-urjc

Diseño de interiores y portada
Glaux Publicaciones Académicas, SLU

Contacto editorial
Laura de la Cruz / Tomás Zarza
servicio.publicaciones@urjc.es

TABLE OF CONTENTS

INTRODUCTION TO MATHEMATICAL THINKING

Linear algebra can be viewed as the mathematical apparatus needed to solve systems of linear equations, understand their underlying structure, and to applied what is learned in other contexts. Unlike your first brush with the subject in high school, which probably emphasized matrices and Euclidean spaces like the real plane $\mathbb{R}^2$, we will focus on abstract vector spaces and linear transformations. These terms will be defined later, so do not worry if you are not familiar with them. These notes start from the beginning of the subject, assuming no knowledge for linear algebra. The key point is that you are about to immerse yourself in serious mathematics, with an emphasis on your attaining a deep understanding of the definitions, theorems, and proofs.

Upper-level mathematics has much in common with high school mathematics, and students who have been accepted onto a Biomedical Engineering major already have an array of mathematical skills that will serve them well. On the other hand, upper-level mathematics also differs in some important respects. This means that most students need to extend and adapt their existing skills in order to continue to do well. Making such extensions and adaptations can be difficult for those who have never really reflected on the nature of their skills.

One thing you certainly have learned to do is to apply mathematical procedures to calculate answers to standard questions, perhaps by matching key words in the statement of a question with a memorized recipe of calculations from within a set of many recipes. This might have been useful to achieve a good grade in the standardized entry exam to university, which is a controlled and predictable environment. Moreover, some people enjoy this type of work. They like the satisfaction of arriving at a page of correct answers, and they like the security of knowing that if they do everything right then their answers will, indeed, be correct. Sometimes they compare mathematics favorably with other subjects in which things seem to be more a matter of opinion and "there are no right answers."

Other people dislike this aspect of mathematics. They find it dull to do lots of repetitious exercises, and they get more satisfaction from learning about *why* the various procedures work and how they fit together. However, note that knowing how to apply procedures is extremely important because, without fluency in calculations, it is hard to focus your attention on higher-level concepts.

When you start taking upper-level courses, professors will expect you to be fluent in using the procedures you have already learned. They will expect you to be able to accurately manipulate algebraic expressions, to solve equations, and so on. They will expect you to be able to do theses things without having

to stop each time to look up a rule, and they might not be patient with students who are not able to do so. This is not because they are impatient with students in general – most professors will be happy to spend a long time talking with you about new mathematics, or responding to students who say, "I know how to do this, but I've never understood why we do it this way." But they will not expect to re-teach things you have already studied. So you should brush up your knowledge prior to beginning a course, especially if, say, you've done no mathematics all summer.

Once you do begin, you'll find that some upper-level mathematics involves learning new procedures. These procedures, unsurprisingly, will be longer and more complicated than those you met in earlier work. We are not worried about your ability to apply long and complicated procedures, however, because to have got this far you must be able to do that kind of thing. Here, we want to focus on more substantive changes in the ways in which you have to interact with the procedures.

The first substantive difference is that you will have more responsibility for deciding which procedure to apply. Of course, you have learned to do this to some extent already. For instance, you have learned how to multiply out brackets and write things like:

$$(x+2)(x-5) = x^2 - 3x - 10.$$

But hopefully you have also learned that it is *not* sensible to multiply out when trying to simplify a fraction like this:

$$\frac{x^2(x+2)(x-5)}{x^2+2x}.$$

For the fraction, simplification is easier if you keep the factors "visible." Nonetheless, many students automatically multiply out, probably because multiplying out was one of the first things they learned to do when studying algebra. They would, however, be more successful in mathematics courses like calculus and linear algebra if they learned to stop and think first about what would allow them to make the most progress. If this doesn't apply to you for this particular type of problem, does it apply in others? Have you ever done a long calculation and then realized that you didn't need to? Could you have avoided it if you'd stopped to think first? Part of deciding which procedures to apply is giving yourself a moment to think about it before you leap in and do the first thing that comes to mind.

This might not sound like a big deal, but think for a moment about how often you *don't* have to make a choice about what procedure to apply. Often, questions

in books or on tests tell you exactly what to do. They say things like, "Use Cramer's rule to solve this system of equations." Even when a question doesn't tell you outright, it is sometimes obvious from the context. In high school, if your teacher spent a lesson showing you how to apply the Rouché-Frobenius theorem, then gave you a set of questions to do, it was probably safe for you to assume that these would involve automatically applying the Rouché-Frobenius theorem. This helped you out, but it means that much of the time you didn't have to decide what procedure to apply. In the wider world, and in advanced mathematics, making decisions is more highly valued and more often expected. This means that questions presented to you on problem sheets and exams will usually just say "solve this problem" rather than "solve this problem using this procedure."

Another part of deciding which procedures to apply is being able to distinguish between cases that look similar but are best approached in different ways. For example, consider integration, and more specifically integration by parts. You may know that this is used when we want to integrate a product of two functions, one of which gets simpler when we differentiate it, and the other of which does not get any more complicated when we integrate it. For instance, in $\int xe^x dx$, x gets simpler if we differentiate it, and e^x gets no more complicated if we integrate it. You might also know, however, that sometimes mathematical situations look superficially similar, but are best tackled using different procedures. In the integration case, integration by substitution might sometimes be more appropriate. For instance, in $\int xe^{x^2} dx$, we would probably want to use integration by substitution instead of by parts. Can you see why?

Integration by substitution is a good case for another point we would like to make, this time about making decisions *within* procedures. It might be that you read the end of the last paragraph and thought, "But what substitution should I use?" Perhaps your teachers or books always told you what to use, but we would argue that they shouldn't necessarily have to. After a while, you should notice that certain substitutions are useful in certain cases. If you pay attention to the structures of these cases then, even if you wouldn't be sure that you could pick a good substitution for a new case, you should have an idea of some sensible things to try. If you haven't deliberately thought about this before, we suggest you do so now. Get out some questions on integration by substitution and, without actually doing the problems, look at the suggested substitutions. Can you anticipate why those substitutions will work? Can you then anticipate what would work in similar cases?

So how can a student improve their ability to make decisions about and within procedures? We have two suggestions. The first is to try doing exercises from a source where the procedure to be applied is not obvious. A good place to look is in the books listed in the bibliography section of the teaching guide. The second suggestion is to turn ordinary exercises into opportunities for

reflection. When you finish an exercise, instead of just moving on to the next one, stop and think about these questions:

1. Why did that procedure work?
2. What could be changed in the question so that it would still work?
3. What could be changed in the question so that it would *not* work?
4. Could I modify the procedure so that it would work for some of these cases?

All of these questions should help you build flexibility in applying what you know.

We want to come back to the idea that knowing how to apply such procedures is only one part of understanding mathematics. It is an important part, but most students can recognize the difference between learning to apply a procedure mechanically, and understanding why it works. Learning mechanically has some advantages: it is generally quick and relatively straightforward. But it also has disadvantages: if you learn procedures mechanically, it is easier to forget them, to misapply them, and to mix them up (and it will be harder for you to be as successful in this course as you would like to be). Developing a proper understanding of why things work is generally harder and more time-consuming, but the resulting knowledge is easier to remember and more supportive of flexible and accurate reasoning (and will certainly help you get a good grade in this course).

To summarize, before taking upper-level courses like calculus and algebra, you should brush up you knowledge of standard procedures because your professors will expect you to be able to use these fluently. As you progress, you will be expected to take more responsibility for deciding which procedure to apply; it might be a good idea to practice this by working on exercises from sources that do not tell you exactly what to do. You will also be expected to adapt procedures in sensible ways, and to work out how theorems or definitions can be applied without necessarily having seen many worked examples. You will not succeed in upper-level mathematics if you always try to solve problems by finding something that looks similar and copying it. You have to be more thoughtful than that. Mathematics is not just about procedures. Fluency with procedures is important, but in many cases you should aim for a deeper understanding of why procedures work.

Definitions

We will encounter many *definitions*, *theorems*, and *proofs* throughout these notes.

A definition has nothing to do with something being true. It just tells us what a mathematical word means. A definition might define a type of object, or it might define a property. The definition about defines a property about linear equations.

You already know that a definition tells us what a word means because you've been using dictionaries for a very long time. But there are two very, very important differences between dictionary definitions and mathematical definitions. If you want to understand the material in these notes, it is vital that you understand these differences.

The first difference is that when mathematicians (in particular, your algebra professor) state a definition, *they really mean it*. They don't mean that this is a good description of the majority of cases, but that there might be exceptions out there somewhere. This is not how dictionary definitions work. If you took two dictionaries and looked up an everyday concept, either a concrete one (like "table") or an abstract one (like "justice"), would you expect the definitions to be identical? Probably not. Probably, in fact, you would expect to be able to find things in the world that satisfy one definition but not the other. Or to find something that you would want to call a "table" or "justice" but that didn't really satisfy either definition*. Or to find something for which, even though there is a definition, people would disagree.

Now, compare this with what would happen if you took a mathematics book and looked up a definition of "even number." Would you expect to be able to find an even number that did not satisfy the definition? Or an non-even number that did, nonetheless, satisfy the definition? Absolutely not. Exceptions simply don't exist. It's not like you could find a number so big that it could be even without being divisible by 2. If you took two textbooks, you might not expect the two definitions to be phrased in exactly the same way, but you would expect them to be logically equivalent. That is, you would expect that all and only the things that satisfy the first definition would also satisfy the second. So, that's one difference: mathematical definitions mean exactly, *exactly* what they say. There are no exceptions, and different phrasings might exist but these must be logically equivalent.

The other difference, which is partly a consequence of the first, is that mathematical definitions are precise and operable in a way that dictionary definitions are not. This means that they contain some information that we can actually manipulate in an algebraic or logical argument. For instance, consider the following simple definition:

---

* For example, is a 10-meter tall "table" at an art exhibition a table? Sort of, but it wouldn't satisfy functional criteria such as being a flat surface you could rest things on – no-one would be able to reach, and anyway touching the exhibit is probably not allowed.

Definition: A number is *even* if it is divisible by 2.

"Well, yes," you'll be thinking, "obviously." In fact, though, this is probably not how your professor will write this definition. You're more likely to see something like this:

Definition: A number n is *even* if and only if there is an integer k such that $n = 2k$.

You could be forgiven for thinking that this over-complicates things. But it has some advantages. The first is precision. We can see this by comparing with the kind of thing that students sometimes write when they try to define *even*. They tend to say things like "Even is when it's divisible by 2." Clearly that captures the key idea, but it isn't very precise. For a start, what is "it"? Clearly "it" is supposed to be a number, but this number isn't introduced properly. In contrast, the better definition tells us explicitly that we are dealing with a number and gives it a name, n. For another thing, the student definition contains the locution "is when." This tends to sound clunky, in mathematics but also in other fields. In mathematics, if you find yourself writing "it" or "is when," you should probably consider rephrasing.

The second advantage is operability. We can operate with the mathematical definition to prove things. The better definition gives us a way of capturing and manipulating even numbers because it states what divisibility means in algebraic terms. We could, for instance, use it to prove that any integer multiple of an even number must also be even, perhaps by writing something like this:

Suppose that n is an even number. Then (by definition) there is an integer k such that $n = 2k$. Now, let z be any integer and $y = zn$. Then $y = z(2k)$. But we can rewrite this as $y = 2(zk)$. Now zk is an integer, so y is even because it can be written in the required form.

We could have the same ideas if we were working with the imprecise student definition. But the better version gives us a leg-up for writing arguments like this by providing some notation. Moreover, we expect the precise definition to *exclude* all the numbers that are not even (for the number 3, there is no appropriate integer k). In this way, the definition of even number captures the notion of evenness in a reasonable way.

A final point here is that we do not "prove" definitions. We can't, because there's nothing to prove: definitions just capture conventions in which we all agree to use a word to mean exactly the same thing. A professor might, at some point, explain to you how a definition captures an intuitive idea, but this isn't the same as proving it.

Theorems

Whereas a definition states what we mean by a word describing a mathematical object or property, a theorem tells us about a relationship between two or more types of objects and properties; it assumes that we already know what those objects and properties are. Of course, if you don't know the meanings of all the words and symbols in the statement of a theorem you will probably feel that you don't understand the theorem. Notice, however, that even if you don't understand a theorem, you should be able to recognize that it has a theorem-like structure of the form:

If this thing is true, *then* this other thing is true as well.

The bit that goes with the *if* is called the premise or the assumption or the hypothesis. The bit that goes with the *then* is called the conclusion. It's actually not quite this simple in real life, because there are a few ways of phrasing theorems that don't make the "if. . . then. . ." structure so obvious.

Theorems tell us true things about relationships between concepts. Nevertheless, we need to go to the trouble of proving them. Sometimes we do this because theorems are far from obvious. Sometimes, though, theorems are pretty obviously true, and we do it for the more subtle reason that proving them allows us to see how they all fit together to form a coherent theory.

Once definitions are settled, theorems follow by logical necessity. For instance, once we have decided to define even numbers as numbers divisible by 2, we must conclude that any integer multiple of an even number is also even. We don't have any choice about that kind of consequence.

We should say that although we've only used the term *theorem*, there are several words that sometimes get used instead. Some of these are *proposition*, *lemma*, *claim*, *corollary*, and the apparently all-encompassing *result*. *Lemma* is usually used for a small theorem that will then be used to prove a bigger, more important one. *Corollary* is used for a result that follows as a fairly immediate consequence of a big theorem. The other terms are more or less interchangeable (in our opinion).

Thinking about objects can allow us to link new statements to our existing knowledge. However, we can also work towards understanding statements by thinking about their logical structures. To do so, we need to pay attention to mathematical uses of logical language. We will start by looking more carefully at ways in which the word *if* is used in mathematics.

First, consider a statement "If A then B." This is sometimes written as "$A \Rightarrow B$," which is read out loud as "A implies B." The use of an arrow makes

it clearer that we can think of this statement as having a direction, which is important because we might have a situation in which $A \Rightarrow B$ is a true statement but $B \Rightarrow A$ is not. Sometimes both implications do hold, as in the case:

$$x \text{ is even} \Rightarrow x^2 \text{ is even (true)}$$
$$x^2 \text{ is even} \Rightarrow x \text{ is even (true)}$$

However, sometimes one is true but the other is not, as in the case:

$$x < 2 \Rightarrow x < 5 \text{ (true)}$$
$$x < 5 \Rightarrow x < 2 \text{ (false)}$$

These statements are actually somewhat imprecise, because we have not specified what type of object x is. You probably assumed that it must be an integer in the squaring case and a real number in the inequalities case, since that would make sense. But, to be clearer, we could write things like:

$$\text{For every } x \in \mathbb{R},\ x < 5 \Rightarrow x < 2.$$

Doing so makes it easier to see why this particular statement is false: there are some real numbers that are less than 5 but not less than 2. But mathematicians sometimes omit such phrases when writing in note form or when the intended interpretation is obvious.

When they want to discuss both implications at once, mathematicians use a double-headed arrow meaning "is equivalent to," or write "if and only if" or its abbreviation "iff." So these are different ways of writing the same (true) statement about integers:

$$x \text{ is even} \Leftrightarrow x^2 \text{ is even.}$$
$$x \text{ is even if and only if } x^2 \text{ is even.}$$
$$x \text{ is even iff } x^2 \text{ is even.}$$

We have seen the phrase "if and only if" before, in our definitions. Here it is again:

Definition: A number n is *even* if and only if there is an integer k such that $n = 2k$.

To think about the phrase, it might be illuminating to split up this definition and write each implication separately:

A number n is even *if* there exists an integer k such that $n = 2k$.

A number n is even *only if* there exists an integer k such that $n = 2k$.

Can you see how this definition "catches" the numbers that are even, and excludes those that are not?

One final thing to note about a statement of the form "*A* if and only if *B*" is that, if we want to prove it, we can take one of two approaches. We can either construct a proof in which all the lines are equivalent to each other, or prove the two statements $A \Rightarrow B$ and $B \Rightarrow A$ separately.

The uses of "if" and " $\Rightarrow$ " sound straightforward when we are considering simple mathematical statements. But we would like to draw your attention to two potential sources of confusion.

First, some thought is necessary to sort out which of the "if" and the "only if" corresponds to which implication. You should think about this, perhaps by thinking about which one could replace the "implies" arrow in these two versions of the same (true) statement:

$$x < 2 \Rightarrow x < 5 \quad x < 5 \Leftarrow x < 2.$$

Second, it turns out that students do *not* always interpret "if" in a mathematical way in everyday life. In everyday conversation, we tend to speak rather imprecisely, relying on the context to help our listener make the interpretation we intend. For instance, imagine someone says to you,

If you clean the car then you can go out on Friday night.

You could reasonably infer from this that if you don't clean the car then you can't go out Friday night. Clearly that is what the speaker intends. And someone else could infer that if you were allowed out on Friday, they you must have cleaned the car. But, in fact, *neither of these is logically equivalent to the original statement*. Perhaps the easiest way to see this is to look at the logic in parallel with our simple statements about inequalities:

clean car $\Rightarrow$ out Friday	$x < 2 \Rightarrow x < 5$
not clean car $\Rightarrow$ not out Friday	$x \geq 2 \Rightarrow x \geq 5$
clean car $\Leftarrow$ out Friday	$x < 2 \Leftarrow x < 5.$

The second and third statements in each case are not logically the same as the first. Alternatively, you could think about the everyday situation, and see that the original statement says nothing at all about what happens if you don't clean the car, so there would be no contradiction if you didn't clean it but you were still allowed out. Technically, the person bargaining with you should really say:

You can go out on Friday night *if and only if* you clean the car.

Of course, no one talks like that. Which means that you might have less practice than you think accurately interpreting logical statements. Using logical language in a mathematically correct sense is not too hard, thought, because there are cases in which the intended natural language interpretation *is* the same as the mathematical one. Consider the statement:

If Juan is from Sevilla then Juan is from Andalucía.

No one hearing this would dream of inferring either of these:

If Juan is not from Sevilla then Juan is not from Andalucía.

If Juan is from Andalucía then Juan is from Sevilla.

But these inferences are analogous to those we looked at for the statement about cleaning the car. Make sure you can see how.

In the Sevilla case, the natural interpretation of the statement is the same as the mathematical one. We can also use it to illustrate a general point about logical equivalence of different implications. For any statement of the form $A \Rightarrow B$ we can consider three related statements called its *converse*, *inverse*, and *contrapositive*. This is what each one means, using the Sevilla example for illustration.

original	$A \Rightarrow B$	from Sevilla $\Rightarrow$ from Andalucía
converse	$B \Rightarrow A$	from Andalucía $\Rightarrow$ from Sevilla
inverse	not $A \Rightarrow$ not B	not from Sevilla $\Rightarrow$ not from Andalucía
contrapositive	not $B \Rightarrow$ not A	not from Andalucía $\Rightarrow$ not from Sevilla.

This is what each one means using one of our simple mathematical example instead:

original	$A \Rightarrow B$	$x < 2 \Rightarrow x < 5$
converse	$B \Rightarrow A$	$x < 5 \Rightarrow x < 2$
inverse	not $A \Rightarrow$ not B	$x \geq 2 \Rightarrow x \geq 5$
contrapositive	not $B \Rightarrow$ not A	$x \geq 5 \Rightarrow x \geq 2$.

This should help you remember that if $A \Rightarrow B$ is a true statement, then its contrapositive will also be true (in fact, they are logically equivalent), but its inverse and its converse might not be.

Finally, I should point out that your professor will be careful about uses of "if" and "if and only if" in theorems and proofs, but perhaps not so careful in

definitions. In a definition, they might just write "if" instead of "if and only if." This works for the same reason that everyday communication works: everyone knows that, in a definition, this is what is intended.

Next we want to discuss the phrases "for all" and "there exists." These are called *quantifiers*, because they tell us how many of something we're talking about. They are so common in mathematics that we have symbols for them: we use "$\forall$" (the universal quantifier) to mean "for all" and "$\exists$" (the existential quantifier) to mean "there exists."

In simple statements, quantifiers are easy to think about. Here is a simple quantifier statement:

$$\forall x \in \mathbb{Z}, x^2 \geq 0.$$

In this statement, we write $\forall x \in \mathbb{Z}$ to specify exactly which objects we are talking about. We could just write $\forall x$, and sometimes people do when it is obvious what kind of numbers (or other objects) a statement is about. But doing so could be ambiguous. In this case, the statement could just as well be about real or complex numbers, which raises issues of the truth or otherwise of the statement: "$\forall x \in \mathbb{Z}, x^2 \geq 0$" is true, "$\forall x \in \mathbb{C}, x^2 \geq 0$" is not. So it is good practice to be specific.

More complicated quantified statements can be harder to think about. Here is a definition written in words and in an abbreviated form using the new symbol (which might be more naturally read as "for every" in this case):

Definition: A function $f : \mathbb{R} \to \mathbb{R}$ is *increasing* if and only if for every $x_1, x_2 \in \mathbb{R}$ such that $x_1 < x_2$, we have $f(x_1) \leq f(x_2)$.

Definition (abbreviated): $f : \mathbb{R} \to \mathbb{R}$ is *increasing* if and only if $\forall x_1, x_2 \in \mathbb{R}$ such that $x_1 < x_2, f(x_1) \leq f(x_2)$

Here is a simple quantified statement involving the existential quantifier:

$$\exists x \in \mathbb{Z} \text{ such that } x^2 = 25.$$

This is a true statement, because when mathematicians say "there exist," they mean "there exists at least one." Here, there are two different integers x that satisfy the statement. In other cases, there might be hundreds. Students sometimes find it strange that we say "there exists" without specifying how many because, if we know exactly how many there are, it seems rude not to say so. However, advanced mathematics is at least partly about general relationships between concepts, rather than about finding "answers" as such. We can see other reasons why it makes sense to use "there exists" without extra

specification if we look at another definition (again shown both in words and in an abbreviated form):

Definition: A number n is *even* if and only if there exists an integer k such that $n = 2k$.

Definition (abbreviated): $n \in \mathbb{Z}$ is *even* if and only if $\exists k \in \mathbb{Z}$ such that $n = 2k$.

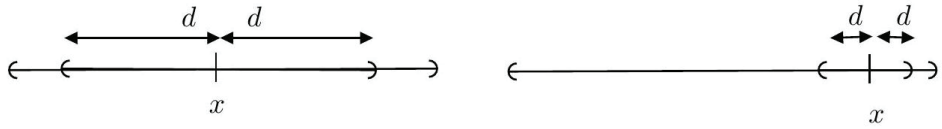
This definition gives us an agreed way of deciding whether a number is even or not. To make that decision, we do not care what the particular k is, we just care whether or not there is one. It also allows us to make general arguments about all even numbers. We might start by saying "Suppose n is even, so $\exists k \in \mathbb{Z}$ such that $n = 2k$." In this case, we don't want to specify what the k is because we want the ensuing argument to be general in the sense that it applies to any number that satisfies to the condition.

Some mathematical statements have more than one quantifier. The following definition (again with an abbreviated version) might be described as doubly quantified or as having two nested quantifiers:

Definition: A set $X \subseteq \mathbb{R}$ is *open* if and only if for every $x \in X$ there exists $d > 0$ such that $(x - d, x + d) \subseteq X$.

Definition (abbreviated): $X \subseteq \mathbb{R}$ is *open* if and only if $\forall x \in X \exists d > 0$ such that $(x - d, x + d) \subseteq X$.

When a statement has more than one quantifier, *the order in which they appear is really important*. In this case, the definition says, "for every x, there exists a d." We should imagine taking a particular x value and finding an appropriate d (maybe depending on x). If we take a different x, we might need a different d (perhaps a smaller one, as in the right-had diagram below).



If the definition instead said "there exists a d for every x," mathematicians would read that as meaning that we could select a single d (independent of x) that works for every x. That is not the same thing at all.

To check your understanding of this, consider the following two statements. One is true, and the other is false. Which is which?

$$\exists y > 0 \text{ such that } \forall x > 0, y < x.$$
$$\forall x > 0 \exists y > 0 \text{ such that } y < x.$$

It is normal to find this difficult, because everyday life would probably not distinguish between these two statements. Without even realizing it, we would make the interpretation that seems most realistic, disregarding logical correctness. Most students therefore have to concentrate for a while before they get the hang of reading what is literally there and making the mathematical correct interpretation.

Although we have written the theorems in the form "If...then ...," you might also see different phrasings. Here are some common ones:

Theorem: If f is an even function, then $\frac{df}{dx}$ is an odd function.

Theorem: Suppose that f is an even function. Then $\frac{df}{dx}$ is an odd function.

Theorem: Every even function has an odd derivative.

These would all be interpreted to mean the same thing: the premise in each case is that the function is even, and the conclusion is that its derivative is odd. It might seem strange that we don't just pick one form and stick to it but, sometimes, one version or another sounds more natural, so mathematicians like to have this flexibility.

You will also see theorems of different types. There are, for instance, *existence theorems* like this one:

Theorem: There exists a number x such that $x^3 = x$.

One way to prove a theorem like this is just to produce an object that satisfies it: the number 1 would do, in this case. It's not always that easy, but it's important to recognize that it might be, because sometimes students tie themselves in knots doing complicated things when a simple answer would do.

There are also theorems about non-existence, like this one:

Theorem: There does not exist a largest prime number.

In fact, a bit of thought, non-existence theorems can be restated in our standard form. This one, for instance, could be written with a universal quantifier:

Theorem: For every prime number n there exists a prime number p such that $p > n$.

Then it could be rephrased into our initial form:

Theorem: If n is a prime number then there exists another prime number p such that $p > n$.

These rephrasing possibilities can be very useful when we want to start proving something: sometimes rewriting in a different way can give us different ideas about sensible things to try. However, the fact that we can often rephrase does *not* mean that you can be sloppy about your mathematical writing. Sloppy paraphrasing can easily change the logical meaning of a statement. Once you become fluent in using logical language in a mathematical way, you will find that you can switch forms without doing violence to the meaning. Until that point you should think carefully about logical precision.

One great thing about having precise meanings for logical terms is that it buys us a lot of mechanistic reasoning power. For instance, if we know that $A \Rightarrow B$ and that $B \Rightarrow C$, then we can deduce that $A \Rightarrow C$. We can do this even if A, B, and C are about really complicated objects that we've never met before and we don't understand. Similarly, if we would like to prove a statement of the form $A \Rightarrow B$, be we are not making much progress, we can remember that the contrapositive (not $B \Rightarrow$ no A) is always equivalent to the original, and try proving that instead.

This is what we mean when we say that we can develop valuable understanding by looking at the logical structure of a statement. If we pay attention to constructions involving "if" or quantifier, we can make use of such regularities in our reasoning. This is (at least partly) what people mean when they talk about *formal* work: we can concentrate on the logical form of a sentence and temporarily ignore its meaningful content. We don't *have to* ignore the meaning, of course, and for most of the above discussion you were probably thinking about meanings as well. But attending to logical form is vital for proper understanding. Some students are lax about this; when reading mathematics, they look mostly at the symbols, ignoring or glossing over the words. This can make their understanding faulty and their writing inacurate, because they mix up important quantifiers or implications. For instance, consider the following theorem:

Rolle's Theorem: Suppose that $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b) and that $f(a) = f(b)$. Then $\exists c \in (a,b)$ such that $f'(c) = 0$.

It is quite common to see students make errors, writing things like this:

Rolle's Theorem: Suppose that $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b). Then $f(a) = f(b)$ and $\exists c \in (a,b)$ such that $f'(c) = 0$.

We can see that these are different by looking at their logical forms: one of the premises in the correct version appears instead as part of the conclusion in the second (look carefully to make sure that you can see this). Clearly that must make a very important difference, so the incorrect version cannot possibly be logically equivalent to the correct version. Nonetheless, it might still be a valid theorem. In this case, however, it isn't, which we can see by thinking in terms of examples. In the incorrect version, the premises introduce a function f that is defined on an interval and is continuous and differentiable on this interval. The conclusion claims that the function values are equal at the endpoints of the interval, but this cannot possibly follow in a valid way from the premises, because there are many functions and intervals that satisfy the premises but do not have this property. For example, $f(x) = x^2$ is continuous on [0,2] and differentiable on (0,2), but certainly isn't the case that $f(0) = f(2)$. We can see how a student might write the incorrect version in the first place, but someone who is thinking about the meaning of their writing should recognize such errors when rereading.

This brings us back to the idea that both logical form and example objects can contribute to mathematical understanding, though focusing on each has different advantages and disadvantages. If you look mainly at examples, you might feel that you understand, but you might fail to appreciate the full generality of a statement of find it difficult to see the logical structure of a whole course. If you look mainly at formal arguments, you might be able to see how everything fits together logically, but you might find yourself complaining that it is very abstract and that you don't really understand what is going on. Your experience of these issues will probably vary from course to course, because some professor give lots of examples and draw lots of diagrams, whereas others give a much more formal presentation. If a professor's approach does not match you preferred way of developing understanding, you might find it useful to strengthen your understanding of the links between the example objects and the formal work.

Proofs

You have constructing mathematical proofs for many years. For instance, to do the calculations needed to prove that the solutions to the equation $x^2 - 20x + 10 = 0$ are $10 + 3\sqrt{10}$ and $10 - 3\sqrt{10}$, you would write something like this:

$$x = \frac{20 \pm \sqrt{400 - 40}}{2} = \frac{20 \pm \sqrt{360}}{2} = \frac{20 \pm 6\sqrt{10}}{2} = 10 \pm 3\sqrt{10}.$$

This calculation uses methods that everyone agrees are valid, so it captures everything we need for a proof that the solutions really are as claimed. To make it look more like and upper-level proof, we could rewrite it like this:

Claim: If $x^2 - 20x + 10 = 0$ then $x = 10 + 3\sqrt{10}$ or $x = 10 - 3\sqrt{10}$.

Proof: Suppose that $x^2 - 20x + 10 = 0$. Then, using the quadratic formula,

$$x = \frac{20 \pm \sqrt{400 - 40}}{2} = 10 \pm 3\sqrt{10}.$$

This version explicitly states the claim, begins with the premise ("Suppose that $x^2 - 20x + 2 = 0$"), and contains few words to justify those steps that are more sophisticated or less obvious.

My point is that there is nothing inherently mysterious about proofs. It is certainly true that most high school mathematics is not presented in this way, but it's also true that most of it could be. We say this because, in upper-level courses, lots of the mathematics you meet will be presented in this form; these lecture notes will be full of theorems and proofs. This might seem rather an abrupt change, and some students get the idea that proof writing is a mysterious dark art to which only the very privileged have access. It isn't. In cases like the one above, all it really involves is writing in a more mathematically professional way – writing less like a student and more like a textbook, if you like.

This is not to belittle the genuine difficulties that undergraduate students face when handling proofs. Obviously the example above is simple one, and the proofs you are asked to understand and construct in upper-level courses will often (though not always) be much harder. It might take you a while to get used to digesting mathematics presented in this way, and to writing your own mathematics more professionally. But there is no reason to think you won't manage it, and this section discusses some things you could pay attention to in order to get used to it quickly.

One thing you'll often asked to do is to prove that a mathematical object satisfies a definition. The question won't be phrased like that, though. It will just say something like, "Prove that the set (2,5) is open." You will have to interpret this to mean "Prove that the set (2,5) satisfies the definition of open set," and review the definition of open set to make sure you are clear about what this involves. That sounds simple, but We have often seen students who, faced with an instruction like "Prove that the set (2,5) is open," don't know what to do. If you don't know how to start on any proof problem, your first thought should often be, what does the definition say?

We have seen the relevant definition, which says:

Definition: A set $X \subseteq \mathbb{R}$ is *open* if and only if $\forall x \in X, \exists d > 0$ such that $(x - d, x + d) \subseteq X$.

How should we write a proof that $(2,5)$ is an open set? Often the best advice is to follow the structure of the definition itself. We want to prove that $(2,5)$ is open, so we want to show that for every $x \in (2,5)$ there exists $d > 0$ such that $(x-d, x+d) \subseteq (2,5)$. When we want to show something is true for every x in some set, we usually start our proof by introducing one, like this:

Claim: (2,5) is an open set.
Proof: Let $x \in (2,5)$ be arbitrary.

Here, *arbitrary* means that we are taking any old x, not some specific one or one with special properties. It is not necessary to write this – lots of people would just write "Let $x \in (2,5)$" – but it does emphasize that the ensuing argument will work for any x in the set.

Now we need to show the existence of an appropriate d. The simplest way to show the existence of something is to produce one. In this case, we need to produce a d that will work for our x. This d will depend on x, and one way to do it is to pick d to be the minimum of the two distances $5-x$ and $x-2$ (think about why). So we might write the rest of our proof like this:

Claim: (2,5) is an open set.
Proof: Let $x \in (2,5)$ be arbitrary. Let d be the minimum of $5-x$ and $x-2$. Then $(x-d, x+d) \subseteq (2,5)$. So $\forall x \in (2,5), \exists d > 0$ such that $(x-d, x+d) \subseteq (2,5)$. So (2,5) is an open set.

Notice that this proof reflect the order and structure of the definition. We are showing something for all x, so we start with an arbitrary one. We are showing that for this x, an appropriate d exists, so we produce one. Because of this, the structure of the proof will be obvious to a mathematician, so you don't really need to write the line "So $\forall x \in (2,5)$…, but you might find it helpful for your own thinking.

Some people tend to include diagrams along with proofs like this, and some don't. That's because diagrams can be illuminating, but they are not necessary, and they are not a substitute for writing a proof out properly (diagrams are not proofs); mathematicians (in particular you algebra and calculus professors) want to see a written argument that is clearly linked to the appropriate definition.

Another standard proof type is known as *direct proof*. In a direct proof we start by assuming that the premise(s) hold and move, via a sequence of valid manipulations or logical deductions, to the desired conclusion. Here are some theorems for which we've already studied a direct proof:

Theorem: If n is an even number, then any integer multiple of n is even.

Theorem: If $x^2 - 20x + 10 = 0$ then $x = 10 + 3\sqrt{10}$ or $x = 10 - 3\sqrt{10}$.

Theorem: (2,5) is an open set.

Stop and think for a moment here. Can you write out proofs of these statements without looking? If you can't do it immediately, can you remember the gist of how we proceed in each case and reconstruct the rest? If you give yourself a minute for each one, we bet you can remember more than you would initially have thought. Students often have too little faith in their own ability to recall mathematical ideas and reconstitute arguments around them.

One important thing to note is that *direct* describes the eventual proof we write, it does not necessarily describe the process of constructing the proof. You might be able to write down the premises and just follow you nose to a proof, but is more likely that you'll have to try out some of the things like: write everything in terms of definitions, think about some examples, maybe draw a diagram, and so on. You should then, however, work out how to write your final proof in a way that makes its logical structure clear for a reader. It is probably a good idea to treat this writing as a separate task, one that is worthy of your attention over and above simply getting to an answer or a proof.

A second thing to note is that direct proofs can have somewhat more complicated structures within them. The most obvious such structure occurs in a *proof by cases*, which means what it sounds like it means: we divide up the cases we're dealing with into sensible groups and work with each one separately within the main proof. Consider the following theorem:

Theorem: For every $x, y \in \mathbb{R}, \max\{x, y\} = \dfrac{x + y + |x - y|}{2}$.

Of course, we need to define the symbols that appear in the statement.

Definition: For every $x, y \in \mathbb{R}$,

$$max\{x, y\} = \begin{cases} x, & \text{if } x \geq y, \\ y, & \text{if } y < x. \end{cases}$$

Definition: For every $x \in \mathbb{R}$,

$$|x| = \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0. \end{cases}$$

The first definition gives the meaning to $\max\{x, y\}$ as "choose the largest number between x and y", and the second definition is just the usual absolute value function. The theorem states that max can be expressed in terms

of the absolute value function. In order to prove the theorem, you need to reinterpret it as follows:

Theorem: For every $x, y \in \mathbb{R}$, the expression $\dfrac{x+y+|x-y|}{2}$ satisfies the definition of $\max\{x, y\}$.

Notice that the definition of $\max\{x, y\}$ is given in a piece-wise format, where two cases are obvious: $x \geq y$ and $y > x$. The absolute value function is defined in a piece-wise format as well (and probably it is the first time you see it written formally in this way). This should hint you into considering a proof by cases. Before reading the proof below, you should try to come up with one by yourself.

Proof: *Case 1:* Suppose that $x \geq y$. Then $x - y \geq 0$ and, thus, by the definition of the absolute value function, $|x-y| = x - y$. So,

$$\frac{x+y+|x-y|}{2} = \frac{x+y+x-y}{2} = \frac{2x}{2} = x.$$

Case 2: Suppose that $x < y$. Then $x - y < 0$ and, thus, by the definition of the absolute value function, $|x-y| = -(x-y) = y - x$. So,

$$\frac{x+y+|x-y|}{2} = \frac{x+y+y-x}{2} = \frac{2y}{2} = y.$$

Putting both cases together, we have that for all $x, y \in \mathbb{R}$,

$$\frac{x+y+|x-y|}{2} = \begin{cases} x, & \text{if } x \geq y, \\ y, & \text{if } y < x. \end{cases}$$

Therefore, $\dfrac{x+y+|x-y}{2}$ satisfies the definition of $\max\{x, y\}$.

So when should you consider a proof by cases? Sometimes you will find that you have to. In this illustration, for instance, we don't have much choice; the $\max\{x, y\}$ values are different depending on whether $x \geq y$ or $x < y$, so we have to handle theses cases separately. In other situations, it might not be necessary to use a proof by cases, but it might be convenient anyway because there is some sort of natural split, perhaps between positive and negative numbers (like in the definition of the absolute value function), or between odd and even ones. Finally, it might be worth starting a proof by cases if you think that you can construct an argument for some of the objects

to which the theorem applies but not for others. A start is better than nothing and, once you've got an argument written down for one case, you might find that reflecting on it gives you ideas about how to continue.

One final tip is that, if you have produced a proof by cases, or if you're looking at one produced by someone else, it might be a good idea to ask whether the number of cases could be reduced. You don't have to do this, of course – if your proof is valid as it is, that's fine. But remember that mathematicians also value elegance, and brevity contributes to elegance so it's a worthwhile aim.

The next type of proof we want to talk about is *proof by contradiction*. This is a type of indirect proof, so called because we do not proceed directly from the premises to the conclusion. Instead, we make a temporary assumption that our desired conclusion (or some part of it) is false, and we show that this leads us to a contradiction. From this we can deduce that the temporary assumption must have been wrong, and thus that the desired conclusion is true. This sounds rather convoluted, but you're accustomed to making this kind of argument informally in everyday life. Here's a simple case:

Your friend: Daniel was at home in Vicálvaro all weekend.

You: No he wasn't, I saw him in Alcorcón on Saturday afternoon.

Here you are implicitly using a proof by contradiction. The temporary assumption is that Daniel was in Vicálvaro. Your argument says that if we make that assumption, then we can deduce that he wasn't in Alcorcón (if you like, this uses the "theorem" that people can't be in two places at the same time). But this is contradicted by the fact that you saw him in Alcorcón. So the assumption that he was in Vicálvaro must have been wrong.

Next, we will look at a mathematical example, which involves the definition of *rational number*. The notation $\mathbb{Q}$ denotes the set of all rational numbers and $\mathbb{Z}$ denotes the set of integers. We introduce this definition here:

Definition: $x \in \mathbb{Q}$ if and only if $\exists p,q \in \mathbb{Z}$ (with $q \neq 0$) such that $x = p/q$.

The theorem and proof below use the symbol $\notin$ to mean "is not an element of," and in this case $y \notin \mathbb{Q}$ means that y is an *irrational number*. The theorem and proof both implicitly assumes that all the numbers we are working with are real (this is common in early work with rational and irrational numbers). As with any theorem and proof, you should read everything carefully, checking that you understand what is going on in each step.

Theorem: If $x \in \mathbb{Q}$ and $y \notin \mathbb{Q}$ then $x + y \notin \mathbb{Q}$.

Proof: Let $x \in \mathbb{Q}$, so $\exists p,\ q \in \mathbb{Z}$ (with $q \neq 0$) such that $x = p/q$ Let $y \notin \mathbb{Q}$. Suppose for contradiction that $x + y \in \mathbb{Q}$. This means that $\exists r,s \in \mathbb{Z}$ (with $s \neq 0$) such that $x + y = \frac{r}{s}$. But then

$$y = \frac{r}{s} - x = \frac{r}{s} - \frac{p}{q} = \frac{rq - ps}{sq}.$$

Now $rq - ps \in \mathbb{Z}$ and $sq \in \mathbb{Z}$ because $p,q,r,s \in \mathbb{Z}$. Also $sq \neq 0$ because $q \neq 0$ and $s \neq 0$. So $y \in \mathbb{Q}$. But this contradicts the theorem premise. So it must be the case that $x + y \notin \mathbb{Q}$.

In this proof, the temporary assumption is this one:

Suppose for contradiction that $x + y \in \mathbb{Q}$.

Making that temporary assumption leads us, by some sensible use of definitions and algebra, to this line:

So $y \in \mathbb{Q}$.

This (as stated) contradicts the theorem premise, so it allows us to conclude that our temporary assumption must have been wrong, like this:

So it must be the case that $x + y \notin \mathbb{Q}$.

It should be clear that in order to properly understand a proof like this, you have to do more than you might have had to do in earlier mathematics. In high school, most of your mathematical reading will have involved checking some algebra. Here, you have to be more sophisticated.

You certainly should check to make sure that you can see how the algebra works and that there are no mistakes. But that's not really where the action is in a proof like this. To fully understand it, you need to understand its global structure. You need to be able to identify what assumptions are made where, identify where the contradiction arises and what exactly is being contradicted, and understand how it all fits together to prove that the theorem is true.

The final standard proof type we want to discuss is *proof by induction*. Depending on your previous experience, you may have met this already. If so, you'll have used it to prove things like this:

$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$$

If not, you might not have seen this notation so here is a quick explanation. Thy symbol $\sum$ is called "sigma" and is a Greek upper case letter S, used here to denote a sum from $i=1$ to $i=n$. Written out, the left-hand side of the above means

$$\sum_{i=1}^{n} i^2 = 1^2 + 2^2 + 3^2 + \cdots + n^2.$$

What we've done here is substitute $i=1$, then $i=2$, then $i=3$, and so on, up to $i=n$, where we stop.

So our original expression actually captures infinitely many propositions:

$$P(1)\ 1^2 = \frac{1(1+1)(2+1)}{6}$$

$$P(2)\ 1^2 + 2^2 = \frac{2(2+1)(4+1)}{6}$$

$$P(3)\ 1^2 + 2^2 + 3^2 = \frac{3(3+1)(6+1)}{6}$$

$$P(4)\ 1^2 + 2^2 + 3^2 + 4^2 = \frac{4(4+1)(8+1)}{6}, \text{ and so on.}$$

Having a theorem that captures infinitely many cases isn't unusual—many of the other theorems we've looked at do the same. The difference here is that the form of the statement allows us to put the propositions in an ordered list $P(1), P(2), P(3), P(4), \ldots$

Proof by induction works as follows. First we prove $P(1)$. This is often easy. Then we do something clever. We don't try to prove any of the other propositions directly. Instead, we take a general number k and prove that if $P(k)$ is true, then $P(k+1)$ must be true too. This gives us $P(1) \Rightarrow P(2)$ and, since we've already proved $P(1)$, we can conclude that $P(2)$ is true as well. It also gives us $P(2) \Rightarrow P(3)$, so we can conclude that $P(3)$ is true as well. You get the idea. We have made an infinite chain of propositions, which are all true because the first one is true and the implications are all true:

$$P(1) \Rightarrow P(2) \Rightarrow P(3) \Rightarrow P(4) \Rightarrow \cdots$$

Proof by induction is one of those ideas that students usually find intuitive straightforward when it is explained in the abstract. However, they often find it difficult to use in any particular case, so we will look closely at the example we started with. In that example, it is easy to prof that $P(1)$ is true:

$$1^2 = 1 \text{ and } 1 = \frac{1(1+1)(2+1)}{6}$$

Proving that the implication $P(k) \Rightarrow P(k+1)$ is a bit harder. We would like to assume that $P(k)$ is true and use this to prove $P(k+1)$ is true. We would start doing some rough work at this point, writing something like this:

Will assume $P(k)$, which means $\sum_{i=1}^{k} i^2 = \frac{k(k+1)(2k+1)}{6}$.

Want to prove $P(k+1)$, which means

$$\sum_{i=1}^{k+1} i^2 = \frac{(k+1)\big([k+1]+1\big)\big(2[k+1]+1\big)}{6},$$

which, by rewriting the left-hand side, means we want

$$\left(\sum_{i=1}^{k} i^2\right) + (k+1)^2 = \frac{(k+1)\big([k+1]+1\big)\big(2[k+1]+1\big)}{6},$$

which, by the assumption about $P(k)$, means we want

$$\frac{k(k+1)(2k+1)}{6} + (k+1)^2 = \frac{(k+1)\big([k+1]+1\big)\big(2[k+1]+1\big)}{6}.$$

Then I've just got some algebra to do to show that the last two things are, in fact, equal (you might like to try it).

This, however, is definitely a situation in which the way you think about constructing a proof is not necessarily the same as the way you should write it out. The thinking above is completely logical, but presenting it like that wouldn't work very well because it doesn't match the structure of what we are trying to prove. When we write out a proof that $P(k) \Rightarrow P(k+1)$, we really want out proof to start with a clear assumption of $P(k)$ and proceed through some nice, tidy deduction to $P(k+1)$.

For proofs by induction, we favor a layout that makes that structure very clear. We would write something like this:

Theorem: $\forall n \in \mathbb{N}, \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.

Proof (by induction): Let $P(n)$ be the statement $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.

Note that $1^2 = 1 = \frac{1(1+1)(2+1)}{6}$ so $P(1)$ is true.

Now let $k \in \mathbb{N}$ be arbitrary and assume that $P(k)$ is true, that is, that

$$\sum_{i=1}^{k} i^2 = \frac{k(k+1)(2k+1)}{6}.$$

Then

$$\begin{aligned}
\sum_{i=1}^{k+1} i^2 &= \left(\sum_{i=1}^{k} i^2 \right) + (k+1)^2 \\
&= \frac{k(k+1)(2k+1)}{6} + (k+1)^2 \text{ by the assumption} \\
&= \frac{k(k+1)(2k+1) + 6(k+1)^2}{6} \\
&= \frac{(k+1)(2k^2 + 7k + 6}{6} \\
&= \frac{(k+1)(k+2)(2k+3)}{6} \\
&= \frac{(k+1)([k+1]+1)(2[k+1]+1)}{6}
\end{aligned}$$

So $\forall k \in \mathbb{N}, P(k) \Rightarrow P(k+1)$. Hence, by mathematical induction, $P(n)$ is true $\forall n \in \mathbb{N}$.

There are a couple of things to notice about this. First, the proof contains only a few words, but these help to make the structure clear. Second, all the algebra is in a single chain of equalities that starts with the left-hand side of the statement of $P(k+1)$ and ends with the right-hand side. You might like to think about why it makes sense to do the manipulations in this order, given that we know what we're going for. You might also like to think about why the last expression in the chain is not necessary but might be useful for a reader who wants to link the proof back to the theorem.

Confusion does tend to arise with proof by induction because there are lots of things to think about. We find that students are confused most often by the point which we write, "Assume that $P(k)$ is true." Students often read this and think, "But that's what we want to prove, how come we're allowed to assume it?" In fact, at that stage of the proof, we are not proving that $P(k)$ is true, we are proving that $P(k) \Rightarrow P(k+1)$. Make sure you can see the difference. Also, students sometimes confuse themselves, usually because they have allowed ambiguities to creep into their writing by using the word "it" in phrases like "so it is true for n." There are many possible candidates for the

meaning of "it" in a typical proof by induction, so you should be more specific. Writing "So $P(n)$ is true" is one way to do that. (Notice that the proof above does not include the word "it" – we are very specific about what we have deduced at each stage.)

So when should you use proof by induction? In some cases this will be obvious, because you are likely to have a section about it in at least one course. Also you will come across cases in which you want to prove that something is true for all $n \in \mathbb{N}$, which is a giveaway that induction is worth a try. Be aware, however, that problems for which induction is useful can vary quite a bit. First, there is no particular reason for a proof by induction to start at $n = 1$. You might be asked to prove that something is true for every $n \in \mathbb{N}$ such that $n > 4$, for instance. In that case, you can just make $P(5)$ your base case and proceed as before, except that at some point, perhaps in the *induction step* (the point in the proof when you apply the assumption that $P(k)$ is true, referred to as the *induction hypothesis*), you will find that you need $n > 4$ to justify some manipulation you want to make. Second, while some of the first problems you meet will involve working with a sum, proof by induction is useful for many other types of problem. All we really need is a situation in which we have infinitely many statements that can be listed in the order of the natural numbers, which can happen in all sorts of ways. For instance, consider these tasks:

Prove that for every natural number $n > 10, 2^n > n^3$

Prove that for every $n \in \mathbb{N}, 5^{3n} + 2^{n+1}$ is divisible by 3.

For the first task, we would write

$$\text{Let } P(n) \text{ be the statement that } 2^n > n^3.$$

Then we would prove directly that $P(11)$ is true, that is, that $2^{11} > 11^3$. Then we would work out how to prove that if $k > 10$ and $2^k > k^3$, then $2^{k+1} > (k+1)^3$.

For the second task, we would write

$$\text{Let } P(n) \text{ be the statement that } 5^{3n} + 2^{n+1} \text{ is divisible by 3.}$$

Then we would prove directly that $P(1)$ is true, that is, that $5^3 + 2^2$ is divisible by 3. Then we would work out how to prove that if $5^{3k} + 2^{k+1}$ is divisible by 3, then $5^{3(k+1)} + 2^{(k+1)+1}$ is divisible by 3. Notice, in this case, that the statement $P(k)$ is **not** just "$5^{3k} + 2^{k+1}$." In fact, $5^{3k} + 2^{k+1}$ isn't a statement at all - we couldn't prove it because it's just an expression (for any particular k it is a number; you can't "prove" a number, and it makes no sense to say that one number implies another). The statement is "$5^{3k} + 2^{k+1}$ is divisible by 3."

In our experience, starting out with a clear statement of $P(n)$ often makes the difference between success and failure in constructing a proof by induction, especially when dealing with a new type of problem. This isn't that surprising, of course – before you start any problem, you should always make sure that you are clear about what you are trying to do. In any case, you won't always be told what method to use, so you should be on the lookout for less familiar cases like these, and you should train yourself to notice when proof by induction might be useful.

Some advice

Earlier we said that a student should never sit in from of a problem and think "I don't know what to do." There are always things to try and, to be a good student, you must be willing to try them. In fact, you must be willing to try things that turn out not to work. In our experience, students are sometimes unwilling to do this, for three main reasons.

First, some students dislike the insecurity of not knowing exactly what to do. It makes them nervous. They want to know in advance what is going to work, and sometimes they ask for a teacher's assurance about this ("Is this the right way to do it"). The problem with seeking such assurance all the time is that you never find out what you could have done if you'd had a go, which means that you never get any more confident, which means that you end up in a vicious circle, having to ask for support all the time.

Second, some students do not want to waste time. We understand this – obviously no one wants to spend ages on one thing, especially when there are so many interesting things to do at university. But it is a big mistake to think that trying something that turns out not to work is a waste of time. Time spent learning is never wasted. If you try a method that doesn't work then, provided you are thoughtful about it, you learn *why* it doesn't work, which means you know something new about the applicability of the method. And you might gain some insight about the problem so that you have a better idea about what to try next. Of course, it is a mistake to keep plugging away at a method that clearly isn't working – research shows that good problem solvers stop frequently to re-evaluate whether their current approach seems to be getting them anywhere. But it's an even bigger mistake not to start.

Third, some students do not want to mess up their paper. They want to know that once they begin writing, they will be able to carry on writing and arrive at a nice, neat, correct solution. If this applies to you, then we're afraid you'll have to get over it. Real mathematical thinking is not tidy. It is full of false starts and partial attempts and realizations that what does not seem to be working just here would, in fact, form a useful part of a solution if put together with

something that failed then minutes ago, or yesterday, or last week. It is very important to embrace this if you want to keep improving as a mathematical problem solver. You need to get partial solution attempts on paper for the simple, practical reason that your brain cannot handle many things at once. You have an enormous amount of knowledge stored in what is known as you long-term memory, but your working memory, where you actually do the new thinking, has a seriously limited capacity. It will not be big enough to hold all the information about a complicated mathematical problem while simultaneously working out how to solve it. When you write down definitions or theorems or calculations that might help you solve a problem or construct a proof, you are using the paper to supplement your cognitive powers by, in effect, extending the capacity of your working memory. So don't be worried about writing things that are wrong or that turn out not to be useful. You can always write up a neat version of your solution or proof later.

Your professors will explain proofs that are long and logically complicated. They will also explain proofs that rely on some really clever insight. Sometimes they will explain proofs that are long and logically complicated *and* that rely on some really clever insight. This tends to worry students. They think, "Well, okay, I can see how that works, but *I would never have thought of it*." This makes them wonder whether they're good enough at mathematics. But you shouldn't worry, because you're not supposed to be able to reinvent the whole of modern mathematics by having all the original ideas yourself. Even a mathematics PhD student wouldn't be expected to have many totally original insights. As an undergraduate, when faced with a poof like this, your job is to appreciate the clever insight, to understand why it works, to think about how modifications of it might work in slightly different circumstances, and to relate it to ideas used elsewhere in the course or in your major. Te reassure you further, here is a list of things the you *will* be expected to do.

First, you will be expected to do routine mathematical calculations much like those you have seen in lower-level mathematics. As we said at the beginning of this introduction, you should be prepared for these calculations to be longer and more involved that those you have experienced before, and you should be prepared to have to adapt the calculation procedure if a step in it not valid for a new case.

Second, you will be expected to adapt proofs that you have seen to closely related cases. For instance:

- Having seen the proof that $(2,5)$ is an open set, you might be expected to prove that any interval of the form (a,b) is an open set.
- Having seen the proof that $\max\{x,y\} = \dfrac{x+y+|x-y}{2}$, you might be expected to find a formula, using the absolute value, to express the function

$$\phi(x) = \begin{cases} x, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

In such cases, you will often be able to treat the proof you have seen like a template, and change some numbers appropriately. However, to reiterate one of the main points of this introduction, you shouldn't do this thoughtlessly - you should make sure that each step in the proof really does work for the new cases, and be ready to make minor adjustments if it doesn't. It is important to be careful in cases where some number might be zero, for instance, or when dividing both sides of an inequality by a number that might be negative.

Third, you will be expected to adapt proofs you have seen to cases that are related, but not so closely. For instance:

- Having seen the proof that for all $n \in \mathbb{N}, \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$, you might be expected to prove that $a^n - b^n = (a-b)\sum_{i=1}^{n} a^{n-i}b^{i-1}$.
- Having seen the proof that $(2,5)$ is an open set, you might be expected to prove that $A = \{x \in \mathbb{R} : x^2 - 1 < 0\}$ is open.

In cases like these, the proof you have seen will certainly be helpful, but you will not be able to treat like a template. You might be able to construct a proof that is very similar in its basic structure, but you will have to think further to work out exactly what needs changing.

Fourth, you will be expected to show that definitions are satisfied. You will sometimes be expected to do this with definitions you have not seen before, if your professor thinks that they are sufficiently straightforward.

Fifth, you will be expected to construct proofs of theorems for which you haven't seen a closely related model, or to solve problems for which you haven't seen your professor solve in class. As we've said, the biggest mistake you could make here would be to sit around thinking "We haven't been shown how to do this." No one will ask you to do things that are completely beyond you.

Sixth and finally, on an exam you might be asked to solve some challenging problems. Sometimes a question might lead you through a solution in steps, or might offer a fairly big hit to help you with a key idea or useful trick. Sometimes it might just ask outright, which means you'll have to be able to remember the key ideas or useful tricks yourself and come up with the rest. To do so you will need to have effectively read and understood the material in your lecture notes.

1. SYSTEMS OF LINEAR EQUATIONS

The theory of linear equations plays an important and motivating role in the subject of this course. In fact, many problems in linear algebra are equivalent to studying a system of linear equations. Thus the techniques introduced in this chapter will be applicable to the more abstract treatment given later. On the other hand, some of the results of the abstract treatment will give us new insights into the structure of "concrete" systems of linear equations.

This chapter investigates systems of linear equations and describes in detail the Gaussian elimination algorithm which is used to find their solution. Matrices, together with certain operations on them, are also introduced here, since they are closely related to systems of linear equations and their solutions.

All our equations will involve specific numbers called *constants* or *scalars*. For simplicity, we will assume in this chapter that all our scalars are real numbers. The set of real numbers will be denoted, as usual, by $\mathbb{R}$. The solutions of our equations will also involve n-tupes

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

of real numbers called *column vectors*. The set of all such n-tuples is denoted by $\mathbb{R}^n$. We note that the results in this chapter also hold for equations over the complex numbers $\mathbb{C}$.

1.1 Linear equations and their solutions

By *linear equations* in the unknowns or variables $x_1, x_2, \ldots, x_n$ we mean an equation that can be written in the following standard form:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{1.1}$$

where $a_1, a_2, \ldots, a_n, b$ are scalars. The scalar a_k is called the *coefficient* of x_k and b is called the *independent term* of the equation.

A solution of the linear equation (1.1) is a list of scalars $(s_1, s_2, \ldots, s_n)$ with the property that the following statement (obtained by substituting each s_i for x_i in the equation) is true:

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = b$$

This set of values is then said to ***satisfy*** the equation.

The set of all such solutions is called ***solution set*** or ***general solution***, or, simply, the ***solution*** of the equation.

The above notions implicitly assume that there is an ordering of the variables. If the number of variables is small, then, in order to avoid subscripts, we will usually use the variables x, y, as ordered to denote two unknowns, x, y, z, as ordered, to denote three unknowns, x, y, z, t , as ordered, to denote four unknowns, and x, y, z, s, t , as ordered, to denote five unknowns.

Example 1.1

The equations

$$4x_1 - 5x_2 + 2 = x_1 \text{ and } x_2 = 2\left(\sqrt{6} - x_1\right) + x_3$$

are both linear because they can be rearranged algebraically as in (1.1):

$$3x_1 - 5x_2 = -2 \text{ and } 2x_1 + x_2 - x_3 = 2\sqrt{6}$$

The equations

$$4x_1 - 5x_2 = x_1x_2 \text{ and } x_3 = 2\sqrt{x_1} - 6$$

are not linear because of the presence of x_1x_2 in the first equation and $\sqrt{x_1}$ in the second.

Example 1.2

The equation $x + 2y - 4z + t = 3$ is linear in the four unknowns x, y, z, t. The list of scalars $(3,2,1,0)$ is a solution of the equation since

$$3 + 2(2) - 4(1) + 0 = 3$$

is a true statement. However the list $(1,2,4,5)$ is not a solution of the equation since

$$1 + 2(2) - 4(4) + 5 = 3$$

is not a true statement.

At this point, your algebra professor expects you to have carefully and thoroughly read the introduction section of these notes.

Definition 1.1

A linear equation is said to be *degenerate* if it has the form

$$0x_1 + 0x_2 + \cdots + 0x_n = b,$$

that is, if every coefficient is equal to zero.

The solutions of degenerate equations are described in the following theorem.

Theorem 1.1

Consider the degenerate linear equation $0x_1 + 0x_2 + \cdots + 0x_n = b$.

1. If the constant $b \neq 0$, then the equation has no solution.
2. If the constant $b = 0$, then every list of scalars $(s_1, s_2, \ldots, s_n)$ is a solution.

Proof. We begin by proving the first statement. Let $(s_1, s_2, \ldots, s_n)$ be any list of scalars.

Suppose that $b \neq 0$. Substituting $x_i = s_i$ in the equation we obtain:

$$0s_1 + 0s_2 + \cdots + 0s_n = b \text{ or } 0 + 0 + \cdots + 0 = b \text{ or } 0 = b.$$

This is not a true statement since $b \neq 0$. Hence no list of scalars is a solution.

Now we prove the second statement. Suppose $b = 0$. Substituting $(s_1, s_2, \ldots, s_n)$ in the equation we obtain:

$$0s_1 + 0s_2 + \cdots + 0s_n = 0 \text{ or } 0 + 0 + \cdots + 0 = 0 \text{ or } 0 = 0$$

which is a true statement. Thus every list of scalars is a solution, as claimed. ■

1.2 Systems of linear equations and elementary operations

We now consider a system of m linear equations, say, $L_1, L_2, \ldots, L_m$, in the n unknowns $x_1, x_2, \ldots, x_n$ which can be put in the *standard form*:

$$\begin{aligned} L_1 &: a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ L_2 &: a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ L_m &: a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{aligned} \tag{1.2}$$

where the a_{ij}, b_i are constants.

A *solution* (or *particular solution*) of the above system is a list of scalars $(s_1, s_2, \ldots, s_n)$ which is simultaneously a solution of all the equations of the system. The set of all such solutions is called the *solution set* or the *general solution* of the system.

Example 1.3

Consider the system of equations:

$$L_1 : x + 2y - 4z + t = 3,$$
$$L_2 : 2x + y + 5z - 4t = 13.$$

We can verify that $(3,2,1,0)$ is a solution by substituting $x = 3, y = 2, z = 1$, and $t = 0$ in the equations of the system. Similarly, we can verify that $(3,2,1,-3)$ is not a solution.

Systems of linear equations in the same unknowns are said to be *equivalent* if the systems have the same solution set. One way of producing a system which is equivalent to a given system, with linear equations $L_1, L_2, \ldots, L_m$, is by applying a sequence of the following three basic operations called elementary operations:

Definition 1.2: Elementary operations

(E1) Interchange the i th equation and the j th equation: $L_i \leftrightarrow L_j$.

(E2) Multiply the i th equation by a non zero scalar $k : kL_i \rightarrow L_i, k \neq 0$.

(E3) Replace the i th equation by itself plus k times the j th equation: $(kL_j + L_i) \rightarrow L_i$.

Remark 1.1

It is important that you understand that if one system of equations is obtained from another by applying elementary operations, then the two systems will be equivalent. This means that they will have the same solution set. The reason for this is that each elementary operation is *reversible*. The basic application of this fact is that we can *replace one system with an equivalent system that is easier to solve*.

Roughly speaking, the simplest equivalent version of a system is obtained using the elementary operations as follows. Use the x_1 term in the first equation of a system to eliminate the x_1 terms in the other equations. Then use the x_2 term in the second equation to eliminate the x_2 terms in the other equations, and so on, until you finally obtain the least amount of variables in

common among the equations. The following example illustrates this strategy. Make sure to understand that we replace x_1, x_2, x_3 with x, y, z.

Example 1.4

The solution of the system

$$\begin{aligned} L_1 &: x + 2y - 4z = -4 \\ L_2 &: 5x + 11y - 21y = -22 \\ L_3 &: 3x - 2y + 3z = 11 \end{aligned}$$

is obtained as follows:

Use x in the first equation to eliminate it from the other equations. First we replace L_2 with $-5L_1 + L_2$:

$$\begin{array}{lll} L_1: & x + 2y - 4z = -4 & \text{Operations:} \\ L_2: & 5x + 11y - 21z = -22 & \left(-5L_1 + L_2\right) \to L_2 \\ L_3: & 3x - 2y + 3z = 11 & \end{array} \qquad \begin{array}{lll} L_1: & x + 2y - 4z = -4 \\ \text{new } L_2: & y - z = -2 \\ L_3: & (3x - 2y + 3z = 11 \end{array}$$

Next we keep the new L_2 and replace L_3 with $-3L_1 + L_3$:

$$\begin{array}{lll} L_1: & x + 2y - 4z = -4 & \text{Operations:} \\ L_2: & y - z = -2 & \left(-3L_1 + L_3\right) \to L_3 \\ L_3: & 3x - 2y + 3z = 11 & \end{array} \qquad \begin{array}{lll} L_1: & x + 2y - 4z = -4 \\ L_2: & y - z = -2 \\ \text{new } L_3: & -8y + 15z = 23 \end{array}$$

Use y in the second equation to eliminate it from the other equations. We replace L_1 with $-2L_2 + L_1$ and replace L_3 with $8L_2 + L_3$. We show both replacements at once:

$$\begin{array}{lll} L_1: & x + 2y - 4z = -4 & \text{Operations:} \\ L_2: & y - z = -2 & \left(-2L_2 + L_1\right) \to L_1 \\ L_3: & -8y + 15z = 23 & \left(8L_2 + L_3\right) \to L_3 \end{array} \qquad \begin{array}{lll} \text{new } L_1: & x - 2z = 0 \\ L_2: & y - z = -2 \\ \text{new } L_3: & 7z = 7 \end{array}$$

At this point, it is evident that $z = 1$ and that you can use this to solve for the other two variables. Nevertheless, our philosophy of using elementary operations to reach a system of equations with the least amount of variables in common among them will be the keystone to developing all the theory in the rest of the course. So, even if it seems somewhat unnecessary right row, we will continue to solve the system with elementary operations for illustrative purposes.

To simplify calculations, we can replace L_3 with $\frac{1}{7}L_3$:

$$\begin{array}{lll} L_1: & x - 2z = 0 & \text{Operations:} \\ L_2: & y - z = -2 & \frac{1}{7}L_3 \to L_3 \\ L_3: & 7z = 7 & \end{array} \qquad \begin{array}{lll} L_1: & x - 2z = 0 \\ L_2: & y - z = -2 \\ \text{new } L_3: & z = 1 \end{array}$$

Use z in the third equation and use it eliminate from the other equations. We replace L_1 with $2L_3 + L_1$ and replace L_2 with $L_3 + L_2$:

$$\begin{array}{lll} L_1: & x - 2z = 0 & \text{Operations:} \qquad \text{new } L_1: \quad x = 2 \\ L_2: & y - z = -2 & \left(2L_3 + L_1\right) \to L_1 \quad \text{new } L_2: \quad y = -1 \\ L_3: & z = 1 & \left(L_3 + L_2\right) \to L_2 \qquad\quad L_3: \quad z = 1 \end{array}$$

We have clearly arrived at the simplest equivalent system of equations. Notice that the three equations have no variables in common. This is the ideal scenario since we can readily deduce that the solution of the system is $x = 2, y = -1, z = 1$.

Suppose that you start with the system of equation (1.2) and use elementary operations to obtain a simpler equivalent version such that the equations have the least amount of variables in common. Then the first variable with non zero coefficient in each equation is called ***leading variable*** and the variables that are not leading variables are called ***free variables***.

Recall from the beginning of this chapter that we have implicitly assumed that there is an ordering of the variables. Moreover, there may be more than one simpler system equivalent to (1.2). For this reason, we impose some additional requirements for writing a simpler equivalent version of (1.2). First, we will require that the equations should be rearranged so that the leading variables are order not only horizontally from left to right but also vertically from top to bottom. This can be accomplished by interchanging the necessary equations (recall that interchanging equations is an elementary operation). Second, we will require that the coefficient of each leading variable is 1*. This is possible by rescaling the necessary equations. These two additional conditions ensure that the resulting simpler equivalent system is ***the simplest*** one, independently of the sequence of elementary operations that you use to arrive at it.

This notion of simplest equivalent system will be, in our opinion, the most useful tool at your disposal to develop and understand the material in this course. We will illustrate how to obtain such simplest equivalent system in the following examples.

* This format of writing the simplest version is called the *row reduced echelon form*, but we are reserving this name for matrices.

Example 1.5

Consider the following system of equations:

$$\begin{aligned} L_1: &\quad y-4z=8, \\ L_2: &\quad 2x-3y+2z=1 \\ L_3: &\quad 5x-8y+7z=1 \end{aligned}$$

1. Find the simplest system of equations that is equivalent to the given one.
2. Indicate the leading variables and the free variables.
3. Find the solution of the given system of equation.

1. First, you need to remember that the variables need to be in the correct order horizontally and vertically, and that the coefficients of the leading variables should be equal to 1. This means that the variable x should appear as the leading variable of the first equation. This is not the case for the given system. One way to fix this is to interchange L_1 with either L_2 or L_3. Either option is good.

$$\begin{aligned} L_1: &\quad y-4z=8 \quad \text{Operations:} \quad \text{new } L_1: \quad 2x-3y+2z=1 \\ L_2: &\quad 2x-3y+2z=1 \quad L_1 \leftrightarrow L_2 \quad \text{new } L_2: \quad y-4z=8 \\ L_3: &\quad 5x-8y+7z=1 \qquad\qquad L_3: \quad 5x-8y+7z=1 \end{aligned}$$

The coefficient of x in L_1 should be 1 , so we rescale as follows:

$$\begin{aligned} L_1: &\quad 2x-3y+2z=1 \quad \text{Operations:} \quad \text{new } L_1: \quad x-\tfrac{3}{2}y+z=\tfrac{1}{2} \\ L_2: &\quad y-4z=8 \quad \tfrac{1}{2}L_1 \leftrightarrow L_1 \quad L_2: \quad y-4z=8 \\ L_3: &\quad 5x-8y+7z=1 \qquad\qquad L_3: \quad 5x-8y+7z=1 \end{aligned}$$

Use x in the first equation to eliminate it from the other equations.

$$\begin{aligned} L_1: &\quad x-\tfrac{3}{2}y+z=\tfrac{1}{2} \quad \text{Operations:} \quad L_1: \quad x-\tfrac{3}{2}y+z=\tfrac{1}{2} \\ L_2: &\quad y-4z=8 \quad \left(-5L_1+L_3\right) \to L_3 \quad L_2: \quad y-4z=8 \\ L_3: &\quad 5x-8y+7z=1 \qquad\qquad \text{new } L_3: \quad -\tfrac{1}{2}y+2z=-\tfrac{3}{2} \end{aligned}$$

Use y in the second equation and use it eliminate from the other equations.

$$\begin{aligned} L_1: &\quad x-\tfrac{3}{2}y+z=\tfrac{1}{2} \quad \text{Operations:} \quad \text{new } L_1: \quad x-5z=\tfrac{25}{2} \\ L_2: &\quad y-4z=8 \quad \left(\tfrac{3}{2}L_2+L_1\right) \to L_1 \quad L_2: \quad y-4z=8 \\ L_3: &\quad -\tfrac{1}{2}y+2z=-\tfrac{3}{2} \qquad\qquad \text{new } L_3: \quad 0=-\tfrac{5}{2} \end{aligned}$$

We have ran out of non-degenerate equations, there is the least amount of variables in common among the equation, the variables are in the correct order, and the coefficients of the leading variables are 1. Therefore we have found the simplest equivalent system.

2. The leading variables are x and y because they are the first variables in each nondegenerate equation with non zero coefficient. The remainder variable, z, is a free variable.

3. It turns out that $L_3 : 0 = -\frac{5}{2}$ is a degenerate equation with no solution which means that there is no values for x, y, z that will *simultaneously* satisfy all equations. Therefore, the given system of equations has no solution.

Before continuing on to the next example, we need to discuss free variables further. Suppose that after reducing (1.2) to its simplest equivalent version, you find that all of its degenerate equations (if any) are of the form $0x_1 + 0x_2 + \cdots + 0x_n = 0$, and that at least one of its variables is a free variable. Then the system has infinite number of solutions since each of the free variables may be assigned any real number. The general solution of the system is obtained as follows. Once the system has been reduced to its simplest form, arbitrary values, called ***parameters***, are assigned to the free variables, and the non free variables can be solved in terms of these parameters.

Example 1.6

Find the general solution of the system

$$\begin{aligned} L_1 &: \quad x + 4y - 3z + 2t = 5 \\ L_2 &: \quad -2x - 8y + 7z - 8t = -8 \end{aligned}$$

If the solution of the system is not obvious, then our best option to find it is to reduce the system to its simplest form. This is done as follows: First, we eliminate x from the second equation.

$$\begin{array}{lll} L_1 : & x + 4y - 3z + 2t = 5 & \text{Operations:} \\ L_2 : & -2x - 8y + 7z - 8t = -8 & \left(2L_1 + L_2\right) \to L_2 \end{array} \qquad \begin{array}{ll} L_1 : & x + 4y - 3z + 2t = 5 \\ L_2 : & z - 4t = 2 \end{array}$$

It is not possible to eliminate y from L_1 since the leading variable of L_2 is z, which means that our next task is to eliminate z from L_1.

$$\begin{array}{lll} L_1 : & x + 4y - 3z + 2t = 5 & \text{Operations:} \\ L_2 : & z - 4t = 2 & \left(3L_2 + L_1\right) \to L_1 \end{array} \qquad \begin{array}{ll} L_1 : & x + 4y - 10t = 11 \\ L_2 : & z - 4t = 2 \end{array}$$

The last system is the simplest equivalent version of the original one. Make sure to understand why this is so. The first variables with nonzero coefficients in each equation are x and z which means that these are the leading variables and that y and t are free variables. From the discussion above, we deduce that the system has infinite number solutions and

the general solution is obtained as follows. We assign arbitrary values to the free variables y and t, say,

$$y = \lambda \text{ and } t = \mu, \text{ where } \lambda, \mu \in \mathbb{R}$$

Then we solve for x and z in terms of λ and μ. We write the general solution:

$$\begin{aligned} x &= 11 - 4\lambda + 10\mu, \\ y &= \lambda, \\ z &= 2 + 4\mu, \\ t &= \mu. \end{aligned}$$

It is always a good idea to check that your final result is indeed the correct answer. In this case, you should verify that these expressions for x, y, z, t satisfy **both** equations of the original system. If you obtain two true statements then your answer is correct.

We finish this section by answering two fundamental questions about systems of linear equations:

1. Does at least one solution exist?

2. If a solution exists, is it the only one; that is, is the solution unique?

The answer to these questions are stated as a theorem.

Theorem 1.2

Any system of linear equations has either: (i) a unique solution, (ii) no solution, or (iii) an infinite number of solutions.

Proof. Given any system of linear equations, we apply elementary operations to reduce it to its simplest equivalent form. After this is done, we may get one of the following cases:

I. There is at least one degenerate equation of the form $0x_1 + 0x_2 + \cdots + 0x_n = b$ with $b \neq 0$. This degenerate equation has no solution and, consequently, the system has no solution.

II. There is no degenerate equation of the form $0x_1 + 0x_2 + \cdots + 0x_n = b$ with $b \neq 0$. We have two subcases.

   a. There is at least one free variable. This means that the system has an infinite number of solutions.

b. There are no free variables. This means that all the variables are leading variables. This is only possible if the simplest equivalent version is of the form

$$\begin{aligned} x_1 &= s_1 \\ x_2 &= s_2 \\ &\vdots \\ x_n &= s_n, \end{aligned}$$

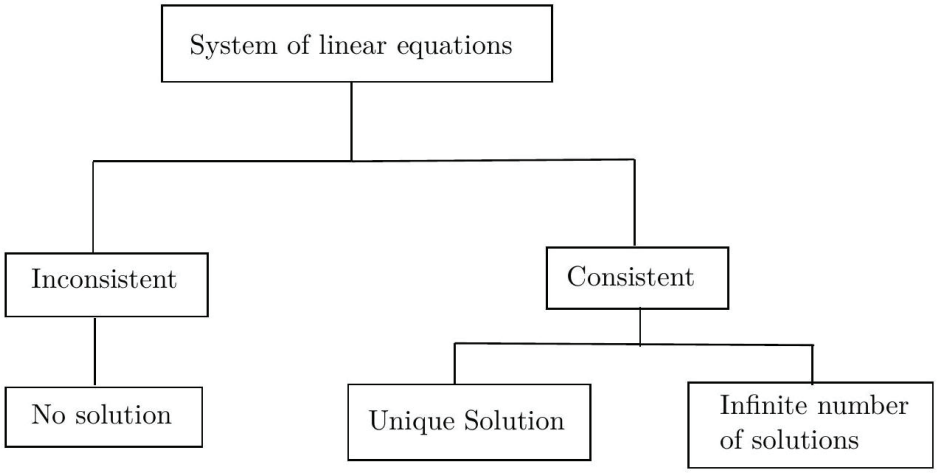
where $s_1, s_2, \ldots, s_n$ are real numbers. Then each variable has a unique value which means that the solution of the system is unique.

No other cases are possible, which proves the theorem. ■

Remark 1.2

Case II in the previous proof does not say that all the equations are non-degenerate. It only says that there are no degenerate equations that have no solution. In fact, this case may include degenerate equations of the form $0x_1 + 0x_2 + \cdots + 0x_n = 0$. In the future, make sure you understand what a definition, theorem or proof says and *does not* say.

In light of Theorem 1.2, systems of linear equations can be classified according to the existence of solutions. A system is said to be *consistent* if it has at least one solution, and is said to be *inconsistent* if it has no solution.



1.3 Matrices and matrix operations

A *matrix* is a rectangular array of numbers:

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

Such matrix may be denoted by $A = \left(a_{ij}\right)$. Note that the element a_{ij}, called the (i,j)-entry, appears in the i-th row and the j-th column. A matrix with m rows and n columns is called and $m \times n$ matrix.

The first non zero entry of a row is called its ***leading entry***. A row whose all of its entries consists of zeros has no leading entries.

Definition 1.3

A matrix is in *echelon form* (or *row echelon form*) if it has the following properties:

1. All zero rows, if any, are at the bottom of the matrix.
2. Each leading entry of a row is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zeros.

If a matrix in echelon form satisfies the following additional conditions, then it is in *reduced echelon form* (or *row reduced echelon form*):

4. The leading entry in each non zero row is 1 .
5. Each leading 1 is the only non zero entry in its column.

Example 1.7

The following matrices are in echelon form. The leading entries (represented with ■) may have any non zero value; the non leading entries (represented with *) may have any values (including zero).

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

The following matrices are in reduced echelon form because the leading entries are 1's, and there are 0 's below and above each leading 1, that is, each leading 1 is the only non zero entry in its column.

$$\begin{bmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 & * & 0 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 1 & 0 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & 0 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \end{bmatrix}$$

Just like with systems of equations, we can operate with the rows of a matrix by applying the following three basic operations.

Definition 1.4: Elementary row operations

(E1) Interchange the i th row and the j th row: $R_i \leftrightarrow R_j$.

(E2) Multiply the i th row by a non zero scalar $k : kR_i \rightarrow R_i, k \neq 0$.

(E3) Replace the i th row by itself plus k times the j th row: $\left(kR_j + R_i\right) \rightarrow R_i$.

A matrix A is said to be *row equivalent* to a matrix B, written $A \sim B$, if B can be obtained from A by applying a finite sequence of elementary row operations.

Any non zero matrix is row equivalent to more than one matrix in echelon form that can be obtained by using different sequences of row operations. However, the reduced echelon form one obtains from a matrix is unique, independently of the row operations used to obtain it. We state this as a theorem without proof.

Theorem 1.3

Each matrix is row equivalent to one and only one reduced echelon matrix.

When row operations on a matrix produce an echelon form, further row operations to obtain the reduced echelon form do not change the positions of the leading entries. Since the reduced echelon form is unique, *the leading entries are always in the same positions in any echelon form obtained from a given matrix*. These leading entries correspond to leading 1's in the reduced echelon form.

Definition 1.5

A *pivot position* in a matrix A is a location in A that corresponds to the leading 1 in the row reduced echelon form of A. A *pivot column* is a column of A that contains a pivot position.

Heuristically, pivot positions are non zero entries of a matrix that are used to create zeros as needed via elementary row operations to obtain a row reduced echelon form.

Example 1.8

Consider the matrix

$$A = \begin{bmatrix} 1 & 2 & -3 & 0 \\ 2 & 4 & -2 & 2 \\ 3 & 6 & -4 & 3 \end{bmatrix}.$$

1. Find *an* echelon form of A.
2. Find *the* row reduced echelon form of A.
3. Indicate the pivot positions of A.

1. There are more than one possible echelon form; each one is obtained depending on the sequence of elementary row operations used. We compute one of them. First, we start by using $a_{11} = 1$ as pivot to create 0 's below a_{11}, that is, apply $-2R_1 + R_2 \to R_2$ and $-3R_1 + R_3 \to R_3$ to obtain

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

Now use the new (2,3)-entry as a pivot to create 0 's below it, that is, apply $-\frac{5}{4}R_2 + R_3 \to R_3$ to obtain

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

The matrix is now in echelon form.

2. We apply further elementary row operations to the echelon form above to obtain the unique row reduced echelon form. Each pivot position should be 1,

so we apply $\frac{1}{4}R_2 \to R_2$ and $2R_3 \to R_3$ to obtain

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

At this point, the fastest way to obtain the row reduced echelon form is to apply $\frac{1}{2}R_3 + R_2 \to R_2$ to obtain

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

followed by applying $3R_2 + R_1 \to R_1$ to obtain

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

The matrix is now in row reduced echelon form.

3. The locations in A that correspond to the leading 1's in the row reduced echelon form are $a_{11} = 1, a_{23} = -2$, and $a_{34} = 3$. These are the pivot positions of A.

Matrices will usually be denoted by capital letters $A, B, \ldots$ Two matrices A and B are *equal*, written $A = B$, if they have the same size, that is, the same number of rows and columns, and if the corresponding elements are equal. Thus the equality of two $m \times n$ matrices is equivalent to a system of mn equalities, one for each pair of elements. For instance, the statement

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

is equivalent to the following system of equations:

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

The solution of the system is $x = 2, y = 1, z = 3, w = -1$. (Do not just accept that these are the answers. Try to solve the system by yourself.)

A matrix with one row is also referred to as a *row matrix*, and with one column as a *column vector*. In particular, a scalar can be viewed as a 1×1 matrix.

Let $A=\left(a_{ij}\right)$ and $B=\left(b_{ij}\right)$ be two matrices with the same size, say, $m\times n$ matrices. The *sum* of A and B, written $A+B$, is obtained by adding the corresponding entries:

$$A+B=\begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}+b_{m1} & a_{2m}+b_{2m} & \cdots & a_{mn}+b_{mn} \end{bmatrix},$$

or, more compactly, $A+B=\left(a_{ij}+b_{ij}\right)$.

The *product* of a matrix $A=\left(a_{ij}\right)$ by a scalar k, written kA, is the matrix obtained by multiplying each entry of A by k : $kA=\left(ka_{ij}\right)$. We also define $-A=\left(-1\right)A$ and $A-B=A+\left(-B\right)$.

The $m\times n$ matrix whose entries are all zeros is called the zero matrix. We usually write the zero matrix as 0 and its size is usually understood from context. The zero matrix is similar to the scalar 0. For any matrix $A, A+0=0+A=A.$

Basic properties of matrices under the operations of matrix addition and scalar multiplication follow.

Theorem 1.4

For any $m\times n$ matrices A,B,C and scalars, α, β

a) $\left(A+B\right)+C=A+\left(B+C\right)$ b) $A+0=A$

c) $A+\left(-A\right)=0$ d) $A+B=B+A$

e) $\alpha\left(A+B\right)=\alpha A+\alpha B$ f) $\left(\alpha+\beta\right)A=\alpha A+\beta A$

g) $\left(\alpha\beta\right)A=\alpha\left(\beta A\right)$ h) $1A=A$

Proof. We only prove the first property here. The rest are left as an exercise for the reader. Let $A=\left(a_{ij}\right), B=\left(b_{ij}\right)$, and $C=\left(c_{ij}\right)$ be matrices of the same size. Then

$$\left(A+B\right)+C=\left(\left(a_{ij}+b_{ij}\right)+c_{ij}\right)=\left(a_{ij}+\left(b_{ij}+c_{ij}\right)\right)=A+\left(B+C\right).$$

■

There are several ways to define matrix-vector multiplication. You have probably seen one way in high school. Nevertheless, we choose the following

definition because it is the most convenient for developing the theoretical ideas throughout the course. It should not be difficult for you verify that the definition you learned in high school is equivalent to the one we present here.

Definition 1.6

If A is an $m \times n$ matrix, with columns $\mathbf{a}_1,\ldots,\mathbf{a}_n$, and if $\mathbf{x}$ is a column vector in $\mathbb{R}^n$, then the *product of* A *and* $\mathbf{x}$, denoted by $A\mathbf{x}$, is defined as follows,

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 \, \mathbf{a}_2 \ldots \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

In words, the product of a matrix A and a column vector $\mathbf{x}$ *is the column vector that is equal to the weighted sum of the columns of* A *using the entries* $\mathbf{x}$ *as weights*. Note that $A\mathbf{x}$ is defined only if the number of columns of A equals the number of entries in $\mathbf{x}$.

We take this opportunity to give this type of weighted sum a special name. Given vectors $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_p$ in $\mathbb{R}^n$ and given scalars $c_1,c_2,\ldots,c_p$, the vector $y \in \mathbb{R}^n$ defined by

$$\mathbf{y} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p$$

is called a *linear combination* of $\mathbf{v}_1,\mathbf{v}_2,\ldots,\mathbf{v}_p$ with *weights* $c_1,c_2,\ldots,c_p$.

Example 1.9

We compute

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

By the definition of matrix-vector multiplication, we must compute the linear combination of the columns of the matrix using the entries of the vector as weights. Therefore,

$$\begin{bmatrix} 3 & 1 & 0 \\ 2 & 2 & 5 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = 2\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 0 \\ 5 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 21 \\ 14 \end{bmatrix}.$$

Remark 1.3

A common error that we see repeatedly is students treating vectors and matrices like fractions. For instance, some students will notice that the entries of the vector in the final result of the previous example share a common factor of 7 ; then they will proceed to "cancel" this common factor out like they would do with the numerator and denominator of a fraction and write the following as a final answer:

$$\begin{bmatrix} 3 & 1 & 0 \\ 2 & 2 & 5 \\ 1 & 0 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}.$$

This is, of course, absolutely wrong. If you write this, your algebra professor will realize that life is too short, immediately resign from teaching, and join a traveling circus. But not before changing your grade in the course to a zero.

Example 1.10

The product of a row matrix and a column vector is

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Make sure you understand how the definition was used to find this product.

The properties of the matrix-vector multiplication in the next theorem are important and will be used throughout the course.

Theorem 1.5

For an $m \times n$ matrix A, vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, and scalar c,

a. $A(\mathbf{u}+\mathbf{v}) = A\mathbf{u} + A\mathbf{v}$

b. $A(c\mathbf{u}) = c(A\mathbf{u})$

Proof. Let u_i and v_i be the i-th entries of $\mathbf{u}$ and $\mathbf{v}$, respectively. To prove the first statement, use the definition of matrix-vector multiplication to compute $A(\mathbf{u}+\mathbf{v})$:

$$
\begin{aligned}
A(\mathbf{u}+\mathbf{v}) &= \begin{bmatrix}\mathbf{a}_1\mathbf{a}_2\ldots\mathbf{a}_n\end{bmatrix}\begin{bmatrix}u_1+v_1\\u_2+v_2\\\vdots\\u_n+v_n\end{bmatrix}\\
&= (u_1+v_1)\mathbf{a}_1+(u_2+v_2)\mathbf{a}_2+\cdots+(u_n+v_n)\mathbf{a}_n\\
&= (u_1\mathbf{a}_1+u_2\mathbf{a}_2+\cdots+u_n\mathbf{a}_n)+(v_1\mathbf{a}_1+v_2\mathbf{a}_2+\cdots+v_n\mathbf{a}_n)\\
&= A\mathbf{u}+A\mathbf{v}.
\end{aligned}
$$

The proof of the second statement is left as an exercise for the reader. ■

You have certainly seen how to multiply matrices in high school. Nevertheless, such a definition is not very useful for theoretical purposes. We now give an equivalent abstract definition of matrix multiplication.

Definition 1.7

Let A be an $m\times p$ matrix, and let B be a $p\times n$ matrix. The product of A and B, denoted by AB, is the unique $m\times n$ matrix such that, for every vector $\mathbf{x}\in\mathbb{R}^n, (AB)\mathbf{x}=A(B\mathbf{x})$.

It is possible to find a (less abstract) representation of AB in terms of matrix-vector multiplication. We provide this representation in the following theorem.

Theorem 1.6

Let A be an $m\times p$ matrix, and let B be a $p\times n$ matrix with columns $\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n$, then the product AB is the unique $m\times n$ matrix whose columns are $A\mathbf{b}_1,A\mathbf{b}_2,\ldots,A\mathbf{b}_n$. That is,

$$AB=\begin{bmatrix}A\mathbf{b}_1 & A\mathbf{b}_2 & \ldots & A\mathbf{b}_n\end{bmatrix}.$$

Proof. Following the definition of AB, let $\mathbf{x}$ be an arbitrary vector in $\mathbb{R}^n$. Let

$$\mathbf{y}=B\mathbf{x}=\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_n\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}=x_1\mathbf{b}_1+x_2\mathbf{b}_2+\cdots+x_n\mathbf{b}_n.$$

Then,

$$
\begin{aligned}
AB\mathbf{x} &= A(B\mathbf{x})=A\mathbf{y}\\
&= A(x_1\mathbf{b}_1+x_2\mathbf{b}_2+\cdots+x_n\mathbf{b}_n)\\
&= x_1A\mathbf{b}_1+x_2A\mathbf{b}_2+\cdots+x_nA\mathbf{b}_n\\
&= \begin{bmatrix}A\mathbf{b}_1 & A\mathbf{b}_2 & \ldots & A\mathbf{b}_n\end{bmatrix}\mathbf{x}.
\end{aligned}
$$

This is true for every choice of $\mathbf{x}$ and each product $A\mathbf{b}_i$ is uniquely determined. This proves the theorem. ■

Example 1.11

Compute AB, where

$$A=\begin{bmatrix}2 & 3\\ 1 & -5\end{bmatrix} \text{ and } B=\begin{bmatrix}4 & 3 & 6\\ 1 & -2 & 3\end{bmatrix}.$$

Write $B=\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3\end{bmatrix}$, and compute

$$A\mathbf{b}_1=\begin{bmatrix}2 & 3\\ 1 & -5\end{bmatrix}\begin{bmatrix}4\\ 1\end{bmatrix}=\begin{bmatrix}11\\ -1\end{bmatrix}, \quad A\mathbf{b}_2=\begin{bmatrix}2 & 3\\ 1 & -5\end{bmatrix}\begin{bmatrix}3\\ -2\end{bmatrix}=\begin{bmatrix}0\\ 13\end{bmatrix}, \quad A\mathbf{b}_3=\begin{bmatrix}2 & 3\\ 1 & -5\end{bmatrix}\begin{bmatrix}6\\ 3\end{bmatrix}=\begin{bmatrix}21\\ -9\end{bmatrix}.$$

Then,

$$AB=\begin{bmatrix}A\mathbf{b}_1 & A\mathbf{b}_2 & A\mathbf{b}_3\end{bmatrix}=\begin{bmatrix}11 & 0 & 21\\ -1 & 13 & -9\end{bmatrix}.$$

The following theorem lists the standard properties of matrix multiplication. We note that I_m denotes the unique $m\times m$, called the *identity matrix*, such that $I_m\mathbf{x}=\mathbf{x}$ for all $\mathbf{x}$ in $\mathbb{R}^n$. The identity matrix has the following representation:

$$I_m=\begin{bmatrix}1 & 0 & 0 & \cdots & 0\\ 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots & 1\end{bmatrix}.$$

Theorem 1.7

Let A be an $m\times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

a. $A(BC)=(AB)C$ (associative law of multiplication)

b. $A(B+C)=AB+AC$ (left distributive law)

c. $(B+C)A=BA+CA$ (right distributive law)

d. $r(AB)=(rA)B=A(rB)$ for any scalar r

e. $I_mA=A=AI_m$ (identity for matrix multiplication)

Proof. Let $C = \begin{bmatrix} \mathbf{c}_1 & \cdots & \mathbf{c}_p \end{bmatrix}$. By Theorem 1.6,

$$BC = \begin{bmatrix} B\mathbf{c}_1 & \cdots & B\mathbf{c}_p \end{bmatrix}$$
$$A(BC) = \begin{bmatrix} A(B\mathbf{c}_1) & \cdots & A(B\mathbf{c}_p) \end{bmatrix},$$

By the definition of matrix multiplication, $A(B\mathbf{x}) = (AB)\mathbf{x}$ for all $\mathbf{x}$, so

$$A(BC) = \begin{bmatrix} (AB)\mathbf{c}_1 & \cdots & (AB)\mathbf{c}_p \end{bmatrix} = (AB)C.$$

The rest of the properties are left for the reader to prove. ■

The left-to-right order in matrix products is critical because AB and BA are usually not the same. On one hand, the sizes of A and B are such that AB is well defined but BA may not be even possible to compute. For instance, if A is a 3×5 matrix and B is a 5×1 matrix, then we can compute AB because the number of columns of A is equal to the number of rows of B, but it is not possible to compute BA because the number of columns of B is not equal to the number of rows of A. on the other hand, the columns of AB are linear combinations of the columns of A, whereas the columns of BA are constructed from the columns of B. If $AB = BA$, we say that A and B *commute* with one another.

Remark 1.4

We have already discussed that, in general, matrix products do not commute, that is, $AB \neq BA$. There two more important differences between matrix algebra and ordinary algebra that you need to keep in mind.

1. The cancellation laws *do not* hold for matrix multiplication. That is, if $AB = AC$, then it is not true, in general, that $B = C$. For instance, let

$$A = \begin{bmatrix} 2 & -3 \\ -4 & 6 \end{bmatrix}, B = \begin{bmatrix} 8 & 4 \\ 5 & 5 \end{bmatrix}, C = \begin{bmatrix} 5 & -2 \\ 3 & 1 \end{bmatrix}.$$

   Verify that $AB = AC$ and yet $B \neq C$.

2. If a product AB is the zero matrix, you *cannot* conclude, in general, that either $A = 0$ or $B = 0$. For instance, let

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

   Verify that $AB = 0$ and yet $A \neq 0$ and $B \neq 0$.

Every now and then we need to compute a particular entry of the product AB without computing the whole matrix product. In this case, we do so by

recalling the matrix multiplication rule you learned in high school (which can be deduced from Theorem 1.6).

Theorem 1.8

Let $A=\left(a_{ij}\right)$ be an $m\times p$ matrix and let $B=\left(b_{ij}\right)$ be a $p\times n$ matrix. If $(AB)_{ij}$ denotes the (i,j)-entry in AB, then

$$(AB)_{ij}=\begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{ip}\end{bmatrix}\begin{bmatrix} b_{1j}\\ b_{2j}\\ \vdots\\ b_{pj}\end{bmatrix}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{ip}b_{pj}.$$

Given an $m\times n$ matrix $A=\left(a_{ij}\right)$, the transpose of A, denoted by $A^{\top}$, is the $n\times m$ matrix $A^{\top}=\left(a_{ij}^{\top}\right)$, where $a_{ij}^{\top}=a_{ji}$. That is, $A^{\top}$ is obtained by changing the columns of A into columns, or, equivalently, by changing the rows of A into columns. So, if we write $A=\left[\mathbf{a}_1\dots\mathbf{a}_n\right]$, then

$$A^{\top}=\begin{bmatrix}\mathbf{a}_1^{\top}\\ \vdots\\ \mathbf{a}_n^{\top}\end{bmatrix}.$$

Theorem 1.9

Let A and B denote matrices whose sizes are appropriate for the following sums and products.

a. $\left(A^{\top}\right)^{\top}=A$

b. $(A+B)^{\top}=A^{\top}+B^{\top}$

c. For any scalar $r,(rA)^{\top}=rA^{\top}$

d. $(AB)^{\top}=B^{\top}A^{\top}$

Proof. We only prove the last statement. The rest are left for the reader as exercises. First, observe that

$$\left[(AB)^{\top}\right]_{ij}=(AB)_{ji}=a_{j1}b_{1i}+a_{j2}b_{2i}+\cdots+a_{jp}b_{pi}=\begin{bmatrix} b_{1i} & b_{2i} & \dots & b_{pi}\end{bmatrix}\begin{bmatrix} a_{j1}\\ a_{j2}\\ \vdots\\ a_{jp}\end{bmatrix}.$$

On the other hand,

$$\left[B^\top A^\top\right]_{ij} = \begin{bmatrix} b_{i1}^\top & b_{i2}^\top & \cdots & b_{ip}^\top \end{bmatrix} \begin{bmatrix} a_{1j}^\top \\ a_{2j}^\top \\ \vdots \\ a_{pj}^\top \end{bmatrix} = \begin{bmatrix} b_{1i} & b_{2i} & \cdots & b_{pi} \end{bmatrix} \begin{bmatrix} a_{j1} \\ a_{j2} \\ \vdots \\ a_{jp} \end{bmatrix},$$

which shows that the (i,j)-entry of $(AB)^\top$ is equal to the corresponding entry of $B^\top A^\top$. Therefore, $(AB)^\top = B^\top A^\top$. ■

1.4 Systems of linear equations and matrices

Consider again a system of m linear equations and n unknowns:

$$\begin{aligned} L_1 \quad & : a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ L_2 \quad & : a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ L_m \quad & : a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m. \end{aligned} \tag{1.3}$$

We can interpret (1.3) the equations resulting from considering two equal column vectors:

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Then, L_i is obtained by setting the i-th entry of the vector on the left side equal to the i-th entry of the vector on the right side. We can also operate on the left-hand side vector as follows:

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \begin{bmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix}$$

$$= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

So, if we let

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

we can write (1.3) more compactly as the vector equation

$$A\mathbf{x} = \mathbf{b}. \tag{1.4}$$

The matrix A is called the *matrix coefficient*, $\mathbf{x}$ is the *unknown vector*, and $\mathbf{b}$ is the *independent term*. We make the following key observation:

Observation 1.1

Every solution of the system (1.3) is solution of the vector equation (1.4), and vice versa.

The *augmented matrix* of (1.3) is the following matrix:

$$\begin{bmatrix} A & \mathbf{b} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{bmatrix}.$$

That is, the augmented matrix of the system (1.3) is the matrix which consists of the matrix A concatenated by the vector $\mathbf{b}$. Observe that the system (1.3) is completely determined by the vector equation (1.4) which, in turn, is completely determined by its augmented matrix. In light of this, we have the following theorem.

Theorem 1.10

If A is an $m \times n$ matrix, with columns $\mathbf{a}_1,\ldots,\mathbf{a}_n$, and if $\mathbf{b}$ is in $\mathbb{R}^m$, the matrix equation

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

has the same solution set as the vector equation

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$\begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n & \mathbf{b}\end{bmatrix}.$$

From now on, we will use the language and theory of matrices to study systems of linear equations. We reframe the question of existence of solutions in terms of matrices. Given a particular vector **b**, it is easy to deduce from Theorem 1.10 that the equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if **b** is a linear combination of the columns of A.

Example 1.12

Consider the following system of equations

$$\begin{aligned} x_1 + 2x_2 + 7x_3 &= b_1 \\ -2x_1 + 5x_2 + 4x_3 &= b_2, \end{aligned}$$

where b_1 and b_2 are arbitrary real numbers. Let

$$A = \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \end{bmatrix},\ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_2 \end{bmatrix},\ \text{and}\ \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Then the system can be written more compactly as the vector equation $A\mathbf{x} = \mathbf{b}$. Notice that **b** is a generic vector. Keeping this vector arbitrary will allow us to deduce general results about the system of equations since these results will not depend on a particular choice of **b**. Here, we will study the existence of solutions by solving the system with a generic vector **b**. To do so, we row reduce the augmented $\begin{bmatrix}A\mathbf{b}\end{bmatrix}$ as follows:

$$\begin{bmatrix} 1 & 2 & 7 & b_1 \\ -2 & 5 & 4 & b_2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 & b_1 \\ 0 & 9 & 18 & 2b_1 + b_2 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 7 & b_1 \\ 0 & 1 & 2 & \frac{2b_1+b_2}{9} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & \frac{5b_1-2b_2}{9} \\ 0 & 1 & 2 & \frac{2b_1+b_2}{9} \end{bmatrix}$$

The last matrix is the row reduced echelon form which, in turn, corresponds to an augmented matrix of a simpler equivalent system of equations. In fact, it corresponds to the simplest equivalent system. Indeed, if we write down the system that is determined by the row reduced echelon form:

$$x_1 + 3x_3 = \frac{5b_1 - 2b_2}{9}$$
$$x_2 + 2x_3 = \frac{2b_1 + b_2}{9},$$

it becomes clear that the equations have the least amount of variables in common, the variables are in the correct order, and the coefficients of the leading variables are 1's. The presence of the free variable x_3 indicates that this system has an infinite number of solutions, which are as follows:

$$x_1 = \frac{5b_1 - 2b_2}{9} - 3\lambda,$$
$$x_2 = \frac{2b_1 + b_2}{9} - 2\lambda,$$
$$x_3 = \lambda,\ \lambda \in \mathbb{R}.$$

Observe that it is possible to find a solution without putting any restriction on the vector **b**, which allows us to conclude that for each **b** in $\mathbb{R}^2$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution. Moreover, by the definition of the product $A\mathbf{x}$, we can write

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \left(\frac{5b_1 - 2b_2}{9} - 3\lambda \right) \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \left(\frac{2b_1 + b_2}{9} - 2\lambda \right) \begin{bmatrix} 2 \\ 5 \end{bmatrix} + \lambda \begin{bmatrix} 7 \\ 4 \end{bmatrix},\ \lambda \in \mathbb{R}.$$

The most useful interpretation of the equation above is that each **b** in $\mathbb{R}^2$ can be written as a linear combination of the columns of A. In this case, there is an infinite number of ways to write each **b** as a linear combination of the columns of A since there is an infinite number of options for the choice of values for the parameter λ.

You should always be on the look out for important observations when studying the material. In the previous example, observe that the pivot positions in the row reduced echelon form correspond to leading variables and that non pivot positions correspond to free variables. A moment's thought should convince you that this is true in general. This is such an important observation that we must highlight it.

Observation 1.2

The pivot columns of *coefficient matrix* of a system correspond to the leading variables of the system, and the non pivot columns correspond to free variables. Therefore, a system has an infinite number of solutions if its coefficient matrix has at least one non pivot column. Conversely, if every column of the coefficient matrix is a pivot column, then every variable is a leading variable and, thus, the system has a unique solution.

Warning: The last sentence of the observation above is about the *columns of the coefficient matrix* not about the augmented matrix nor about the rows of the coefficient matrix. On one hand, if every column of the augmented matrix is a pivot column, then it has a row of the form

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

which corresponds to a degenerate equation of the form $0 = 1$, which clearly has no solution. On the other hand, the uniqueness of the solution cannot be deduced if the coefficient matrix has a pivot in every row. The system in Example 1.12 has an augmented matrix with a pivot in every row and yet it has an infinite number of solutions.

Example 1.13

Let

$$A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

Is the equation $A\mathbf{x} = \mathbf{b}$ consistent for all possible $\mathbf{b}$?

Row reduce the augment matrix of $A\mathbf{x} = \mathbf{b}$:

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & 4b_1 + b_2 \\ 0 & 7 & 5 & 3b_1 + b_3 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & 4b_1 + b_2 \\ 0 & 0 & 0 & b_1 - \frac{1}{2}b_2 + b_3 \end{bmatrix}$$

The third row in the last matrix corresponds to a degenerate equation of the form

$$0 = b_1 - \frac{1}{2}b_2 + b_3.$$

If we proceed further applying elementary operations, this row will not change. Therefore, the system is consistent only if the entries of $\mathbf{b}$ satisfy the equation $0 = b_1 - \frac{1}{2}b_2 + b_3$. Clearly, not every $\mathbf{b}$ satisfies this condition (an easy example is $\mathbf{b} = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^\top$). If A had a pivot in all three rows, we would not care about the calculations in the augmented column because in this case an echelon form of the augmented matrix could not have a row such as $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$ corresponding to a degenerate equation with no solutions.

1.5 Solution sets of linear systems

The system of linear equations (1.3) is said to be *homogeneous* if all the independent terms are equal to zero, that is, if it can be written in the form $A\mathbf{x} = \mathbf{0}$, where A is its $m \times n$ coefficient matrix and $\mathbf{0}$ is the zero vector in $\mathbb{R}^m$. The homogeneous equation $A\mathbf{x} = \mathbf{0}$ always has a solution, namely $\mathbf{x} = \mathbf{0}$, the zero vector in $\mathbb{R}^n$, called the trivial solution. Any other solution, if it exists, is called a non trivial solution.

For a given equation $A\mathbf{x} = \mathbf{0}$, the important question is whether there is a non trivial solution. A non trivial solution exists if and only if the trivial solution is not a unique solution. The proof of Theorem 1.2 leads immediately to the following fact:

Remark 1.5

The homogeneous equation $A\mathbf{x} = \mathbf{0}$ has a non trivial solution if and only if the equation has at least one free variable.

Example 1.14

Determine if the following homogeneous system has a non trivial solution. Then describe the solution set.

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

Row reduce the augmented matrix to row reduced echelon form:

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

$$\sim \begin{bmatrix} 1 & 2 & -3 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & -5 & 4 & -3 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & 3 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & -1 & -8 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & 3 & 0 \\ 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 8 & 0 \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 & 0 & 0 & 11 & 0 \\ 0 & 1 & 0 & 7 & 0 \\ 0 & 0 & 1 & 8 & 0 \end{bmatrix}$$

The last matrix is the row reduced echelon form which, in turn, is the augmented matrix of the simplest equivalent system:

$$\begin{aligned} x + 11w &= 0, \\ y + 7w &= 0, \\ z + 8w &= 0. \end{aligned}$$

Notice that w is a free variable. It is possible to deduce this in advanced from the row reduced echelon form since the fourth column is **not** a pivot column (which means that w is not a leading variable; thus, w is free). By the previous remark, the system has a non trivial solution. After solving the simplified system, we deduce that the solution set consists of

$$\begin{aligned} x &= -11\lambda, \\ y &= -7\lambda, \\ z &= -8\lambda, \\ w &= \lambda \in \mathbb{R}, \end{aligned}$$

or, equivalently, the solution of the equation $A\mathbf{x} = 0$ where A is the coefficient matrix of the system, consists of all vector in $\mathbb{R}^4$ of the form

$$\mathbf{x} = \begin{bmatrix} -11\lambda \\ -7\lambda \\ -8\lambda \\ \lambda \end{bmatrix} = \lambda \begin{bmatrix} -11 \\ -7 \\ -8 \\ 1 \end{bmatrix}, \ \lambda \in \mathbb{R}.$$

That is, the solution of $A\mathbf{x} = 0$ consists of all the scalar multiples of the vector

$$\begin{bmatrix} -11 \\ -7 \\ -8 \\ 1 \end{bmatrix}.$$

Example 1.15

A single linear equation can be treated as a system with one equation. Describe the solution set of the homogeneous system

$$10x_1 - 3x_2 - 2x_3 = 0.$$

Obviously, there is no need to write its augmented matrix or to apply elementary row operations. We can directly solve for x_1 (respecting the implicit ordering of the variables). In other words, x_1 is the leading variable and

the rest of variables are free:

$$x_1 = \frac{3}{10}\lambda + \frac{1}{5}\mu,\ \lambda, \mu \in \mathbb{R},$$
$$x_2 = \lambda,$$
$$x_3 = \mu.$$

As a vector equation $A\mathbf{x} = 0$. (here $0 = 0 \in \mathbb{R}$), where

$$A = \begin{bmatrix} 10 & -3 & -2 \end{bmatrix}.$$

the general solution is

$$\mathbf{x} = \begin{bmatrix} \frac{3}{10}\lambda + \frac{1}{5}\mu \\ \lambda \\ \mu \end{bmatrix} = \lambda \begin{bmatrix} \frac{3}{10} \\ 1 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} \frac{1}{5} \\ 0 \\ 1 \end{bmatrix}.$$

In words, the general solution consists of all possible linear combinations of two vectors:

$$\begin{bmatrix} \frac{3}{10} \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{5} \\ 0 \\ 1 \end{bmatrix}.$$

When a non homogeneous system has more than one solution, two distinct solutions satisfy a simple but important relation between them, and the set of solutions of the homogeneous system characterizes this relation. We make this precise in the following theorem.

Theorem 1.11

Suppose that the equation $A\mathbf{x} = \mathbf{b}$ is consistent for some given non zero $\mathbf{b}$, and let $\mathbf{u}$ and $\mathbf{v}$ be two distinct solutions. Then there is some $\mathbf{h}$ satisfying $A\mathbf{h} = 0$, such that $\mathbf{u} = \mathbf{v} + \mathbf{h}$.

Proof. Let $\mathbf{u}$ and $\mathbf{v}$ be two distinct solutions of the equation $A\mathbf{x} = \mathbf{b}$. Then

$$A(\mathbf{u} - \mathbf{v}) = \mathbf{b} - \mathbf{b} = 0.$$

This means that $\mathbf{h} = \mathbf{u} - \mathbf{v}$ is a solution of the homeneous equation $A\mathbf{x} = 0$. This proves the theorem. ■

Another way to sate the previous theorem is as follows.

Theorem 1.12

Suppose that the equation $A\mathbf{x} = \mathbf{b}$ is consistent for some given non zero $\mathbf{b}$, and let $\mathbf{v}$ be a particular solution. Then the solution set of $A\mathbf{x} = \mathbf{b}$ consists of the set of all vectors of the form $\mathbf{u} = \mathbf{v} + \mathbf{h}$, where $\mathbf{h}$ is any solution of the homogeneous equation $A\mathbf{x} = 0$.

Warning: The previous theorem applies only when $A\mathbf{x} = \mathbf{b}$. If the equation has no solution, then the solution set is empty.

Example 1.16

Recall the system in Example 1.12:

$$\begin{aligned} x_1 + 2x_2 + 7x_3 &= b_1 \\ -2x_1 + 5x_2 + 4x_3 &= b_2, \end{aligned}$$

which is consistent for each $b_1, b_2 \in \mathbb{R}$. Its matrix coefficient is

$$A = \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \end{bmatrix},$$

and, in vector form, its general solution is

$$\mathbf{x} = \begin{bmatrix} \frac{5b_1 - 2b_2}{9} \\ \frac{2b_1 + b_2}{9} \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix}, \ \lambda \in \mathbb{R}$$

We now verify that this general solution is of the form $\mathbf{u} + \mathbf{h}$, where $\mathbf{u}$ satisfies $A\mathbf{u} = \mathbf{b}$ and $\mathbf{h}$ satisfies $A\mathbf{h} = 0$. Let

$$\mathbf{u} = \begin{bmatrix} \frac{5b_1 - 2b_2}{9} \\ \frac{2b_1 + b_2}{9} \\ 0 \end{bmatrix} \text{ and } \mathbf{h} = \lambda \begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix}.$$

Then

$$A\mathbf{u} = \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \end{bmatrix} \begin{bmatrix} \frac{5b_1 - 2b_2}{9} \\ \frac{2b_1 + b_2}{9} \\ 0 \end{bmatrix} = \left(\frac{5b_1 - 2b_2}{9} \right) \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \left(\frac{2b_1 + b_2}{9} \right) \begin{bmatrix} 2 \\ 5 \end{bmatrix} + 0 \begin{bmatrix} 7 \\ 4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

and

$$A\mathbf{h} = \begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \end{bmatrix} \begin{bmatrix} -3\lambda \\ -2\lambda \\ \lambda \end{bmatrix} = -3\lambda \begin{bmatrix} 1 \\ -2 \end{bmatrix} - 2\lambda \begin{bmatrix} 2 \\ 5 \end{bmatrix} + \lambda \begin{bmatrix} 7 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

1.6 Invertible matrices

An $n \times n$ matrix A is said to be *invertible* if there is an $n \times n$ matrix C such that

$$CA = I \text{ and } AC = I$$

where $I = I_n$ is the identity matrix. In this case, C is an *inverse* of A.

Theorem 1.13

The inverse of an $n \times n$ matrix A is unique.

Proof. Suppose that C and B are both inverses of A. Then,

$$B = BI = B(AC) = (BA)C = IC = C.$$

Therefore, $B = C$ which means that the inverse is uniquely determined by A. ■

Since the inverse of a matrix A is uniquely determined by A, we can assign it a special notation. This unique inverse is denoted by A^{-1}, so that

$$A^{-1}A = I \text{ and } AA^{-1} = I$$

A matrix that is not invertible is sometimes called a ***singular*** matrix, and an invertible matrix is called a ***non singular*** matrix.

Invertible matrices are indispensable in linear algebra – mainly for algebraic calculations and formula derivation. In particular, they play an important role in solving systems of equations.

Theorem 1.14

If A is an invertible $n \times n$ matrix, then for each $\mathbf{b} \in \mathbb{R}^r$, the equation $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

Proof. For any $\mathbf{b} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^n, \mathbf{x} = A^{-1}\mathbf{b}$ is solution since:

$$A\mathbf{x} = A\left(A^{-1}\mathbf{b}\right) = \left(AA^{-1}\right)\mathbf{b} = I\mathbf{b} = \mathbf{b}.$$

If $\mathbf{u}$ is another solution, then

$$A\left(\mathbf{u} - A^{-1}\mathbf{b}\right) = \mathbf{b} - \mathbf{b} = 0.$$

Therefore,

$$\mathbf{u} - A^{-1}\mathbf{b} = I\left(\mathbf{u} - A^{-1}\mathbf{b}\right) = A^{-1}A\left(\mathbf{u} - A^{-1}\mathbf{b}\right) = A^{-1}0 = 0.$$

This implies $\mathbf{u} = A^{-1}\mathbf{b}$, which means that the solution $\mathbf{x} = A^{-1}\mathbf{b}$ is unique. ■

The following theorem provides three useful properties of invertible matrices.

Theorem 1.15

a. If A is an invertible matrix, then A^{-1} is invertible and

$$\left(A^{-1}\right)^{-1} = A.$$

b. If A and B are $n \times n$ invertible matrices, then AB is invertible and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

c. If A is invertible, then $A^{\top}$ is invertible and

$$\left(A^{\top}\right)^{-1} = \left(A^{-1}\right)^{\top}.$$

Proof. a. We already know that A satisfies

$$AA^{-1} = I \text{ and } A^{-1}A = I.$$

Therefore, the inverse of A^{-1} is A, that is, $\left(A^{-1}\right)^{-1} = A$.

b. On one hand,

$$\left(B^{-1}A^{-1}\right)AB = B^{-1}\left(A^{-1}A\right)B = B^{-1}IB = B^{-1}B = I.$$

On the other hand,

$$AB\left(B^{-1}A^{-1}\right) = A\left(B^{-1}B\right)A^{-1} = AIA^{-1} = AA^{-1} = I.$$

Therefore, $(AB)^{-1} = B^{-1}A^{-1}$.

c. We compute

$$\left(A^{-1}\right)^{\top} A^{\top} = \left(AA^{-1}\right)^{\top} = I^{\top} = I.$$

Similarly,

$$A^{\top}\left(A^{-1}\right)^{\top} = \left(A^{-1}A\right)^{\top} = I^{\top} = I.$$

Therefore, $\left(A^{\top}\right)^{-1} = \left(A^{-1}\right)^{\top}$. ■

We have treated inverses abstractly. Now we describe a procedure for computing them.

Theorem 1.16

If A is an invertible matrix, then

$$\begin{bmatrix} A & I \end{bmatrix} \sim \begin{bmatrix} I & A^{-1} \end{bmatrix}.$$

Consequently, A is invertible if and only if it is row equivalent to the identity matrix I.

Proof. Write $I = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \end{bmatrix}$. Recall that A is invertible if and only if there is a unique matrix C such that

$$AC = I \text{ and } CA = I.$$

Write $C = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \end{bmatrix}$. Since

$$AC = \begin{bmatrix} A\mathbf{c}_1 & A\mathbf{c}_2 & \dots & A\mathbf{c}_n \end{bmatrix},$$

then $AC = I$ if and only if all the equations

$$A\mathbf{c}_i = \mathbf{e}_i \text{ for } i = 1,2,\dots,n.$$

have unique solutions. Each equation has an augmented matrix $\begin{bmatrix} A\mathbf{e}_i \end{bmatrix}$. We can apply the same sequence of row elementary operations to reduce all the augmented matrix to a row reduced echelon form. But instead of applying these same sequences of row operations n times, we could just apply them once to the *augmented* matrix

$$\begin{bmatrix} A & \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_n \end{bmatrix} = \begin{bmatrix} A & I \end{bmatrix}.$$

Recall that each $A\mathbf{c}_i = \mathbf{e}_i$ has a unique solution if and only if all the columns of A are pivot columns implying that $A \sim I$ and

$$\left[A \quad I\right] \sim \left[I \quad \mathbf{c}_1 \quad \mathbf{c}_2 \quad \dots \quad \mathbf{c}_n\right] = \left[A \quad C\right].$$

This proves the theorem because $C = A^{-1}$. ■

We use this procedure to recover the well-known formula for the inverse of a 2×2 matrix.

Example 1.17

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Suppose that $ad - bc \neq 0$. We compute

$$\left[A \quad I_2\right] = \begin{bmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} ac & bc & c & 0 \\ ac & ad & 0 & a \end{bmatrix} \sim \begin{bmatrix} ac & bc & c & 0 \\ 0 & ad-bc & -c & a \end{bmatrix}$$

$$\sim \begin{bmatrix} ac & bc & c & 0 \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} \sim \begin{bmatrix} ac & 0 & c + \frac{bc^2}{ad-bc} & \frac{-abc}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$$

$$\sim \begin{bmatrix} ac & 0 & \frac{acd}{ad-bc} & \frac{-abc}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & c + \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}.$$

Hence,

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

We must admit that we cheated. In the last reduction step, we divided by ac without knowing if $ac \neq 0$. Nevertheless, it is possible to deduce the same formula for A^{-1} when $ac = 0$. We leave it to the reader to treat this case. (Hint: consider two cases: $a \neq 0, c = 0$ and $a = 0, c \neq 0$. The case $a = c = 0$ is excluded because this implies $ad - bc = 0$.)

1.7 Determinants

Each $n \times n$ matrix A is assigned a special scalar called the determinant of A, denoted by $\det A$ or $|A|$. We give a recursive definition of a determinant.

Definition 1.8

The determinant of a 1×1 matrix $A=\begin{bmatrix} a \end{bmatrix}$ is defined as $|A|=a$. For $n\geq 2$, the determinant of an $n\times n$ matrix $A=\left(a_{ij}\right)$ is defined as

$$\det A=\sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det A_{1j},$$

where A_{ij} is the $\left(n-1\right)\times\left(n-1\right)$ matrix obtained from A by deleting the i-th row and the j-th column.

Example 1.18

Let us recover the well-known expression for the determinant of a 2×2 matrix. Let

$$A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then $A_{11}=d$ and $A_{12}=c$. Using the definition of $\det A$, we get

$$\det A=a\det A_{11}-b\det A_{12}=ad-bc.$$

Example 1.19

Let us show another well-known result. Let

$$A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},\ \mathbf{x}=\begin{bmatrix} x_1 \\ x_2 \end{bmatrix},\ \text{and}\ \mathbf{b}=\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

If $\det A=a_{11}a_{22}-a_{12}a_{21}\neq 0$, then it is possible to express the solution of $A\mathbf{x}=\mathbf{b}$ in terms of determinants. This result is known as Cramer's rule. Here we show how to obtain x_1. The remaining variable is left for the reader to compute.

Write $I_2=\begin{bmatrix} \mathbf{e}_1\mathbf{e}_2 \end{bmatrix}$. Notice that if $\mathbf{x}$ satisfies $A\mathbf{x}=\mathbf{b}$, we have

$$A\begin{bmatrix} \mathbf{x} & \mathbf{e}_2 \end{bmatrix}=\begin{bmatrix} A\mathbf{x} & A\mathbf{e}_2 \end{bmatrix}=\begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix},$$

were we have used the definition of matrix multiplication. On the other hand, we also have

$$A\begin{bmatrix} \mathbf{x} & \mathbf{e}_2 \end{bmatrix}=\begin{bmatrix} A\mathbf{x} & A\mathbf{e}_2 \end{bmatrix}=\begin{bmatrix} a_{11}x_1+a_{12}x_2 & a_{12} \\ a_{21}x_1+a_{22}x_2 & a_{22} \end{bmatrix}$$

Then,

$$\det\begin{bmatrix} a_{11}x_1 + a_{12}x_2 & a_{12} \\ a_{21}x_1 + a_{22}x_2 & a_{22} \end{bmatrix} = \det\begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix}.$$

We use the definition of determinant to compute the left hand side of the equation:

$$\begin{aligned} \det\begin{bmatrix} a_{11}x_1 + a_{12}x_2 & a_{12} \\ a_{21}x_1 + a_{22}x_2 & a_{22} \end{bmatrix} &= a_{22}\left(a_{11}x_1 + a_{12}x_2\right) - a_{12}\left(a_{21}x_1 + a_{22}x_2\right) \\ &= \left(a_{11}a_{22} - a_{12}a_{21}\right)x_1 + \left(a_{12}a_{22} - a_{12}a_{22}\right)x_2 \\ &= x_1 \det A. \end{aligned}$$

Therefore,

$$x_1 \det A = \det\begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix}$$

which gives us the expression for x_1:

$$x_1 = \frac{\det\begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix}}{\det A}.$$

Observe that the definition of determinant of a matrix A involves the first row of A. It turns out that it is possible to use any row or any column of A to compute its determinant. To state the next theorem, we introduce the following notation. The (i, j)-*cofactor* of A is the number C_{ij} given by

$$C_{ij} = (-1)^{i+j} \det A_{ij}.$$

We omit the proof of the following fundamental theorem to avoid a lenthy digression.

Theorem 1.17

The determinant of an $n \times n$ matrix $A = \left(a_{ij}\right)$ can be computed by a cofactor expansion across any row or down any column. The expansion across the i-th row is

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.$$

The cofactor expansion down the j-th column is

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}.$$

Example 1.20

Use a cofactor expansion to compute $\det A$, where

$$A = \begin{bmatrix} 1 & 0 & 5 \\ 2 & 4 & -1 \\ -2 & 0 & 2 \end{bmatrix}.$$

It is always desirable to minimize computational effort. If we choose to expand the determinant down the second column, then the cofactor expansion involves only one term because $a_{12} = a_{32} = 0$:

$$\det A = 4C_{22} = 4 \cdot (-1)^{2+2} \det \begin{bmatrix} 1 & 5 \\ -2 & 2 \end{bmatrix} = 4(12) = 48.$$

Let $A = (a_{ij})$ be an $m \times n$ matrix. We say that a_{ii} are elements in the *main diagonal*. A matrix is said to be *lower triangular* if $a_{ij} = 0$ when $i > j$, that is, when all the elements above the main diagonal are all zero. A matrix is said to be *upper triangular* if $a_{ij} = 0$ when $i < j$, that is, when all the elements below the main diagonal are all zero. Some examples of lower triangular matrices are

$$\begin{bmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 4 & 6 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 & 0 \\ 3 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Some examples of upper triangular matrices are

$$\begin{bmatrix} 1 & 3 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 3 & 5 \\ 0 & 1 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Matrices in echelon form are upper triangular matrices.

Theorem 1.18

If A is an $n \times n$ triangular matrix, then $\det A$ is the product of the entries on the main diagonal of A.

Proof. We present the proof only for upper triangular matrices. The proof for lower triangular matrix is similar.

We prove this by induction on the size of the matrix. We start by verifying the first non trivial case. Then, for a 2×2 upper triangular matrix, we have

$$\det\begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{bmatrix} = a_{11}a_{12} - 0\cdot a_{22} = a_{11}a_{12}\,.$$

Suppose that the theorem has been established for $(k-1)\times(k-1)$ upper triangular matrices. For a $k\times k$ upper triangular matrix, expand its determinant across its last row:

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1k} \\ 0 & a_{22} & a_{23} & \cdots & a_{2k} \\ 0 & 0 & a_{33} & \cdots & a_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{kk} \end{bmatrix} = a_{kk}C_{kk} = a_{kk}(-1)^{2k}\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1(k-1)} \\ 0 & a_{22} & a_{23} & \cdots & a_{2(k-1)} \\ 0 & 0 & a_{33} & \cdots & a_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{(k-1)(k-1)} \end{bmatrix}$$

The last matrix is a $(k-1)\times(k-1)$ upper triangular matrix. So by the induction hypothesis, we obtain

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1k} \\ 0 & a_{22} & a_{23} & \cdots & a_{2k} \\ 0 & 0 & a_{33} & \cdots & a_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{kk} \end{bmatrix} = a_{kk}a_{11}a_{22}\cdots a_{(k-1)(k-1)} = a_{11}a_{22}\cdots a_{kk}.$$

Then, by mathematical induction, the theorem is proved for all $n\times n$ upper triangular matrices. ■

One of the most important properties of determinants is they "behave well" under the applications of elementary row operations.

Theorem 1.19

Let A be a square matrix.

a. If the matrix B is obtained from A by interchanging two rows, then $\det A = -\det B$.

b. If the matrix B is obtained from A by multiplying one row by a scalar k, then $\det A = \frac{1}{k}\det B$.

c. If the matrix B is obtained from A by adding a multiple of a row to another row, then $\det A = \det B$.

A common strategy in hand calculations is to reduce A to echelon form and then to use the fact that the determinant of a triangular matrix is the product of the entries in its main diagonal.

Example 1.21

Compute $\det A$, where

$$A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}.$$

First, we interchange the first and the last row, which changes the sign of the determinant:

$$\det A = -\det \begin{bmatrix} 1 & -4 & 0 & 6 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 2 & -8 & 6 & 8 \end{bmatrix}.$$

Next, we replace rows in order to produce 0 ' under the pivot in the first column. These replacements do not change the determinant:

$$\det A = -\det \begin{bmatrix} 1 & -4 & 0 & 6 \\ 0 & 3 & 5 & -8 \\ 0 & -12 & 1 & 16 \\ 0 & 0 & 6 & -4 \end{bmatrix}.$$

We continue to replace rows to produce 0 's under $a_{22} = 3$. This does not change the determinant:

$$\det A = -\det \begin{bmatrix} 1 & -4 & 0 & 6 \\ 0 & 3 & 5 & -8 \\ 0 & 0 & 21 & -16 \\ 0 & 0 & 6 & -4 \end{bmatrix}.$$

If we rescale the last row by a factor of $\frac{1}{2}$, the determinant changes as follows:

$$\det A = -2\det \begin{bmatrix} 1 & -4 & 0 & 6 \\ 0 & 3 & 5 & -8 \\ 0 & 0 & 21 & -16 \\ 0 & 0 & 3 & -2 \end{bmatrix}.$$

The third row can be replaced by itself plus -7 times the last row, which does not change the determinant:

$$\det A = -2\det\begin{bmatrix} 1 & -4 & 0 & 6 \\ 0 & 3 & 5 & -8 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 3 & -2 \end{bmatrix}.$$

Finally, we obtain a triangular matrix if we interchange the last two rows, which changes the sign of the determinant:

$$\det A = 2\det\begin{bmatrix} 1 & -4 & 0 & 6 \\ 0 & 3 & 5 & -8 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & -2 \end{bmatrix}.$$

Therefore,

$$\det A = 2(1)(3)(3)(-2) = -36.$$

We are ready to state the main theorem of this section.

Theorem 1.20

A square matrix A is invertible if and only if $\det A \neq 0$.

Proof. Suppose that A has been reduced to an echelon form U by row replacements and row interchanges (no rescaling). Clearly, this is always possible. Each row interchange produces a change of sign; so if there are r row interchanges, by Theorem 1.19, we have

$$\det A = (-1)^r \det U.$$

Since U is in echelon form, it is an upper triangular matrix, and so $\det U$ is the product of its diagonal entries $u_{11}, u_{22}, \ldots, u_{nn}$.

If A is invertible, then all the entries u_{ii} are pivots because $A \sim I$, which means that $u_{ii} \neq 0$ for all $i = 1, 2, \ldots, n$. Therefore,

$$\det A = (-1)^r u_{11}u_{22}\ldots u_{nn} \neq 0.$$

If A is not invertible, the at least one entry, say u_{kk}, is not a pivot, which means that $u_{kk} = 0$. Consequently, $\det A = 0$.

We conclude this section by stating two more properties of the determinant.

Theorem 1.21

If A and B are $n \times n$ matrices, then

a. $\det A^{\top} = \det A$

b. $\det AB = \big(\det A\big)\big(\det B\big)$

2. VECTOR SPACES

This chapter introduces the underlying algebraic structure of linear algebra – that of a finite dimensional vector space. The definition of a vector space involves an arbitrary field whose elements are called *scalars*. The following notation will be used (unless otherwise stated or implied):

$\mathbb{K}$	the field of scalars
$a, b, c, \ldots$	the elements of $\mathbb{K}$
V	the given vector space
$\mathbf{u}, \mathbf{v}, \mathbf{w}$	the elements of V

Nothing essential is lost if the reader assumes that $\mathbb{K}$ is the real field $\mathbb{R}$ or the complex field $\mathbb{C}$.

2.1 Vector spaces and subspaces

The following defines the notion of *vector space* or *linear space*.

Definition 2.1

Let $\mathbb{K}$ be a given field ($\mathbb{R}$ or $\mathbb{C}$) and let V be a non empty set with rules of addition, and scalar multiplication which assigns to any $\mathbf{u}, \mathbf{v} \in V$ a $\text{sum} \mathbf{u} + \mathbf{v} \in V$ and to any $\mathbf{u} \in V$ and scalar $k \in \mathbb{K}$ a *product* $k\mathbf{u} \in V$. Then V is called a *vector space over* $\mathbb{K}$ (and the elements of V are called *vectors*) if the following axioms hold:

(A1) For any vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V, (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.

(A2) There is a vector in V, denoted by 0 and called the *zero vector*, for which $\mathbf{u} + 0 = \mathbf{u}$ for any vector $\mathbf{u} \in V$.

(A3) For each vector $\mathbf{u} \in V$ there is a vector in V, denoted by $-\mathbf{u}$, for which $\mathbf{u} + (-\mathbf{u}) = 0$.

(A4) For any vectors $\mathbf{u}, \mathbf{v} \in V, \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

(M1) For any scalar $k \in \mathbb{K}$ and any vectors $\mathbf{u}, \mathbf{v} \in V, k(\mathbf{u} + \mathbf{v}) = k\mathbf{u} + k\mathbf{v}$.

(M2) For any scalars $a, b \in \mathbb{K}$ and any vector $\mathbf{u} \in V, (a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$.

(M3) For any scalars $a, b \in \mathbb{K}$ and any vector $\mathbf{u} \in V, (ab)\mathbf{u} = a(b\mathbf{u})$.

(M4) For the unit scalar $1 \in \mathbb{K}, 1\mathbf{u} = \mathbf{u}$ for any vector $\mathbf{u} \in V$.

The above axioms naturally split into two sets. The first four are only concerned with the additive structure of V. From these four axioms it follows that any sum of vectors of the form

$$\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_m$$

requires no parentheses and does not depend upon the order of the summands.

Theorem 2.1

For any vector space V,

a. The zero vector $0 \in V$ is unique.

b. For any vector $\mathbf{u} \in V$, the *additive inverse* $-\mathbf{u} \in V$ is unique.

c. For any vectors $\mathbf{u},\mathbf{v},\mathbf{w} \in V$, if $\mathbf{u}+\mathbf{w} = \mathbf{v}+\mathbf{w}$ then $\mathbf{u} = \mathbf{v}$.

Proof. a. Suppose that $\hat{0} \in V$ is a vector for which $\mathbf{u}+\hat{0} = \mathbf{u}$ for any vector $\mathbf{u} \in V$. Then,

$$\begin{aligned} \hat{0} &= \hat{0} + 0 && (\text{by A2}) \\ &= 0 && (\text{by the definition of } \hat{0}). \end{aligned}$$

Since $\hat{0} = 0$, the zero vector is unique.

b. Given a vector $\mathbf{u} \in V$, let $\mathbf{w} \in V$ be a vector for which $\mathbf{u}+\mathbf{w} = 0$. Then,

$$\begin{aligned} \mathbf{w} &= \mathbf{w} + 0 && (\text{by A2}) \\ &= \mathbf{w} + \left[\mathbf{u} + (-\mathbf{u})\right] && (\text{by A3}) \\ &= \left[\mathbf{w} + \mathbf{u}\right] + (-\mathbf{u}) && (\text{by A1}) \\ &= 0 + (-\mathbf{u}) && (\text{by the definition of } w) \\ &= -\mathbf{u} && (\text{by A2}). \end{aligned}$$

Since $\mathbf{w} = -\mathbf{u}$, the negative of $\mathbf{u}$ is unique.

c. For any vectors $\mathbf{u},\mathbf{v},\mathbf{w} \in V$, such that $\mathbf{u}+\mathbf{w} = \mathbf{v}+\mathbf{w}$ then

$$\begin{aligned}
\mathbf{u} &= \mathbf{u} + 0 && (\text{by A2}) \\
&= \mathbf{u} + \left[\mathbf{w} + (-\mathbf{w})\right] && (\text{by A3}) \\
&= \left[\mathbf{u} + \mathbf{w}\right] + (-\mathbf{w}) && (\text{by A1}) \\
&= \left[\mathbf{v} + \mathbf{w}\right] + (-\mathbf{w}) && (\text{by } \mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}) \\
&= \mathbf{v} + \left[\mathbf{w} + (-\mathbf{w})\right] && (\text{by A1}) \\
&= \mathbf{v} + 0 && (\text{by A3}) \\
&= \mathbf{v} && (\text{by A2}).
\end{aligned}$$

Therefore, $\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}$ implies $\mathbf{u} = \mathbf{v}$. ■

Also, *subtraction* is defined by

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$$

On the other hand, the remaining four axioms are concerned with the "action" of the field $\mathbb{K}$ on V. Observe that the labelling of the axioms reflects this splitting. Using these additional axioms we prove the following simple properties of a vector space.

Theorem 2.2

Let V be a vector space over a field $\mathbb{K}$.

a. For any scalar $k \in \mathbb{K}$ and $0 \in V, k0 = 0$.

b. For $0 \in \mathbb{K}$ and any vector $\mathbf{u} \in V$, then $0\mathbf{u} = 0$.

c. For any $k \in \mathbb{K}$ and any $\mathbf{u} \in V, (-k)\mathbf{u} = k(-\mathbf{u}) = -(k\mathbf{u})$.

Proof. a. We compute:

$$k0 + 0 = k0 = k(0 + 0) = k0 + k0$$

But $k0 + 0 = k0 + k0$ implies $k0 = 0$ (Theorem 2.1 part c).

b. Similarly, for any vector $\mathbf{u} \in V$,

$$0\mathbf{u} + 0 = 0\mathbf{u} = (0 + 0)\mathbf{u} = 0\mathbf{u} + 0\mathbf{u}$$

which implies $0\mathbf{u} = 0$.

c. On one hand, for any $k \in \mathbb{K}$ and any $\mathbf{u} \in V$,

$$(-k)\mathbf{u} = (-k)\mathbf{u} + k\mathbf{u} - (k\mathbf{u}) = (-k + k)\mathbf{u} - (k\mathbf{u}) = 0\mathbf{u} - (k\mathbf{u}) = -(k\mathbf{u})$$

Therefore, $(-k)\mathbf{u} = -(k\mathbf{u})$. On the other hand,

$$k(-\mathbf{u}) = k(-\mathbf{u}) + k\mathbf{u} - (k\mathbf{u}) = k(\mathbf{u} - \mathbf{u}) - (k\mathbf{u}) = k0 - (k\mathbf{u}) = -(k\mathbf{u})$$

Therefore, $k(-\mathbf{u}) = -(k\mathbf{u})$. ■

Observe that when $k = 1$ in part c of the previous theorem, we have $-\mathbf{u} = (-1)\mathbf{u}$ for each $\mathbf{u} \in V$. That is, multiplying $\mathbf{u}$ by -1 produces its additive inverse.

Now we list a number of important examples of vector spaces which will be used throughout this course.

Example 2.1

Let $\mathbb{K}$ be an arbitrary field. The notation $\mathbb{K}^n$ is frequently used to denote the set of all column matrices with n rows and entries which are elements in $\mathbb{K}$. Here $\mathbb{K}^n$ is viewed as a vector space over $\mathbb{K}$ where vector addition and scalar multiplication defined by

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix}$$

and

$$k\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} k & a_1 \\ k & a_2 \\ & \vdots \\ k & a_n \end{bmatrix}.$$

The zero vector in $\mathbb{K}^n$ is

$$0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and the inverse of a vector is defined by

$$-\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} -a_1 \\ -a_2 \\ \vdots \\ -a_n \end{bmatrix}.$$

The proof that $\mathbb{K}^n$ is a vector space is left to the reader, so we now regard as stating that $\mathbb{R}^n$ with the usual operations is a vector space over $\mathbb{R}$.

Example 2.2

The notation $\mathbf{M}_{m,n}$, or simply $\mathbf{M}$, will be used to denote the set of all $m \times n$ matrices over an arbitrary field $\mathbb{K}$. Then $\mathbf{M}_{m,n}$ is a vector space over $\mathbb{K}$ with respect to the usual operations of matrix addition and scalar multiplication.

Example 2.3

Let $\mathbf{P}$ denote the set of all polynomials

$$a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n,\ n = 0,1,2,3,\ldots,$$

with coefficients a_i is some field $\mathbb{K}$. Then $\mathbf{P}$ is a vector space over $\mathbb{K}$ with respect to the usual operations of addition of polynomials and multiplication of polynomials by constants.

Example 2.4

Let X be any non empty set and let $\mathbb{K}$ be an arbitrary field $\mathbb{K}$. Consider the set $\mathbf{F}(X)$ of all functions from X into $\mathbb{K}$. The sum of two functions $f, g \in \mathbf{F}(X)$ is the function $f + g \in \mathbf{F}(X)$ defined by

$$(f+g)(x) = f(x) + g(x),\ \forall x \in X,$$

and the product of a scalar $k \in \mathbb{K}$ and a function $f \in \mathbf{F}(X)$ is the function $kf \in \mathbf{F}(X)$ defined by

$$(kf)(x) = kf(x),\ \forall x \in X.$$

Then $\mathbf{F}(X)$ with the above operations is a vector space over $\mathbb{K}$.

The zero vector in $\mathbf{F}(X)$ is the zero function 0 which maps each $x \in X$ into $0 \in \mathbb{K}$, that is,

$$0(x) = 0,\ \forall x \in X.$$

Also, for any function $f \in \mathbf{F}(X)$, the function $-f$ defined by

$$(-f)(x) = -f(x),\ \forall x \in X,$$

is the additive inverse of the function f.

Let W be a subset of a vector space V over a field $\mathbb{K}$. Then W is called a *subspace* of V if W is itself a vector space over $\mathbb{K}$ with respect to the operations of vector addition and scalar multiplication on V. Simple criteria for identifying subspaces follow.

Theorem 2.3

Suppose that W is a subset of a vector space V. Then W is a subspace of V if and only if the following hold:

a. $0 \in W$

b. W is *closed under vector addition*, that is:

For every $\mathbf{u}, \mathbf{v} \in W$, the $\text{sum} \mathbf{u} + \mathbf{v} \in W$.

c. W is *closed under scalar multiplication*, that is:

For every $\mathbf{u} \in W, k \in \mathbb{K}$, the product $k\mathbf{u} \in W$.

Conditions b and c may be combined into one condition.

Corollary 2.1

W is a subspace if and only if

a. $0 \in W$

b. W is *closed under linear combinations*, that is:

$$a\mathbf{u} + b\mathbf{v} \in W \text{ for every } \mathbf{u}, \mathbf{v} \in W \text{ and } a, b \in \mathbb{K}$$

Example 2.5

Let V be any vector space. Then the set $\{0\}$ consisting of the zero vector alone, and also the entire space V are subspaces of V.

Example 2.6

Let W be the set of vectors in $\mathbb{R}^3$ consisting of those vectors whose third component is 0; or, in other words

$$W = \left\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} : a, b, \in \mathbb{R} \right\}$$

Notice that $0 \in W$. Further, for any vectors

$$\begin{bmatrix} a \\ b \\ 0 \end{bmatrix}, \begin{bmatrix} c \\ d \\ 0 \end{bmatrix} \in W,$$

and scalars $r,s \in \mathbb{R}$, we have

$$r\begin{bmatrix} a \\ b \\ 0 \end{bmatrix} + s\begin{bmatrix} c \\ d \\ 0 \end{bmatrix} = \begin{bmatrix} ra+sc \\ rb+sd \\ 0 \end{bmatrix} \in W$$

Thus W is a subspace of $\mathbb{R}^3$.

Example 2.7

Let $V = \mathbf{M}_{n,n}$, the space of $n \times n$ matrices. Then the subset W_1 of (upper) triangular matrices and the subset W_2 of matrices satisfying $A^\top = A$, called the set of symmetric matrices, are subspaces of V since they are non empty and closed under linear combinations.

Example 2.8

Recall that $\mathbf{P}$ denotes the vector space of polynomials. Let $\mathbf{P}_n$ denote the subset of $\mathbf{P}$ that consists of all polynomials of degree $\leq n$, for a fixed n. Then $\mathbf{P}_n$ is a subspace of $\mathbf{P}$.

Recall that any solution of a homogeneous system $A\mathbf{x} = \mathbf{0}$. is an element of $\mathbb{R}^n$. Thus, the solution set of $A\mathbf{x} = \mathbf{0}$. is a subset of $\mathbb{R}^n$.

Theorem 2.4

The solution set W of a homogeneous system $A\mathbf{x} = \mathbf{0}$ in n unknowns is a subset of $\mathbb{R}^n$.

Proof. First, $\mathbf{0} \in W$ because $\mathbf{0}$ is always a solution of a homogeneous system: $A\mathbf{0} = \mathbf{0}$.

Next we show that W is closed under linear combinations. So let $\mathbf{u},\mathbf{v} \in W$ and $a,b \in \mathbb{R}$. Since $\mathbf{u}$ and $\mathbf{v}$ are solutions of the homogeneous system, we have $A\mathbf{u} = \mathbf{0}$ and $A\mathbf{v} = \mathbf{0}$. Moreover,

$$A\left(a\mathbf{u} + b\mathbf{v}\right) = aA\mathbf{u} + bA\mathbf{v} = a\mathbf{0} + b\mathbf{0} = \mathbf{0}$$

Consequently, $a\mathbf{u} + b\mathbf{v}$ is a solution of the homogeneous solution and, thus, $a\mathbf{u} + b\mathbf{v} \in W$. ■

2.2 Linear combinations and spans

Let V be a vector space over a field $\mathbb{K}$ and let $\mathbf{v}_1,\ldots,\mathbf{v}_m \in V$. Any vector in V of the form

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m$$

where $a_i \in \mathbb{K}$, is called a *linear combination* of $\mathbf{v}_1,\ldots,\mathbf{v}_m$. The set of all such linear combinations, denoted by

$$\text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$$

is called the *span* of $\mathbf{v}_1,\ldots,\mathbf{v}_m$.

Generally, for any subset S of V, Span $S = \{0\}$ when S is empty and Span S consists of all the linear combinations of vectors in S.

Theorem 2.5

If $\mathbf{v}_1,\ldots,\mathbf{v}_m \in V$ then $\text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$ is a subspace of V.

Proof. Since $0 \in V$ can be written as a linear combination of $\mathbf{v}_1,\ldots,\mathbf{v}_m$:

$$0 = 0\mathbf{v}_1 + \cdots + 0\mathbf{v}_n$$

then $0 \in \text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$. To show that $\text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$ is closed under linear combinations, choose any two vectors $\mathbf{u},\mathbf{v} \in \text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$. Therefore,

$$\begin{aligned}\mathbf{u} &= a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m,\\ \mathbf{v} &= b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_m\mathbf{v}_m\end{aligned}$$

for some scalars $a_i, b_i \in \mathbb{K}$. Then, for any $r,s \in \mathbb{K}$,

$$\begin{aligned}r\mathbf{u} + s\mathbf{v} &= r\left(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_m\mathbf{v}_m\right) + s\left(= b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + b_m\mathbf{v}_m\right)\\ &= \left(ra_1 + sb_1\right)\mathbf{v}_1 + \left(ra_2 + sb_2\right)\mathbf{v}_2 + \cdots + \left(ra_m + sb_m\right)\mathbf{v}_m\,.\end{aligned}$$

Then $r\mathbf{u} + s\mathbf{v}$ is a linear combination of the vectors $\mathbf{v}_1,\ldots,\mathbf{v}_m$, and, consequently, $r\mathbf{u} + s\mathbf{v} \in \text{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_m\}$. ■

On the other hand, given a vector space V, the vectors $\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_r$ are said to *span* or *generate* or to form a *generating set* of V if

$$V = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}.$$

In other words, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ span V if, for every $\mathbf{v} \in V$, there are scalars $a_1, a_2, \ldots, a_r$ such that

$$\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_r\mathbf{u}_r$$

that is, if $\mathbf{v}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$.

Example 2.9

Let

$$H = \left\{ \begin{bmatrix} a - 3b \\ b - a \\ a \\ b \end{bmatrix} : a, b \in \mathbb{R} \right\}.$$

Show that H is a subspace of $\mathbb{R}^4$.

In light of Theorem 2.5, if we can write H as the span of a set of vectors, then we can conclude that H is a subspace of $\mathbb{R}^4$. Hence, write a generic element of H as a linear combination of vectors:

$$\begin{bmatrix} a - 3b \\ b - a \\ a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \; a, b \in \mathbb{R}.$$

This means that a generic element of H can always be expressed as a linear combination of the two vectors on the right. Therefore,

$$H = \text{Span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

So H is a subspace by Theorem 2.5.

Example 2.10

For what value(s) of h will the polynomial $q(t)$ be in the subspace of $\mathbf{P}_2$ spanned by $p_1(t), p_2(t), p_3(t)$, if

$$p_1(t) = 1 - t - 2t^2,\ p_2(t) = 5 - 4t - 7t^2,\ p_3(t) = -3 + t,\ q(t) = -4 + 3t + ht^2.$$

The polynomial $q(t)$ is in the subspace spanned by $p_1(t), p_2(t), p_3(t)$ if and only if there are scalars $a_1, a_2, a_3 \in \mathbb{R}$ such that

$$q(t) = a_1 p_1(t) + a_2 p_2(t) + a_2 p_3(t).$$

The above equality holds if and only if the coefficients multiplying $1, t$, and t^2, respectively, on the left side of the equation are equal to the corresponding coefficients on the right side. Therefore, comparing the corresponding coefficients we get,

$$\begin{aligned} 1: \quad -4 &= a_1 + 5a_2 - 3a_3, \\ t: \quad 3 &= -a_1 - 4a_2 + a_3, \\ t^2: \quad h &= -2a_1 - 7a_2 + 0a_3, \end{aligned}$$

or, in vector form,

$$\begin{bmatrix} 1 & 5 & -3 \\ -1 & -4 & 1 \\ -2 & -7 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -4 \\ 3 \\ h \end{bmatrix}.$$

We study the solution set of this equation by reducing the augmented matrix to its row reduced echelon form:

$$\begin{bmatrix} 1 & 5 & -3 & -4 \\ -1 & -4 & 1 & 3 \\ -2 & -7 & 0 & h \end{bmatrix} \sim \begin{bmatrix} 1 & 5 & -3 & -4 \\ 0 & 1 & -2 & -1 \\ 0 & 3 & -6 & h-8 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 7 & 1 \\ 0 & 1 & -2 & -1 \\ 0 & 0 & 0 & h-5 \end{bmatrix}.$$

Observe that the third row in the last matrix corresponds to the degenerate equation $0 = h - 5$, which has a solution if and only if $h = 5$. Therefore, $q(t) \in \text{Span}\{p_1(t), p_2(t), p_3(t)\}$ if and only if $h = 5$.

2.3 Null space and column space of a matrix

There are two important subspaces associated with any $m \times n$ matrix A. In this section, we introduce both of these subspaces and study some of their properties.

Definition 2.2

The null space of an $m \times n$ matrix A, written as $\text{Nul}A$, is the set of all solutions of the homogeneous equation $A\mathbf{x} = 0$. In set notation,

$$\text{Nul}A = \left\{ \mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = 0 \right\}$$

A more dynamic description of $\text{Nul}A$ is the set of all $\mathbf{x} \in \mathbb{R}^n$ that are *annihilated* by A in the sense that each $\mathbf{x} \in \text{Nul}A$ is "transformed" into the zero vector.

The term *space* in *null space* is appropriate because the null space of an $m \times n$ matrix is a vector subspace of $\mathbb{R}^n$. This is a direct consequence of Theorem 2.4.

Example 2.11

Let H be the set of all vectors in $\mathbb{R}^4$ whose coordinates a,b,c,d satisfy the equations

$$\begin{aligned} a - 2b + 5c - d &= 0, \\ -a - b + c &= 0. \end{aligned}$$

In other words, each element $\mathbf{x}$ of H satisfies the homogeneous equation

$$\begin{bmatrix} 1 & -2 & 5 & -1 \\ -1 & -b & 1 & 0 \end{bmatrix} \mathbf{x} = 0.$$

So by the definition of the null space of a matrix, we have the following expression of H:

$$H = \text{Nul} \begin{bmatrix} 1 & -2 & 5 & -1 \\ -1 & -b & 1 & 0 \end{bmatrix}.$$

Example 2.12

Find a generating set for the null space of the matrix

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}.$$

The first step is to find the general solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$. So we reduce the augmented matrix to its row reduced echelon form:

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

Then, the solution set of $A\mathbf{x} = 0$ is the same as the solution set of the simpler system

$$\begin{aligned} x_1 - 2x_2 - x_4 + 3x_5 &= 0, \\ x_3 + 2x_4 - 2x_5 &= 0, \\ 0 &= 0, \end{aligned}$$

with free variables x_2, x_4, x_5. If we write the leading variables in terms of the free variables, then we can express the general solution of $A\mathbf{x} = 0$ as:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2a + b - 3c \\ a \\ -2b + 2c \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}.$$

This means that every solution of the homogeneous equation can be expressed as a linear combination of the three vectors on the right. Then, the generating set of $\text{Nul}A$ is

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

Using the notation of this and the previous section, we may write

$$\text{Nul}A = \text{Span} \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Observation 2.1

The thoughtful reader would observe that the number of non zero vectors in the generating set of Nul A equals the number of free variables in the system $A\mathbf{x} = 0$.

Another important subspace associated with a matrix is its column space. Unlike the null space, the column space is defined explicitly via linear combinations.

Definition 2.3

The *column space* of an $m \times n$ matrix A, written as $\text{Col}A$, is the set of all linear combinations of the columns of A. If $A = \begin{bmatrix} \mathbf{a}_1 \dots \mathbf{a}_n \end{bmatrix}$, then

$$\text{Col}A = \text{Span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}.$$

The fact that $\text{Col}A$ is a subspace of $\mathbb{R}^m$ follows from Theorem 2.5. Note that by the definition of matrix multiplication, a typical vector in $\text{Col}A$ can be written as $A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$. That is,

$$\text{Col}A = \left\{ \mathbf{b} \in \mathbb{R}^m : \mathbf{b} = A\mathbf{x} \text{ for some } \mathbf{x} \in \mathbb{R}^n \right\}.$$

Example 2.13

Find a matrix A such that $W = \text{Col}A$, where

$$W = \left\{ \begin{bmatrix} 6a - b \\ a + b \\ -7a \end{bmatrix} : a, b \in \mathbb{R} \right\}.$$

We first find a generating set of W and use the vectors in the generating set as columns of the matrix A.

$$W = \left\{ a \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix} + b \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} : a, b \in \mathbb{R} \right\} = \text{Span}\left\{ \begin{bmatrix} 6 \\ 1 \\ -7 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

The matrix A whose columns are the vectors in the generating set of W is

$$A = \begin{bmatrix} 6 & -1 \\ 1 & 1 \\ 0 & -7 \end{bmatrix}.$$

Then $W = \text{Col}A$.

The following example is meant to emphasize the differences between the null space and the column space of a matrix.

Example 2.14

Let $A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$.

a. If the column space of A is a subspace of $\mathbb{R}^k$, what is k?

b. If the null space of A is a subspace of $\mathbb{R}^k$, what is k?

c. Find a non zero vector of $\text{Col}A$ and a non zero vector of $\text{Nul}A$?

d. Is there a vector in $\text{Col}A$ that also belongs to $\text{Nul}A$?

The answers to these questions are:

a. Since the columns of A belong to $\mathbb{R}^3$, then they must generate a subspace of $\mathbb{R}^3$. So $k = 3$.

b. The matrix A has four columns, which means that the product $A\mathbf{x}$ is defined if and only if $\mathbf{x} \in \mathbb{R}^4$. Therefore, the general solution of the homogeneous equation $A\mathbf{x} = \mathbf{0}$ belongs to $\mathbb{R}^4$. Hence, $k = 4$.

c. Finding a non zero element of $\text{Col}A$ is easy. Clearly, any column of A belongs to $\text{Col}A$. Then it is sufficient to choose any column, say, the first one: $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix}$.

Finding a non zero element of $\text{Nul}A$ is harder. For this, we need to solve the homogeneous equation $A\mathbf{x} = \mathbf{0}$. Reduce the augmented matrix to row reduced echelon form:

$$\begin{bmatrix} 2 & 4 & -2 & 1 & 0 \\ -2 & -5 & 7 & 3 & 0 \\ 3 & 7 & -8 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 9 & 0 & 0 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

which corresponds to the system of equations

$$\begin{aligned} x_1 + 9x_3 &= 0, \\ x_2 - 5x_3 &= 0, \\ x_4 &= 0, \end{aligned}$$

with x_3 free. The general solution of the homogeneous equation is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -9\lambda \\ 5\lambda \\ \lambda \\ 0 \end{bmatrix} = \lambda \begin{bmatrix} -9 \\ 5 \\ 1 \\ 0 \end{bmatrix}.$$

Then, a no zero element of the null space of A is $\begin{bmatrix} -9 \\ 5 \\ 1 \\ 0 \end{bmatrix}$.

d. Each element of $\mathrm{Col}A$ is a vector in $\mathbb{R}^3$ and each element of $\mathrm{Nul}A$ is a vector in $\mathbb{R}^4$. Thus, with only three entries, the elements of $\mathrm{Col}A$ could not possible be in $\mathrm{Nul}A$.

Table 1: Contrast between $\mathrm{Nul}A$ an $\mathrm{Col}A$ of an $m \times n$ matrix A.

$\mathrm{Nul}A$	$\mathrm{Col}A$
1. Nul A is a subspace of $\mathbb{R}^n$.	1. $\mathrm{Col}A$ is a subspace of $\mathbb{R}^m$.
2. $\mathrm{Nul}A$ is defined implicitly via the condition ($A\mathbf{x} = 0$) that its elements must satisfy.	2. $\mathrm{Col}A$ is defined explicitly via how to construct its elements.
3. A typical vector $\mathbf{v}$ is $\mathrm{Nul}A$ has that property that $A\mathbf{v} = 0$.	3. A typical vector $\mathbf{v}$ is $\mathrm{Col}A$ has the property that $A\mathbf{x} = \mathbf{v}$ is consistent.
4. Given a specific vector $\mathbf{v}$, it is to tell if $\mathbf{v}$ is in $\mathrm{Nul}A$. Just compute $A\mathbf{v}$.	4. Given a specific vector $\mathbf{v}$, it takes time to tell if $\mathbf{v}$ is in $\mathrm{Col}A$. The solutions of $A\mathbf{x} = \mathbf{v}$ must be studied.
5. $\mathrm{Nul}A = \{0\}$ if and only if the equation $A\mathbf{x} = 0$ has a unique solution.	5. $\mathrm{Col}A = \mathbb{R}^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b} \in \mathbb{R}^m$.

2.4 Linear independence

We now define the notion of linear independence. This concept plays an essential role in the theory of linear algebra and in mathematics in general.

Definition 2.4

Let V be a vector space over a field $\mathbb{K}$. The vectors $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are said to be *linearly independent* if the vector equation

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m = 0 \tag{2.1}$$

has a unique solution: $c_1 = c_2 = \cdots = c_m = 0$. The vectors $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are said to be *linearly dependent* if (2.1) has a non trivial solution.

If 0 is one of the vectors in (2.1), say $\mathbf{u}_1$, then the vectors must be linearly dependent; for

$$1\mathbf{u}_1 + 0\mathbf{u}_2 + \cdots + 0\mathbf{u}_m = 10 + 0 + \cdots + 0 = 0,$$

which means that(2.1) has the non trivial solution $c_1 = 1, c_2 = \cdots = c_m = 0$.

Any non zero vector $\mathbf{u}$ is, by itself, linearly independent; for the equation $c\mathbf{u} = 0$ has the unique solution $c = 0$.

If two of the vectors in (2.1) are equal or one is a scalar multiple of the other, say $\mathbf{u}_1 = k\mathbf{u}_2$, then the vectors are linearly dependent. For, in this case,

$$\mathbf{u}_1 - k\mathbf{u}_1 + 0\mathbf{u}_3 + \cdots + 0\mathbf{u}_m = 0,$$

which means that (2.1) has the non trivial solution $c_1 = 1, c_2 = -k, c_3 = \cdots = c_m = 0$. In particular, two vectors are linearly dependent if and only if one is a multiple of the other.

Clearly, if the vectors $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are linearly independent, then any rearrangement of these vectors is also linearly independent.

If the vectors $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are linearly independent, then any subset of these vectors, say $\mathbf{u}_1,\ldots,\mathbf{u}_r (r < m)$, is linearly independent. Indeed, consider the equation

$$d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_r\mathbf{u}_r = 0.$$

But we can also write

$$d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_r\mathbf{u}_r + 0\mathbf{u}_{r+1} + \cdots + 0\mathbf{u}_m = 0.$$

Since $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$ are linearly independent, this equation has only the trivial solution, which means that $d_1 = \cdots = d_r = 0$ and, thus, $\mathbf{u}_1,\ldots,\mathbf{u}_r$ are linearly

independent. By contrapositive (see the Introduction), if some subset of $\mathbf{u}_1,\ldots,\mathbf{u}_m \in V$, say $\mathbf{u}_1,\ldots,\mathbf{u}_r (r < m)$, is linearly dependent, then $\mathbf{u}_1,\ldots,\mathbf{u}_m$ is linearly dependent.

The following theorem will often be useful. It states that a linearly dependent list of vectors, with the first vector not $\mathbf{0}$, one of the vectors is in the span of the previous ones.

Theorem 2.6

An indexed set $\{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ of two or more vectors, with $\mathbf{u}_1 \neq \mathbf{0}$, is linearly dependent if and only if some $\mathbf{u}_j (j > 1)$ is a linear combination of the preceding vectors $\mathbf{u}_1,\ldots,\mathbf{u}_{j-1}$.

Proof. Suppose that, for some $j > 1, \mathbf{u}_j$ is a linear combination of $\mathbf{u}_1,\ldots,\mathbf{u}_{j-1}$. Then there are scalars $d_1,\ldots,d_{j-1}$, such that

$$\mathbf{u}_j = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_{j-1}\mathbf{u}_{j-1},$$

or, equivalently,

$$-d_1\mathbf{u}_1 - \cdots - d_{j-1}\mathbf{u}_{j-1} + \mathbf{u}_j + 0\mathbf{u}_{j+1} + \cdots + 0\mathbf{u}_m = 0.$$

Therefore, (2.1) has the non trivial solution $c_1 = -d_1,\ldots,c_{j-1} = -d_{j-1}, c_j = 1, c_{j+1} = \cdots = c_m = 0$, which means that $\mathbf{u}_1,\ldots,\mathbf{u}_m$ are linearly dependent.

On the other hand, suppose that $\mathbf{u}_1,\ldots,\mathbf{u}_m$ are linearly dependent. Then there are scalars $a_1, a_2,\ldots,a_m$, not all equal to 0, such that

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_m\mathbf{u}_m = 0.$$

Not all $a_2, a_3,\ldots,a_m$ can be equal to 0. Otherwise, we would have $a_1\mathbf{u}_1 = 0$; but $\mathbf{u}_1 \neq 0$ which means that $a_1 = 0$. Then, in this case, $a_1 = \cdots = a_m = 0$ making the vectors linearly independent which contradicts our assumption. Let j be the largest element of $\{2,\ldots,m\}$ such that $a_j \neq 0$. Then,

$$\mathbf{u}_j = -\frac{a_1}{a_j}\mathbf{u}_1 - \cdots - \frac{a_{j-1}}{a_j}\mathbf{u}_{j-1}.$$

That is, there is some $\mathbf{u}_j$ that is a linear combination of the preceding vectors $\mathbf{u}_1,\ldots,\mathbf{u}_{j-1}$. ■

Now we come to a key result. It says that, if in a set of vectors one can be written as a linear combination of the remaining ones, then we can throw out that vector without changing the span of the original list.

Theorem 2.7

Let $S = \{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ be a set in a vector space V, and let $H = \text{Span } S$. If one of the vectors - say $\mathbf{u}_k$ - is a linear combination of the remaining vectors in S, then the set formed from S by removing $\mathbf{u}_k$ still spans H.

Proof. By rearranging the list of vectors in S, if necessary, we may suppose that $\mathbf{u}_m$ is a linear combination of $\mathbf{u}_1,\ldots,\mathbf{u}_{m-1}$, say

$$\mathbf{u}_m = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_{m-1}\mathbf{u}_{m-1}$$

Given any $\mathbf{w} \in H$, we may write

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m$$

for suitable scalars $c_1, c_2, \ldots, c_m$. Substituting the expression for $\mathbf{u}_m$ we obtain

$$\mathbf{w} = (c_1 + c_m a_1)\mathbf{u}_1 + (c_2 + c_m a_2)\mathbf{u}_2 + \cdots + (c_{m-1} + c_m a_{m-1})\mathbf{u}_{m-1}.$$

Thus, $\mathbf{w}$ is a linear combination of $\mathbf{u}_1,\ldots,\mathbf{u}_{m-1}$. This means that $\{\mathbf{u}_1,\ldots,\mathbf{u}_{m-1}\}$ spans H because $\mathbf{w}$ was an arbitrary element of H. ■

Consider a set of vectors $S = \{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ in a vector space V. We now know that for any vector $\mathbf{w} \in V$, the equation

$$w = x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_m\mathbf{u}_m, \tag{2.2}$$

has a solution if and only if $x \in \text{Span}S$. So the notion of generating set deals with the existence of solutions of vector equations such as (2.2). So, what about uniqueness? It turns out that vector equations such as (2.2) with unique solutions can be characterized in terms of the notion of linear independence.

Theorem 2.8

Let $S = \{\mathbf{u}_1,\ldots,\mathbf{u}_m\}$ be a set in a vector space V. Then each $\mathbf{w} \in \text{Span}S$ has a unique expression

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m,$$

for suitable scalars $c_1,\ldots,c_m$ if and only if S is a linearly independent set.

Proof. Suppose that each $\mathbf{w} \in \text{Span}S$ has a unique expression as a linear combination of the vectors in S. In particular, $\mathbf{0} \in \text{Span}S$ so

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m = 0,$$

which clearly is satisfied with $c_1 = c_2 = \ldots = c_m = 0$. But this expression is unique which means that S is a linearly independent set of vectors.

Now suppose that S is a linearly independent set. If a vector $\mathbf{w} \in \text{Span}S$ can be expressed as

$$\mathbf{w} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_m\mathbf{u}_m,$$

and

$$\mathbf{w} = d_1\mathbf{u}_1 + d_2\mathbf{u}_2 + \cdots + d_m\mathbf{u}_m,$$

then subtracting the expressions we get

$$0 = (c_1 - d_1)\mathbf{u}_1 + (c_2 - d_2)\mathbf{u}_2 + \cdots + (c_m - d_m)\mathbf{u}_m.$$

But S is a linearly independent set, so the last homogeneous equations has the unique solution $c_1 - d_1 = c_2 - d_2 = \ldots = c_m - d_m = 0$. Clearly, this means that $c_1 = d_1, c_2 = d_2, \ldots, c_m = d_m$. ■

Example 2.15

Consider the three vectors in $\mathbb{R}^3$:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Clearly, the homogeneous equation

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

has only the trivial solution $x_1 = x_2 = x_3 = 0$. Therefore, $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. are independent vectors. In fact, since it is obvious that $\mathbb{R}^3 = \text{Span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, we conclude that each vector in $\mathbb{R}^3$ has a unique expression as linear combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.

Example 2.16

Suppose that the scalars $c_0, c_1, \ldots, c_n$ satisfy

$$c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n = 0$$

This equality means that the polynomial on the left has the same values as the zero polynomial on the right. It is well known that a polynomial of degree n with more than n roots is the zero polynomial. Therefore, $c_0 = c_1 = \cdots = c_n = 0$. This proves that the set $\{1, t, t^2, \ldots, t^n\}$ is a linearly independent set in $\mathbf{P}_n$. It is obvious that $\mathbf{P}_n = \text{Span}\{1, t, t^2, \ldots, t^n\}$. Then each polynomial of degree at most n has a unique expression as a linear combination of $1, t, t^2, \ldots, t^n$.

Recall that any linear dependence relationship among the columns of a matrix A can be expressed in the form $A\mathbf{x} = \mathbf{0}$, where $\mathbf{x}$ is a column of weights. When A is row reduced to a matrix B, the columns of B are often totally different from the columns of A. However, the equations $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have exactly the same set of solutions. That is, the columns of A have *exactly the same linear dependence relationships as the columns of* B. So we have the following important observation:

Observation 2.2

Elementary row operations on a matrix preserve the linear dependence relations among the columns of the matrix.

The previous observation provides a useful shortcut for reducing a matrix to its row reduced echelon form when the linear dependence relations among the columns of a matrix are easy to detect.

Example 2.17

Consider the matrix

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}.$$

a. Find the smallest generating set for $\text{Col}A$.

b. Use the linear dependence relations among the columns of A for compute its row reduced echelon form.

These are the solutions:

a. Recall that $\text{Col}A = \text{Span}\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5\}$. By Theorem 2.7, the set of columns formed by removing those columns that are linear combinations of the remaining ones still spans $\text{Col}A$. Then, since

$$\mathbf{a}_2 = 4\mathbf{a}_1 \text{ and } \mathbf{a}_4 = 2\mathbf{a}_1 - \mathbf{a}_3,$$

we can remove $\mathbf{a}_2$ and $\mathbf{a}_4$ from the spanning set. Therefore, $\text{Col}A = \text{Span}\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$. Since $\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5$ are linearly independent (you should show this), no more columns should be removed. This means that $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ is the smallest generating set of $\text{Col}A$.

b. The columns of the row reduced echelon form $B = [\mathbf{b}_1 \mathbf{b}_2 \mathbf{b}_3 \mathbf{b}_4 \mathbf{b}_5]$ should have the same linear dependence relations as the columns of A. Therefore,

$$\mathbf{b}_2 = 4\mathbf{b}_1,\ \mathbf{b}_4 = 2\mathbf{b}_1 - \mathbf{b}_3,$$

and $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5$ are linearly independent. The pivot columns of a matrix in row reduced echelon form are linearly independent columns since they are the columns of the identity matrix. Moreover, all the rows of zeros should be at the bottom of the row reduced echelon matrix. From these two facts we deduce that $\mathbf{b}_1, \mathbf{b}_3, \mathbf{b}_5$ are pivots and correspond to the first three columns of the identity matrix, respectively. Putting everything together, we obtain

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

2.5 Bases and dimension

In a vector space V, sets that guarantee that each vector in V con be expressed uniquely in terms of these sets are one of the most important objects in linear algebra.

Definition 2.5

An ordered set $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ of vectors is a *basis* of V if the following two conditions hold:

1. $\mathcal{B}$ is a linearly independent set.
2. $V = \text{Span}\mathcal{B}$.

The following characterization of a basis is a direct consequence of Theorem 2.8.

Theorem 2.9

An ordered set $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ of vectors is a *basis* of V if and only if every vector $\mathbf{v} \in V$ can be written uniquely as a linear combination of the basis vectors.

We have already seen an example of a basis of $\mathbb{R}^3$ in Example 2.15. Such basis consisting of the columns of the identity matrix is called the *canonical basis* of $\mathbb{R}^3$.

Example 2.16 presents a basis of $\mathbf{P}_n$. Such basis consisting of consecutive powers of the variable t is called the *standard basis* of $\mathbf{P}_n$.

Let us consider less obvious examples of bases.

Example 2.18

Let

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}.$$

Determine if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathbb{R}^3$.

We must show that these vector are linearly independent and generate $\mathbb{R}^3$. Both of these conditions hold if and only if the equation $[\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3]\mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b} \in \mathbb{R}^3$. This equation has a solution for each $\mathbf{b}$ if each row of the matrix $A = [\mathbf{v}_1 \mathbf{v}_2 \mathbf{v}_3]$ has a pivot and the solution is unique if each column of A is a pivot column. In other words, the equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for each $\mathbf{b}$ if $A \sim I$. So we row reduce A to its row reduced echelon form:

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

Therefore $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis of $\mathbb{R}^3$.

Example 2.19

Let

$$p_0(t) = (1-t)^3,\ p_1(t) = 3t(1-t)^2,\ p_2(t) = 3t^2(1-t),\ p_3(t) = t^3.$$

Determine if $\{p_0(t), p_1(t), p_2(t), p_3(t)\}$ is a basis for $\mathbf{P}_3$.

We need to determine if each polynomial in $\mathbf{P}_3$ can be expressed uniquely as a linear combination of $p_0(t), p_1(t), p_2(t), p_3(t)$. Let $q(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3$ and suppose that the scalars c_0, c_1, c_2, c_3 satisfy

$$q(t) = c_0 p_0(t) + c_1 p_1(t) + c_2 p_2(t) + c_3 p_3(t).$$

Comparing the coefficients multiplying each power of t on both sides of the equality, we obtain

$$\begin{aligned} c_0 &= a_0, \\ -3c_0 + 3c_1 &= a_1, \\ 3c_0 - 6c_1 + 3c_2 &= a_2, \\ -c_0 + 3c_1 - 3c_2 + c_3 &= a_3, \end{aligned}$$

or in matrix form,

$$A\mathbf{c} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 \\ 3 & -6 & 3 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

It is not difficult to see that the matrix A has a pivot in every row and every column, which means that the equation above has a unique solution for each choice of the polynomial $q(t)$. Therefore, $\{p_0(t), p_1(t), p_2(t), p_3(t)\}$ is a basis for $\mathbf{P}_3$.

The following theorem characterizes an important basis of the column space of a matrix.

Theorem 2.10

The pivot columns of a matrix A form a basis for $\mathrm{Col}A$.

Proof. Let B be the reduced echelon form of A. The set of pivot columns of B is linearly independent, for no vector in the set is a linear combination of the vectors that precede it. Since A is row equivalent to B, the pivot columns of A are linearly independent as well, because any linear dependence relation among the columns of A corresponds to a linear dependence relation among the columns of B. For this same reason, every non pivot column of A is a linear combination of the pivot columns of A. Thus the non pivot columns of A may be discarded from the spanning set for $\mathrm{Col}A$. This leaves the pivot columns of A as a basis for $\mathrm{Col}A$. ■

Warning: Be careful to use *pivot columns of A itself* for the basis of $\text{Col}A$. Row operations can change the column space of a matrix. We illustrate this in the following example.

Example 2.20

Consider again the matrix given in Example 2.17:

$$A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}.$$

We showed that

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

It can be seen from the row reduced echelon form that the first, third, and last columns are pivot columns, which means that $\{\mathbf{a}_1, \mathbf{a}_3, \mathbf{a}_5\}$ is a basis for $\text{Col}A$. However, the columns of the row reduced echelon form all have zeros in their last entries, so they cannot span the column space of A.

Every generating set in a vector space can be reduced to a basis of the vector space. Indeed, removing vectors that are linear combinations of the remaining vectors will not change the span of the original set. The deletion of vectors from the generating set must stop when the set becomes linearly independent. If additional vectors are deleted, it will not be a linear combination of the remaining vectors, and hence the smaller set will no longer span V. Thus a basis is a generating set that is as small as possible.

Also, every linearly independent list of vectors can be extended to a basis of the vector space. This is done by adjoining a vector that is not in the span of the original list. The extended list of vectors is still linearly independent because the new vector cannot be written as a linear combination of the preceding ones. This step is repeated until a generating set of V is obtained. The addition of vectors must stop as soon as a generating set for V is obtained. For if one more vector is added, then the new set cannot be linearly independent because the old set spans V, and the new vector is therefore a linear combination of the vectors preceding it. Thus, a basis is a linearly independent set that is as large as possible.

Remark 2.1

Let $\mathcal{B} = \{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$ be a basis of a vector space V. Then every generating set of V has at least n elements, and every linearly independent set in V has at most n elements.

We say that a vector space V is *finite-dimensional* if it has a basis with a finite number of elements.

The notion of basis would be useless if different basis of the same vector space V had different numbers of elements. Fortunately, that turns out not to be the case.

Theorem 2.11

Any two bases of a finite-dimensional vector space have the same number of elements.

Proof. Suppose that V is a finite-dimensional vector space. Let $\mathcal{B}_1$ and $\mathcal{B}_2$ be two bases of V. Then $\mathcal{B}_1$ and $\mathcal{B}_2$ are linearly independent in V, so the number of vectors in $\mathcal{B}_1$ is at most the number of elements in $\mathcal{B}_2$. Moreover, $\mathcal{B}_1$ and $\mathcal{B}_2$ span V, then the number of vectors in $\mathcal{B}_1$ is at least the number of elements in $\mathcal{B}_2$. Thus the number of vectors in $\mathcal{B}_1$ and $\mathcal{B}_2$ must be equal. ■

Now we know that the number of vectors in a basis of a finite-dimensional vector space is an inherent property of V (that is, it does not depend on the choice of basis). Therefore we give the following definition.

Definition 2.6

The dimension of a finite-dimensional vector space, denoted by $\dim V$, is the number of vectors in any basis of V. In other words, if $\mathcal{B} = \{\mathbf{u}_1,\ldots,\mathbf{u}_n\}$ is a basis of V, then $\dim V = n$.

If V is a finite-dimensional vector space, then every generating set of vectors of V with $n = \dim V$ elements is a basis of V. For suppose that $\mathbf{v}_1,\ldots,\mathbf{v}_n$ span V. However, every basis of V has n elements, so in this case the reduction must be the trivial one, meaning that no elements are deleted from $\mathbf{v}_1,\ldots,\mathbf{v}_n$. In other words, $\mathbf{v}_1,\ldots,\mathbf{v}_n$ is a basis of V, as desired.

Moreover, every linearly independent set of vectors in V with n elements is a basis of V. For suppose that $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ is a linearly independent set. Then this set can be extended to a basis of V. However, every basis of V has n elements, so in this case the extension must be the trivial one, meaning that no vectors are adjoined to $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$. In other words, $\{\mathbf{v}_1,\ldots,\mathbf{v}_n\}$ is a basis of V.

Example 2.21

The canonical basis for $\mathbb{R}^n$ contains n vectors, so $\dim \mathbb{R}^n = n$. The standard polynomial basis $\{1, t, t^2, \ldots, t^n\}$ shows that $\dim \mathbf{P}_n = n+1$.

Example 2.22

Find the dimension of the subspace

$$H = \left\{ \begin{bmatrix} a-3b+6c \\ 5a+4d \\ b-2c-d \\ 5d \end{bmatrix} : a,b,c,d \in \mathbb{R} \right\}.$$

The strategy is to produce a basis for H by finding a generating set and then reduce it to a linearly independent generating set. The number of elements in the final set is equal to $\dim H$.

First, we find a generating set for H. A generic vector in H can be written as

$$\begin{bmatrix} a-3b+6c \\ 5a+4d \\ b-2c-d \\ 5d \end{bmatrix} = a\begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix} + d\begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix}, \; a,b,c,d \in \mathbb{R}.$$

Then

$$H = \text{Span}\left\{ \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix} \right\}.$$

This is not a basis, however. Notice that the third vector is a multiple of the second vector, so it can be removed from the set. The remaining vectors are linearly independent. Thus, we have found a basis for H; namely,

$$\left\{ \begin{bmatrix} 1 \\ 5 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 4 \\ -1 \\ 5 \end{bmatrix} \right\}.$$

Since it has three vectors, we conclude that $\dim H = 3$.

Example 2.23

Find the dimensions of the null space and the column space of the matrix

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

The pivot columns of A form a basis of $\mathrm{Col}A$, so the dimension of the column space is the number of pivot columns. Row reduce the matrix to row reduced echelon form:

$$A \sim B = \begin{bmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Therefore, $\dim\mathrm{Col}A = 2$.

Now we find a basis for $\mathrm{Nul}A$. If we row reduce the augmented matrix $\begin{bmatrix} A & 0 \end{bmatrix}$ we obtain $\begin{bmatrix} B & 0 \end{bmatrix}$. The corresponding system of linear equations is:

$$\begin{aligned} x_1 - 2x_2 - x_4 + 3x_5 &= 0, \\ x_3 + 2x_4 - 2x_5 &= 0, \\ 0 &= 0, \end{aligned}$$

with x_2, x_4, x_5 free. Its general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2a + b - 3c \\ a \\ -2b + 2c \\ b \\ c \end{bmatrix} = a \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \quad a, b, c \in \mathbb{R}.$$

We already know that the number of generating vectors produced with this method is equal to the number of free variables. Furthermore, this method produces automatically a set of linearly independent vectors because the free variables are the weights on the spanning vectors. For instance, look at the second, fourth, and fifth entries of the three vectors above, and note that 0 is a linear combination of these vectors if and only if $a = b = c = 0$. Therefore, these vectors form a basis for $\mathrm{Nul}A$ and $\dim\mathrm{Nul}A = 3$.

2.6 Coordinates

The purpose of bases in vector spaces is to provide a method of computation, and we are going to learn to use them in this section. We will consider two

topics: how to express a vector in terms of a given basis, and how to relate two different bases of the same vector space.

Let V be a vector space with $\dim V = n$ over a field $\mathbb{K}$, and suppose that

$$\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$$

is a basis of V. Then any vector $\mathbf{v} \in V$ can be expressed uniquely as a linear combination of the basis vectors in $\mathcal{B}$, say

$$\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n.$$

The n scalars $a_1, a_2, \ldots, a_n$ are called the *coordinates* of $\mathbf{v}$ relative to $\mathcal{B}$: and they form the vector

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix},$$

called the *coordinate vector* of $\mathbf{v}$ relative to $\mathcal{B}$.

Example 2.24

Let $\mathcal{E}_3 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the canonical basis for $\mathbb{R}^3$. That is,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

For any vector $\mathbf{x} \in \mathbb{R}^3$, it is obvious that $[\mathbf{x}]_{\mathcal{E}_3} = \mathbf{x}$.

Example 2.25

Let $\mathcal{S}_n = \{1, t, t^2, \ldots, t^n\}$ be the standard basis of $\mathbf{P}_n$. For any polynomial $q(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$, we have

$$[q(t)]_{\mathcal{S}_n} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Example 2.26

Let

$$\mathbf{v} = \begin{bmatrix} 1 \\ 4 \end{bmatrix},\ \mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},\ \mathbf{u}_2 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$$

Find the coordinates of $\mathbf{v}$ relative to the basis $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$.

We must find scalars a_1, a_2 such that $\mathbf{v} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2$, or, in matrix form

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Row reduce the augmented matrix:

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

Therefore,

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 7 \\ -2 \end{bmatrix}.$$

We want to emphasize the $\mathbf{v} \neq [\mathbf{v}]_{\mathcal{B}}$. The correct relation between $\mathbf{v}$ and its coordinates $[\mathbf{v}]_{\mathcal{B}}$ is $\mathbf{v} = 7\mathbf{u}_1 - 2\mathbf{u}_2$.

The following example is to remind you that the material in this course is applicable to every vector space, not just to the well-known ones like $\mathbb{R}^n$ and $\mathbf{P}_n$.

Example 2.27

Let us consider the vector space V generated by the set of linearly independent functions $\mathcal{B} = \left\{e^{-i\pi t}, e^{-i\frac{\pi}{2}t}, 1, e^{i\frac{\pi}{2}t}, e^{i\pi t}\right\}$. Observe that $\mathcal{B}$ is a basis for V and $\dim V = 5$. Some examples of vector in V are

$$\cos\left(\frac{\pi n}{2}t\right),\ \sin\left(\frac{\pi n}{2}t\right),\ n = -2, -1, 0, 1, 2.$$

We will use Euler's formula

$$e^{i\theta} = \cos\theta + i\sin\theta,\ \theta \in \mathbb{R},\ i = \sqrt{-1},$$

to find the coordinates of some of these vectors relative to $\mathcal{B}$.

By Euler's identity

$$\cos\left(\pi t\right)=\frac{e^{i\pi t}+e^{-i\pi t}}{2}\quad\text{and}\quad\sin\left(\frac{\pi}{2}t\right)=\frac{e^{i\frac{\pi}{2}t}-e^{-i\frac{\pi}{2}t}}{2i}.$$

Therefore,

$$[\cos\left(\pi t\right)]_{\mathcal{B}}=\begin{bmatrix}1/2\\0\\0\\0\\1/2\end{bmatrix}\quad\text{and}\quad\left[\sin\left(\frac{\pi}{2}t\right)\right]_{\mathcal{B}}=\begin{bmatrix}0\\-\frac{1}{2i}\\0\\\frac{1}{2i}\\0\end{bmatrix}.$$

Another example is

$$\left[1+2\cos\left(\pi t\right)-2i\sin\left(\frac{\pi}{2}t\right)\right]_{\mathcal{B}}=\begin{bmatrix}1\\1\\1\\-1\\1\end{bmatrix}.$$

The coordinates relative to any basis have the property that they preserve the main structure of vector spaces: linear combinations.

Theorem 2.12

Let $\mathcal{B}=\left\{\mathbf{u}_1,\ldots,\mathbf{u}_n\right\}$ be a basis for a vector space V. For any $\mathbf{v},\mathbf{w}\in V$ and scalar r:

1. $\left[\mathbf{v}+\mathbf{w}\right]_{\mathcal{B}}=\left[\mathbf{v}\right]_{\mathcal{B}}+\left[\mathbf{w}\right]_{\mathcal{B}}$
2. $\left[r\mathbf{v}\right]_{\mathcal{B}}=r\left[\mathbf{v}\right]_{\mathcal{B}}$

Proof. Take two typical vectors in V, say

$$\begin{aligned}\mathbf{v}&=c_1\mathbf{u}_1+c_2\mathbf{u}_2+\cdots+c_n\mathbf{u}_n\\\mathbf{w}&=d_1\mathbf{u}_1+d_2\mathbf{u}_2+\cdots+d_n\mathbf{u}_n\end{aligned}$$

Then, using vector operations,

$$\mathbf{v}+\mathbf{w}=\left(c_1+d_1\right)\mathbf{u}_1+\left(c_2+d_2\right)\mathbf{u}_2+\cdots+\left(c_n+d_n\right)\mathbf{u}_n.$$

It follows that

$$[\mathbf{v}+\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} c_1+d_1 \\ c_2+d_2 \\ \vdots \\ c_n+d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = [\mathbf{v}]_{\mathcal{B}} + [\mathbf{w}]_{\mathcal{B}}.$$

If r is any scalar, then

$$r\mathbf{v} = (rc_1)\mathbf{u}_1 + (rc_2)\mathbf{u}_2 + \cdots + (rc_n)\mathbf{u}_n.$$

So

$$[r\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} r\,c_1 \\ r\,c_2 \\ \vdots \\ r\,c_n \end{bmatrix} = r\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = r[\mathbf{v}]_{\mathcal{B}}.$$

■

We now come to a very important computational method: *change of basis*. Identifying vectors in V with column vectors in $\mathbb{K}^n$ is useful when a natural basis is presented to us, but not when the given basis is poorly suited to the problem at hand. In that case, we will want to change basis. So let us suppose that we are given two bases for the same vector space V, say $\mathcal{B} = \{\mathbf{b}_1,\ldots,\mathbf{b}_n\}$ and $\mathcal{C} = \{\mathbf{c}_1,\ldots,\mathbf{c}_n\}$. We will think of $\mathcal{B}$ as the *old* basis, and $\mathcal{C}$ as a *new* basis. There are two computations which we wish to clarify. We ask first: How are the two bases related? Secondly, a vector $\mathbf{v} \in V$ will have coordinates relative to each of these bases, but of course they will be different. So we ask: How are the two coordinate vectors related? These are the computations called change of basis.

We begin by noting that since the new basis spans V, every vector in the old basis $\mathcal{B}$ is a linear combination of the new basis $\mathcal{C}$. So we can write the coordinates of each vector in $\mathcal{B}$ relative to $\mathcal{C}$:

$$[\mathbf{b}_1]_{\mathcal{C}},\ [\mathbf{b}_2]_{\mathcal{C}},\ \cdots\ [\mathbf{b}_n]_{\mathcal{C}}.$$

Now let

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

be the coordinate vector of $\mathbf{v}$ relative to the old basis $\mathcal{B}$. That is,

$$\mathbf{v} = v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n.$$

Using Theorem 2.12, we compute the coordinate vector of $\mathbf{v}$ relative to the new basis $\mathcal{C}$:

$$[\mathbf{v}]_{\mathcal{C}} = \left[v_1\mathbf{b}_1 + v_2\mathbf{b}_2 + \cdots + v_n\mathbf{b}_n\right]_{\mathcal{C}} = v_1\left[\mathbf{b}_1\right]_{\mathcal{C}} + v_2\left[\mathbf{b}_2\right]_{\mathcal{C}} + \cdots + v_n\left[\mathbf{b}_n\right]_{\mathcal{C}}.$$

If we define the matrix

$$P_{\mathcal{C}}^{\mathcal{B}} = \left[\left[\mathbf{b}_1\right]_{\mathcal{C}} \quad \left[\mathbf{b}_2\right]_{\mathcal{C}} \quad \cdots \quad \left[\mathbf{b}_n\right]_{\mathcal{C}}\right],$$

then we can write

$$[\mathbf{v}]_{\mathcal{C}} = P_{\mathcal{C}}^{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}.$$

Following the same argument as above but interchanging the roles of the bases $\mathcal{B}$ and $\mathcal{C}$, we obtain

$$[\mathbf{v}]_{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{C}}[\mathbf{v}]_{\mathcal{C}},$$

where

$$P_{\mathcal{B}}^{\mathcal{C}} = \left[\left[\mathbf{c}_1\right]_{\mathcal{B}} \quad \left[\mathbf{c}_2\right]_{\mathcal{B}} \quad \cdots \quad \left[\mathbf{c}_n\right]_{\mathcal{B}}\right].$$

Then we make the following observation: $P_{\mathcal{C}}^{\mathcal{B}}$ is an invertible matrix and $\left(P_{\mathcal{C}}^{\mathcal{B}}\right)^{-1} = P_{\mathcal{B}}^{\mathcal{C}}$.

Recapitulating, we have an invertible matrix $P_{\mathcal{C}}^{\mathcal{B}}$, called the ***matrix of change of basis*** from $\mathcal{B}$ to $\mathcal{C}$, whose columns are the coordinate vectors of each element of the old basis $\mathcal{B}$ relative to the new basis $\mathcal{C}$, which transforms $[\mathbf{v}]_{\mathcal{B}}$ into $[\mathbf{v}]_{\mathcal{C}}$. Moreover, the columns of $\left(P_{\mathcal{C}}^{\mathcal{B}}\right)^{-1}$ are the coordinate vectors of each element of the new basis $\mathcal{C}$ relative to the old basis $\mathcal{B}$.

An important consequence of the computations described above is that, given any basis $\mathcal{B}$ of V, it is possible to obtain a new basis $\mathcal{C} = \{\mathbf{c}_1,\ldots,\mathbf{c}_n\}$ by choosing any invertible $n \times n$ matrix $P = [\mathbf{a}_1\ \mathbf{a}_2 \ldots \mathbf{a}_n]$ and letting its columns be the coordinate vectors of the vectors in the new basis relative to the old one. That is,

$$\left[\mathbf{c}_1\right]_{\mathcal{B}} = \mathbf{a}_1, \left[\mathbf{c}_2\right]_{\mathcal{B}} = \mathbf{a}_2, \cdots, \left[\mathbf{c}_n\right]_{\mathcal{B}} = \mathbf{a}_n.$$

In this case, $P = P_{\mathcal{B}}^{\mathcal{C}}$.

Example 2.28

Let $\mathcal{E}_3 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the canonical basis for $\mathbb{R}^3$. That is,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Recall that for any vector $\mathbf{x} \in \mathbb{R}^3$, it is obvious that $[\mathbf{x}]_{\mathcal{E}_3} = \mathbf{x}$. Furthermore, consider another basis for $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$, where

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \ \mathbf{b}_2 = \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix}, \ \mathbf{b}_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}.$$

The change of basis matrix from $\mathcal{B}$ to $\mathcal{E}_3$ is

$$P_{\mathcal{E}_3}^{\mathcal{B}} = \begin{bmatrix} [\mathbf{b}_1]_{\mathcal{E}_3} & [\mathbf{b}_2]_{\mathcal{E}_3} & [\mathbf{b}_3]_{\mathcal{E}_3} \end{bmatrix} = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}.$$

We can also compute $P_{\mathcal{B}}^{\mathcal{E}_3} = \left(P_{\mathcal{E}_3}^{\mathcal{B}}\right)^{-1}$:

$$P_{\mathcal{B}}^{\mathcal{E}_3} = \begin{bmatrix} [\mathbf{e}_1]_{\mathcal{B}} & [\mathbf{e}_2]_{\mathcal{B}} & [\mathbf{e}_3]_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \frac{5}{2} \\ 0 & 0 & -\frac{1}{2} \\ 2 & -1 & 3 \end{bmatrix}.$$

It is interesting to verify that the columns of $P_{\mathcal{B}}^{\mathcal{E}_3}$ are, indeed, the coordinate vectors of the elements of the canonical basis relative to $\mathcal{B}$:

$$\mathbf{e}_1 = \mathbf{b}_1 + 2\mathbf{b}_3, \ \mathbf{e}_2 = -\mathbf{b}_3, \ \mathbf{e}_3 = \frac{5}{2}\mathbf{b}_1 - \frac{1}{2}\mathbf{b}_2 + 3\mathbf{b}_3.$$

Example 2.29

Let

$$\mathbf{b}_1 = \begin{bmatrix} -9 \\ 1 \end{bmatrix}, \ \mathbf{b}_2 = \begin{bmatrix} -5 \\ -1 \end{bmatrix}, \ \mathbf{c}_1 = \begin{bmatrix} 1 \\ -4 \end{bmatrix}, \ \mathbf{c}_2 = \begin{bmatrix} 3 \\ -5 \end{bmatrix},$$

and consider the bases for $\mathbb{R}^2$ given by $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$. Find the change of basis matrices $P_{\mathcal{C}}^{\mathcal{C}}$ and $P_{\mathcal{B}}^{\mathcal{C}}$.

First, the matrix $P_{\mathcal{C}}^{\mathcal{B}}$ involves the coordinate vectors of $\mathbf{b}_1$ and $\mathbf{b}_2$ relative to $\mathcal{C}$. Let $\left[\mathbf{b}_1\right]_{\mathcal{C}} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and $\left[\mathbf{b}_2\right]_{\mathcal{C}} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Then, by defintion

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{b}_1 \text{ and } \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{b}_2.$$

To solve both systems simultaneously, augment the coefficient matrix with $\mathbf{b}_1$ *and* $\mathbf{b}_2$, and row reduce:

$$\begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -9 & -5 \\ -4 & -5 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 6 & 4 \\ 0 & 1 & -5 & -3 \end{bmatrix}.$$

Thus,

$$\left[\mathbf{b}_1\right]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -5 \end{bmatrix} \text{ and } \left[\mathbf{b}_2\right]_{\mathcal{C}} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}.$$

The matrix $P_{\mathcal{C}}^{\mathcal{B}}$ is therefore

$$P_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} \left[\mathbf{b}_1\right]_{\mathcal{C}} & \left[\mathbf{b}_2\right]_{\mathcal{C}} \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -5 & -3 \end{bmatrix}.$$

We can also compute $P_{\mathcal{B}}^{\mathcal{C}} = \left(P_{\mathcal{C}}^{\mathcal{B}}\right)^{-1}$:

$$P_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} \left[\mathbf{c}_1\right]_{\mathcal{B}} & \left[\mathbf{c}_2\right]_{\mathcal{B}} \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} & -2 \\ \frac{5}{2} & 3 \end{bmatrix}.$$

Example 2.30

Let

$$p_0(t) = (1-t)^3,\ p_1(t) = 3t(1-t)^2,\ p_2(t) = 3t^2(1-t),\ p_3(t) = t^3.$$

and consider the bases for $\mathbf{P}_3$ given by $\mathcal{B} = \{p_0(t), p_1(t), p_2(t), p_3(t)\}$ and $\mathcal{S}_3 = \{1, t, t^2, t^3\}$. Use a change of basis matrix to write the polynomial $q(t) = t^2 + 1$ as a linear combination of the basis $\mathcal{B}$.

Observe that the polynomial $q(t)$ is written as a linear combination of the basis $\mathcal{S}_3$. So we need to compute the change of basis matrix $P_{\mathcal{B}}^{\mathcal{S}_3}$. We will use this matrix to compute $[q(t)]_{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{S}_3}[q(t)]_{\mathcal{S}_3}$ since we know that $[q(t)]_{\mathcal{S}_3} = \begin{bmatrix} 1 & 0 & 1 & 0 \end{bmatrix}$.

We will compute $P_{\mathcal{B}}^{\mathcal{S}_3}$ in two steps. We first compute $P_{\mathcal{S}_3}^{\mathcal{B}}$:

Now, we compute $P_{\mathcal{B}}^{\mathcal{S}_3} = \left(P_{\mathcal{S}_3}^{\mathcal{B}}\right)^{-1}$:

$$P_{\mathcal{B}}^{\mathcal{S}_3} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1/3 & 0 & 0 \\ 1 & 2/3 & 1/3 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

Since $[q(t)]_{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{S}_3}[q(t)]_{\mathcal{S}_3} = \begin{bmatrix} 1 & 1 & \frac{4}{3} & 2 \end{bmatrix}$, we have $q(t) = p_0(t) + p_1(t) + \frac{4}{3}p_2(t) + 2p_3(t)$.

3. LINEAR MAPPINGS AND DIAGONALIZATION

In this chapter, we first consider functions that transform elements from one vector space into another one. These functions comply with the fundamental properties of linearity, already considered above, and are therefore called *linear mappings*. In many instances, we are able to formulate matrices to represent linear mappings. Afterwards, we consider a concept that has many applications in the sciences and engineering: the *eigenvector*. Such a vector is invariant under a linear mapping (except for a constant factor called *eigenvalue*) and for certain linear mappings (and associated matrices) it will be possible to find a basis of the vector space such that the associated matrix is diagonal, known as the *diagonalization* of the matrix.

3.1 Linear mappings and matrices

3.1.1 Introduction and basic properties

Example 3.1

Let us start with an example: Let $\mathbf{p} = [x_1, y_1, z_1]$ be a point in the vector space $\mathbb{R}^3$ (for simplicity, we use row vectors here). Then, there exists a unique projection to the vector space $\mathbb{R}^2$ spanned by x and y : such a point is simply $\mathbf{q} = [x_1, y_1]$ (in a three-dimensional Cartesian coordinate system, $\mathbf{q}$ lies in the plane $[x, y, 0]$). We regard $[x_1, y_1, z_1] \to [x_1, y_1]$ as a *mapping* T from $\mathbb{R}^3$ to $\mathbb{R}^2$ and write

$$T : \mathbb{R}^3 \to \mathbb{R}^2 \quad T(\mathbf{p}) = \mathbf{q}.$$

This mapping conserves *linearity*:

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{bmatrix} \to \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

$$\alpha \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} \alpha x_2 \\ \alpha y_2 \\ \alpha z_2 \end{bmatrix} \to \begin{bmatrix} \alpha x_1 \\ \alpha y_1 \end{bmatrix} = \alpha \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}, \ \alpha \in \mathbb{R}$$

We see that *linear mappings* are functions defined on vector spaces that preserve linear combinations. Other names (mostly used synonymously, but strongly depending on the context) are linear transformations, linear operators, or linear maps.

Definition 3.1

Given two vector spaces V and W, we say that $T:V\to W$ is a linear mapping if $\forall \mathbf{u},\mathbf{v}\in V$ and $\forall \alpha\in\mathbb{R}$ it verifies:

(a) $T\left(\mathbf{u}+\mathbf{v}\right)=T\left(\mathbf{u}\right)+T\left(\mathbf{v}\right)$ (additivity).

(b) $T\left(\alpha\mathbf{u}\right)=\alpha T\left(\mathbf{u}\right)$ (homogeneity).

These two properties can be combined in a single statement:

$$T\left(\alpha\mathbf{u}+\beta\mathbf{v}\right)=\alpha T\left(\mathbf{u}\right)+\beta T\left(\mathbf{v}\right)\ \forall \mathbf{u},\mathbf{v}\in V \text{ and } \forall \alpha,\beta\in\mathbb{R}$$

Example 3.2

The mapping $D:\mathbf{P}\to\mathbf{P}$ defined by $D\left(p\left(x\right)\right)=p'\left(x\right)$ is linear.

Proof. We show that the differentiation of polynomials is a linear mapping. Let f and g represent polynomials, written in the canonical basis of polynomials $\mathbf{P}$. Then, we use knowledge of calculus: $D\left(f+g\right)=D\left(f\right)+D\left(g\right)$ and $D\left(\alpha f\right)=\alpha D\left(f\right)$ and differentiation of polynomials (as actually of any differentiable function) is linear. Note that we need to have a well-defined vector space. ■

Example 3.3

The mapping $T:\mathbb{R}^2\to\mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)=\begin{bmatrix}y\\x\end{bmatrix}$$

is linear.

Proof. For all vectors $\begin{bmatrix}x\\y\end{bmatrix}\in\mathbb{R}^2$ and all $\alpha\in\mathbb{R}$, we have:

$$T\left(\begin{bmatrix}x_1\\y_1\end{bmatrix}+\begin{bmatrix}x_2\\y_2\end{bmatrix}\right)=T\left(\begin{bmatrix}x_1+x_2\\y_1+y_2\end{bmatrix}\right)=\begin{bmatrix}y_1+y_2\\x_1+x_2\end{bmatrix}=\begin{bmatrix}y_1\\x_1\end{bmatrix}+\begin{bmatrix}y_2\\x_2\end{bmatrix}=T\left(\begin{bmatrix}x_1\\y_1\end{bmatrix}\right)+T\left(\begin{bmatrix}x_2\\y_2\end{bmatrix}\right)$$

and

$$T\left(\alpha\begin{bmatrix}x_1\\y_1\end{bmatrix}\right)=T\left(\begin{bmatrix}\alpha x_1\\\alpha y_1\end{bmatrix}\right)=\begin{bmatrix}\alpha y_1\\\alpha x_1\end{bmatrix}=\alpha\begin{bmatrix}y_1\\x_1\end{bmatrix}=\alpha T\left(\begin{bmatrix}x_1\\y_1\end{bmatrix}\right)$$

■

This simple example allows us to draw a link to the first chapter. The linear mapping of this example is interchanging x and y components of a vector. This can also be achieved by multiplying a matrix to the vector:

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}.$$

Hence, a linear mapping can be represented (for some relevant cases) by

$$T(\mathbf{u}) = A\mathbf{u},$$

where A is a suitable matrix. We will come back to the matrix associated to a linear mapping below.

Example 3.4

The mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} y \\ x^2 \end{bmatrix}$$

is not linear.

Proof. It is sufficient to show that the additivity property is not fulfilled:

$$T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) = \begin{bmatrix} y_1 + y_2 \\ (x_1 + x_2)^2 \end{bmatrix},$$

while

$$T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} y_1 \\ x_1^2 \end{bmatrix} + \begin{bmatrix} y_2 \\ x_2^2 \end{bmatrix} = \begin{bmatrix} y_1 + y_2 \\ x_1^2 + x_2^2 \end{bmatrix}$$

■

Theorem 3.1

Let $T : V \to W$ be a linear mapping. Then:

(a) $T(0) = 0$.

(b) $T(-\mathbf{u}) = -T(\mathbf{u})$.

(c) $T(a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_n\mathbf{u}_n) = a_1T(\mathbf{u}_1) + a_2T(\mathbf{u}_2) + \cdots + a_nT(\mathbf{u}_n)$.

Observe that in part (a), the nullvector on the left hand side is the nullvector from V, on the left hand side there is the nullvector from W. In particular, part (a) means that for a linear mapping, the nullvector maps to the nullvector. In part (c), the number of vectors n is not specified.

Example 3.5

The mapping $T:\mathbb{R}^2 \to \mathbb{R}^2$ defined by

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

is not linear since

$$T\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

3.1.2 Kernel and image of a linear mapping

Definition 3.2

Let $T:V \to W$ be a linear mapping. We define the kernel and image of T, respectively, as follows:

$$\begin{aligned} \text{Ker}T &= \left\{\mathbf{x} \in V : T(\mathbf{x}) = 0\right\}, \\ \text{Im}T &= \left\{T(\mathbf{x}) \in W : \mathbf{x} \in V\right\}. \end{aligned}$$

Observe that $\text{Ker}T$ is a subspace of V and $\text{Im}T$ is a subspace of W. Once we establish the associated matrix for a linear mapping, we will see how the kernel is analogous to the nullspace and the image to the column space of the matrix.

Theorem 3.2

Let $T:V \to W$ be a linear mapping and let $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ be a system of generators of V. Then, $\{T(\mathbf{u}_1), T(\mathbf{u}_2), \ldots, T(\mathbf{u}_n)\}$ is a system of generators of $\text{Im}T$.

Example 3.6

Find the kernel and image of the linear mapping $T:\mathbb{R}^3 \to \mathbb{R}^3$ defined by

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

For the vector space of this example, we know a very convenient system of generators: the canonical basis of $\mathbb{R}^3$:

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}.$$

We apply the linear mapping to these vectors and obtain three vectors

$$\left\{T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right),T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right),T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right)\right\}=\left\{\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}0\\1\\2\end{bmatrix},\begin{bmatrix}1\\0\\1\end{bmatrix}\right\}.$$

Two of these vectors are linearly dependent, and hence we conclude

$$\mathrm{Im}T=\mathrm{Span}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}0\\1\\2\end{bmatrix}\right\}$$

The kernel is formed by the vectors that map to the nullvector.

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

From that set of equations we obtain

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

with z being arbitrary, and therefore

$$\mathrm{Ker}T=\mathrm{Span}\left\{\begin{bmatrix}1\\0\\-1\end{bmatrix}\right\}$$

Before we proceed, we have to review some general classification of functions:

Definition 3.3

Let $f: A \to B$ be a function. Then,

(a) f is injective if and only if $x \neq y$ implies $f(x) \neq f(y) \forall x, y \in A$.

(b) f is surjective if and only if $\forall b \in B, \exists a \in A$ such that $f(a) = b$.

(c) f is bijective if and only if f is injective and f is surjective.

With this classification it is now possible to confirm the following properties for injective and surjective (and hence, bijective) linear mappings:

Theorem 3.3

Let $T: V \to W$ be a linear mapping. Then:

(a) T is injective if and only if $\text{Ker}T = \{0\}$.

(b) T is surjective if and only if $\text{Im}T = W$.

(c) T is bijective if and only if $\text{Ker}T = \{0\}$ and $\text{Im}T = W$.

There are alternative characterizations for injective, surjective and bijective linear mappings, provided by the following theorem:

Theorem 3.4

Let $T: V \to W$ be a linear mapping. Then:

(a) T is injective if and only if for each independent set $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$, the set $\{T(\mathbf{u}_1), T(\mathbf{u}_2), \ldots, T(\mathbf{u}_n)\}$ is linearly independent.

(b) T is surjective if and only for each system of generators of $V, \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$, the set $\{T(\mathbf{u}_1), T(\mathbf{u}_2), \ldots, T(\mathbf{u}_n)\}$ is a system of generators of W.

(c) T is bijective if and only if for each basis of $V, \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$, the set $\{T(\mathbf{u}_1), T(\mathbf{u}_2), \ldots, T(\mathbf{u}_n)\}$ is a basis for W.

Notice that n for the three cases is not necessarily the same number, as the number of independent vectors can be *smaller* than the dimension, and the number of generatoring vectors can be *larger* than the dimension, respectively.

Let us now move on to generate linear mappings from basic operations.

Definition 3.4

Let $f, g : V \to W$ be two linear mappings and $\lambda \in \mathbb{R}$. Then, we can define the following operations:

(a) $f + g : V \to W, \big(f + g\big)\big(\mathbf{x}\big) = f\big(\mathbf{x}\big) + g\big(\mathbf{x}\big)$

(b) $\lambda f : V \to W, \big(\lambda f\big)\big(\mathbf{x}\big) = \lambda f\big(\mathbf{x}\big)$

Theorem 3.5

With the operations of linear mappings, the set of linear mappings between two vector spaces V and W is itself a vector space.

As for general functions, we can also define the composition of linear mappings and find that its result is also a linear mapping, by the next theorem.

Theorem 3.6

Let $f : V \to W$ and $g : W \to U$ be two linear mappings. Then, their composition $g \circ f : V \to U$, defined by $\big(g \circ f\big)\big(\mathbf{x}\big) = g\Big(f\big(\mathbf{x}\big)\Big)$ is also a linear mapping.

We also would like to define an inverse linear mapping and, for that, we use the general result that if a function is bijective, it has a unique inverse function. In particular, if $f : V \to W$ is a bijective linear mapping, then there is a unique function $f^{-1} : W \to V$ such that:

$$f^{-1}\Big(f\big(\mathbf{v}\big)\Big) = \mathbf{v} \text{ for all } \mathbf{v} \in V,$$
$$f\Big(f^{-1}\big(\mathbf{w}\big)\Big) = \mathbf{w} \text{ for all } \mathbf{w} \in W.$$

It turns out that the inverse of a linear mapping is also a linear mapping.

Theorem 3.7

For each bijective linear mapping Let $f : V \to W$, its inverse mapping $f^{-1} : W \to V$ is also a linear mapping, i.e.,

$$f^{-1}\big(\alpha\mathbf{x} + \beta\mathbf{y}\big) = \alpha f^{-1}\big(\mathbf{x}\big) + \beta f^{-1}\big(\mathbf{y}\big) \ \forall \mathbf{x}, \mathbf{y} \in W \text{ and } \forall \alpha, \beta \in \mathbb{R}.$$

3.1.3 Associated matrices

Here, we will see that linear mappings between finite-dimensional vector spaces have a matrix representation. In fact, given a linear mapping $T : V \to W$, this matrix representation or *associated matrix*, denoted by $M_{\mathcal{C}}^{\mathcal{B}}(T)$ (or just $M_{\mathcal{C}}^{\mathcal{B}}$ for the sake of simplicity) has the useful property that

$$[T(x)]_{\mathcal{C}} = M_{\mathcal{C}}^{\mathcal{B}}[x]_{\mathcal{B}}, \tag{3.1}$$

where $\mathcal{B}$ and $\mathcal{C}$ are bases for V and W, respectively. In other words, for each $\mathbf{x} \in V$, the matrix $M_{\mathcal{C}}^{\mathcal{B}}$ should transform the coordinates of $\mathbf{x}$ into the coordinates of its image under T. In this way, the effect that $M_{\mathcal{C}}^{\mathcal{B}}$ has on $[\mathbf{x}]_{\mathcal{B}}$ is analogous to the effect that T has on $\mathbf{x}$.

An important observation is that $M_{\mathcal{C}}^{\mathcal{B}}$ depends on the bases that are chosen for the vector spaces and, of course, the mapping itself. We now discuss how to construct the associated matrix of a linear mapping relative to the bases $\mathcal{B}$ and $\mathcal{C}$.

Construction of the matrix associated to a linear mapping

Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ be a basis for V. Then, for each $\mathbf{x} \in V$, we can write

$$\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n,$$

which means that its vector of coordinates relative to $\mathcal{B}$ is

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

If we apply the linear mapping T to both sides of $\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n$, then we obtain

$$T(\mathbf{x}) = T(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n),$$

and, due to the linearity of T, we have

$$T(\mathbf{x}) = x_1T(\mathbf{b}_1) + x_2T(\mathbf{b}_2) + \cdots + x_nT(\mathbf{b}_n).$$

Since taking coordinates preserves linear combinations, we deduce that the coordinates of $T(\mathbf{x})$ relative to the basis $\mathcal{C}$ are given by

$$\begin{aligned}[T(\mathbf{x})]_{\mathcal{C}} &= \left[x_1 T(\mathbf{b}_1) + x_2 T(\mathbf{b}_2) + \cdots + x_n T(\mathbf{b}_n)\right]_{\mathcal{C}} \\ &= x_1\left[T(\mathbf{b}_1)\right]_{\mathcal{C}} + x_2\left[T(\mathbf{b}_2)\right]_{\mathcal{C}} + \cdots + x_2\left[T(\mathbf{b}_n)\right]_{\mathcal{C}}.\end{aligned}$$

or, in matrix notation,

$$[T(x)]_{\mathcal{C}} = \left[\left[T(\mathbf{b}_1)\right]_{\mathcal{C}} \quad \left[T(\mathbf{b}_2)\right]_{\mathcal{C}} \quad \ldots \quad \left[T(\mathbf{b}_n)\right]_{\mathcal{C}}\right][x]_{\mathcal{B}}, \tag{3.2}$$

where each $\left[T(\mathbf{b}_i)\right]_{\mathcal{C}}$ is an $m \times 1$ (column) vector.

Comparing (3.1) and (3.2), we see that the $m \times n$ matrix

$$M_{\mathcal{C}}^{\mathcal{B}} = \left[\left[T(\mathbf{b}_1)\right]_{\mathcal{C}} \quad \left[T(\mathbf{b}_2)\right]_{\mathcal{C}} \quad \ldots \quad \left[T(\mathbf{b}_n)\right]_{\mathcal{C}}\right], \tag{3.3}$$

is precisely the associated matrix for T that we were looking for.

Example 3.7

Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear mapping such that

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\2\\3\end{bmatrix},\ T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}2\\1\\0\end{bmatrix},$$

and let $\mathcal{E}_2$ and $\mathcal{E}_3$ be the canonical bases for $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Find $M_{\mathcal{E}_3}^{\mathcal{E}_2}(T) \equiv M_{\mathcal{E}_3}^{\mathcal{E}_2}$. Then, compute $T(\mathbf{x})$ for

$$\mathbf{x} = \begin{bmatrix}1\\1\end{bmatrix}.$$

According (3.3), the columns of $M_{\mathcal{E}_3}^{\mathcal{E}_2}$ are

$$\left[T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)\right]_{\mathcal{E}_3} = \begin{bmatrix}1\\2\\3\end{bmatrix},\ \text{and}\ \left[T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)\right]_{\mathcal{E}_3} = \begin{bmatrix}2\\1\\0\end{bmatrix}.$$

Therefore,

$$M_{\mathcal{E}_3}^{\mathcal{E}_2} = \left[\left[T\left(\begin{bmatrix}1\\0\end{bmatrix}\right)\right]_{\mathcal{E}_3} \left[T\left(\begin{bmatrix}0\\1\end{bmatrix}\right)\right]_{\mathcal{E}_3}\right] = \begin{bmatrix}1 & 2\\2 & 1\\3 & 0\end{bmatrix}$$

In this case, $T(\mathbf{x})$ can be computed directly (even if a definition for T is not explicitly given) using $M_{\mathcal{E}_3}^{\mathcal{E}_2}$, because $T(\mathbf{x}) = [T(\mathbf{x})]_{\mathcal{E}_3}$ and $\mathbf{x} = [\mathbf{x}]_{\mathcal{E}_2}$.

Therefore,

$$T(\mathbf{x})]_{\mathcal{E}_3} = M_{\mathcal{E}_3}^{\mathcal{E}_2}[\mathbf{x}]_{\mathcal{E}_2},$$

becomes

$$T\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = \begin{bmatrix}1 & 2\\ 2 & 1\\ 3 & 0\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}3\\3\\3\end{bmatrix}.$$

Example 3.8

Let $D:\mathbf{P}_3 \to \mathbf{P}_3$ be the linear mapping defined as

$$D\left(p(x)\right) = p'(x).$$

Find $M_{\mathcal{S}_3}^{\mathcal{S}_3}(D) \equiv M_{\mathcal{S}_3}^{\mathcal{S}_3}$, where $\mathcal{S}_3 = \left\{1, x, x^2, x^3\right\}$ is the standard basis for $\mathbf{P}_3$.

Note that, in this example, D is a mapping from $\mathbf{P}_3$ to itself. Moreover, we are asked to use the same basis to represent the elements in $\mathbf{P}_3$ as well as their images under D. By (3.3), we have

$$M_{\mathcal{S}_3}^{\mathcal{S}_3} = \left[[D(1)]_{\mathcal{S}_3} \quad [D(x)]_{\mathcal{S}_3} \quad \left[D\left(x^2\right)\right]_{\mathcal{S}_3} \quad \left[D\left(x^3\right)\right]_{\mathcal{S}_3}\right] = \left[[0]_{\mathcal{S}_3} \quad [1]_{\mathcal{S}_3} \quad [2x]_{\mathcal{S}_3} \quad \left[3x^2\right]_{\mathcal{S}_3}\right].$$

The explicit expression of $M_{\mathcal{S}_3}^{\mathcal{S}_3}$ is obtained by writing out the derivatives in the last equality as linear combinations of $\mathcal{S}_3$:

$$\begin{aligned}0 &= 0\cdot 1 + 0\cdot x + 0\cdot x^2 + 0\cdot x^3,\\ 1 &= 1\cdot 1 + 0\cdot x + 0\cdot x^2 + 0\cdot x^3,\\ 2x &= 0\cdot 1 + 2\cdot x + 0\cdot x^2 + 0\cdot x^3,\\ 3x^2 &= 0\cdot 1 + 0\cdot x + 3\cdot x^2 + 0\cdot x^3.\end{aligned}$$

Therefore,

$$M_{\mathcal{S}_3}^{\mathcal{S}_3} = \begin{bmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3\\ 0 & 0 & 0 & 0\end{bmatrix}.$$

Let us mention an interesting property of this linear mapping. Observe that the first column of $M_{\mathcal{S}_3}^{\mathcal{S}_3}$ consists entirely of zeros and, obviously, it is not a pivot column. Therefore, the equation

$$M_{\mathcal{S}_3}^{\mathcal{S}_3}\mathbf{x}=0,$$

has a non trivial solution (the first column corresponds to a free variable). This means that there is at least one non zero polynomial $p(x)$ such that $D(p(x))=0$. In other words, $\mathrm{Ker}D\neq\{0\}$. It follows from Theorem 3.3 that D is not injective. Indeed, it is well known that for any constant $k, D(k)=0$. Consequently, any polynomial of the form $f(x)=p(x)+k, k\in\mathbb{R}$, satisfies the equation

$$D(f(x))=p'(x).$$

Example 3.9

Let $T:\mathbf{P}_3\to\mathbf{P}_4$ be the linear mapping defined as

$$T(p(x))=xp(x).$$

Let us compute $M_{\mathcal{S}_4}^{\mathcal{B}}(T)\equiv M_{\mathcal{S}_4}^{\mathcal{B}}$ where $\mathcal{B}=\{1,1+x,x+x^2,x^2+x^3\}$ and $\mathcal{S}_4=\{1,x,x^2,x^3,x^4\}$.

Since we are using the standard basis for $\mathbf{P}_4$, it is easy to write the explicit expression for $M_{\mathcal{S}_4}^{\mathcal{B}}$. By (3.3), we have

$$\begin{aligned}M_{\mathcal{S}_4}^{\mathcal{B}}&=\begin{bmatrix}[T(1)]_{\mathcal{S}_4} & [T(1+x)]_{\mathcal{S}_4} & \left[T\left(x+x^2\right)\right]_{\mathcal{S}_4}\end{bmatrix}\begin{bmatrix}T\left(x^2+x^3\right)\end{bmatrix}_{\mathcal{S}_4}\\&=\begin{bmatrix}[x]_{\mathcal{S}_4} & \left[x+x^2\right]_{\mathcal{S}_4} & \left[x^2+x^3\right]_{\mathcal{S}_4}\end{bmatrix}\begin{bmatrix}x^3+x^4\end{bmatrix}_{\mathcal{S}_4}\\&=\begin{bmatrix}0&0&0&0\\1&1&0&0\\0&1&1&0\\0&0&1&1\\0&0&0&1\end{bmatrix}.\end{aligned}$$

The matrices of injective, surjective, and bijective linear mappings

It is possible to characterize the injective and surjective character of a linear mapping $T:V\to W$ in terms of the pivots of its associated matrix independently of the bases we choose for the vector spaces involved.

In the discussion below, $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ will denote a basis for V and $\mathcal{C}$ will denote a basis for W.

Injectivity. By Theorem 3.4, we know that T is an injective linear mapping if and only if $\{T(\mathbf{b}_1), T(\mathbf{b}_2), \ldots, T(\mathbf{b}_n)\}$ is a linearly independent set of vectors. Moreover recall that their coordinate vectors relative to any basis constitute a linearly independent set of column vectors. Hence,

$$\left\{\left[T(\mathbf{b}_1)\right]_{\mathcal{C}}, \left[T(\mathbf{b}_2)\right]_{\mathcal{C}}, \ldots, \left[T(\mathbf{b}_n)\right]_{\mathcal{C}}\right\}$$

is a linearly independet set of vectors in $\mathbb{R}^m$ with $m = \dim W$. These vectors are the columns of the associated matrix $M_{\mathcal{C}}^{\mathcal{B}}$. Therefore, the injectivity of T passes down to its associated matrix $M_{\mathcal{C}}^{\mathcal{B}}$ as follows: T is injective if and only if all the columns of $M_{\mathcal{C}}^{\mathcal{B}}$ are linearly independent, or, equivalently, all the columns of $M_{\mathcal{C}}^{\mathcal{B}}$ are pivot columns. Consequently, the equation

$$[\mathbf{y}]_{\mathcal{C}} = M_{\mathcal{C}}^{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}, \ \mathbf{y} \in \operatorname{Im} T,$$

has a unique solution. If we compute the rank of $M_{\mathcal{C}}^{\mathcal{B}}$ associated with an injective linear mapping T, then we find that

$$\operatorname{rank} M_{\mathcal{C}}^{\mathcal{B}} = \#\text{ pivots of } M_{\mathcal{C}}^{\mathcal{B}} = \#\text{ columns of } M_{\mathcal{C}}^{\mathcal{B}} = \dim V.$$

This has the following important consequence:

If there is an injective linear mapping $T : V \to W$, then

- $n = \dim V \leq m$ because any linearly independent set in W has at most m elements.
- #columns of $M_{\mathcal{C}}^{\mathcal{B}} \leq \#$ rows of $M_{\mathcal{C}}^{\mathcal{B}}$.

Surjectivity. Again, by Theorem 3.4, we know that T is surjective if and only if the set $\{T(\mathbf{b}_1), T(\mathbf{b}_2), \ldots, T(\mathbf{b}_n)\}$ is a generating system of W. Since taking coordinates preserve linear relations, this means that the set

$$\left\{\left[T(\mathbf{b}_1)\right]c, \left[T(\mathbf{b}_2)\right]c, \ldots, \left[T(\mathbf{b}_n)\right]c\right\}$$

is a generating system for $\mathbb{R}^m$ with $m = \dim W$. These vectors are the columns of the associated matrix $M_{\mathcal{C}}^{\mathcal{B}}$, and therefore surjectivity translates into the statement that all rows of $M_{\mathcal{C}}^{\mathcal{B}}$ have a pivot (in order to avoid degenerate rows), which is the same as confirming that the equation

$$[\mathrm{y}]_C = M_{\mathcal{C}}^{\mathcal{B}}\,[\mathrm{x}]_{B,} \quad \mathrm{y} \in \operatorname{Im} T$$

has a solution for all $\mathbf{b} \in W$. Therefore, computing the rank of $M_{\mathcal{C}}^{\mathcal{B}}$, we obtain

$$\text{rank}M_{\mathcal{C}}^{\mathcal{B}} = \#\text{ pivots of } M_{\mathcal{C}}^{\mathcal{B}} = \#\text{ rows of } M_{\mathcal{C}}^{\mathcal{B}} = \dim W.$$

This has the following important consequence:

If there is a surjective linear mapping $T: V \to W$, then

- $n = \dim V \geqslant m$ because any generatig set for W has at least m elements.
- #columns of $M_{\mathcal{C}}^{\mathcal{B}} \geq$ #rows of $M_{\mathcal{C}}^{\mathcal{B}}$.

Bijectivity. From the previous discussion, we deduce that T is bijective (T is both injective and surjective) if and only if the set of vectors $\{T(\mathbf{b}_1), T(\mathbf{b}_2), \ldots, T(\mathbf{b}_n)\}$ forms a basis for W. Putting what we know about injectivity and surjectivity together

$$\text{rank}M_{\mathcal{C}}^{\mathcal{B}} = \#\text{ pivots of } M_{\mathcal{C}}^{\mathcal{B}} = \#\text{ columns of } M_{\mathcal{C}}^{\mathcal{B}} = \#\text{ rows of } M_{\mathcal{C}}^{\mathcal{B}}$$

This has the following significant implication:

If there is a bijective linear mapping $T: V \to W$, then

- $\dim V = \dim W$.
- $M_{\mathcal{C}}^{\mathcal{B}}$ is an invertible square matrix.

Let us consider a few examples.

Example 3.10

Let $T: \mathbb{R}^2 \to \mathbb{R}^3$ be a linear mapping such that

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\2\\3\end{bmatrix}, \; T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}2\\1\\0\end{bmatrix}.$$

We already found the associated matrix with respect to the canonical bases:

$$M_{\mathcal{E}_3}^{\mathcal{E}_2} = \begin{bmatrix}1 & 2\\2 & 1\\3 & 0\end{bmatrix}.$$

Its row reduced echelon form is given by

$$M_{\mathcal{E}_3}^{\mathcal{E}_2} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

We see that all columns of $M_{\mathcal{E}_3}^{\mathcal{E}_2}$ are pivot columns and therefore T is injective. We can also see that $\dim\mathbb{R}^2 < \dim\mathbb{R}^3$, and hence T cannot be surjective (and, thus, not bijective).

Example 3.11

Let $D : \mathbf{P}_3 \to \mathbf{P}_3$ a linear mapping defined by $D\left(p\left(x\right)\right) = p'\left(x\right)$. We already know the associated matrix relative to the standar basis $\mathcal{S}_3 = \left\{1, x, x^2, x^3\right\}$:

$$M_{\mathcal{S}_3}^{\mathcal{S}_3} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

This matrix has less pivot columns than columns (hence, D is not injective) and less pivot rows than rows (hence, D not surjective).

Example 3.12

Let $T : \mathbf{P}_3 \to \mathbf{P}_4$ a linear mapping defined by $T\left(p\left(x\right)\right) = xp\left(x\right)$. We already know the associated matrix relative to the basis $\mathcal{B} = \left\{1, 1+x, x+x^2, x^2+x^3\right\}$ for $\mathbf{P}_3$, and the standard basis $\mathcal{S}_4 = \left\{1, x, x^2, x^3, x^4\right\}$:

$$M_{\mathcal{S}_4}^{\mathcal{B}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

After computing its row reduced echelon form:

$$M_{\mathcal{S}_4}^{\mathcal{B}} \sim \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

we find that all its columns are pivot columns, which means that T is injective. However, $\dim \mathbf{P}_3 < \dim \mathbf{P}_4$, so T cannot be surjective.

3.1.4 Associated matrices and the change of basis

Whenever we compute a matrix associated with a linear mapping $T : V \to W$, we explicitly use a fixed basis for each vector space V and W. This means that the explicit expression of the associated matrix depends on the specific choice of these bases. Nevertheless, since the underlying mapping T is independent of the choice of bases, we should expect to find a relationship between associated matrices relative to two different choices of bases. In this section, we will investigate this in detail and describe how an associated matrix varies under a change of bases. Let us start with an illustrative example.

Example 3.13

Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear mapping defined by

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x + y \\ y - z \end{bmatrix}$$

Let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ and $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2\}$ be bases for $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively. We will consider two cases for illustration: when $\mathcal{B}$ and $\mathcal{C}$ are the canonical bases and when they are not, and give the associated matrices for both cases.

Case 1: $\mathcal{B}$ and $\mathcal{C}$ are the canonical bases $\mathcal{E}_3$ and $\mathcal{E}_2$, respectively. The presentation may appear unnecessarily detailed here, but it is illustrative to see the differences with the more complicated case 2 below. We have

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{b}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{b}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

We calculate, by prescription, $T(\mathbf{b}_i)$, for $i = 1,2,3$:

$$T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}1\\0\end{bmatrix},\ T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\1\end{bmatrix},\ T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\-1\end{bmatrix}.$$

Actually, the resulting vectors are already written in terms the canonical basis $\mathcal{E}_2$, as one can readily confirm:

$$\left[\begin{bmatrix}1\\0\end{bmatrix}\right]_{\mathcal{E}_2} = \begin{bmatrix}1\\0\end{bmatrix}, \left[\begin{bmatrix}1\\1\end{bmatrix}\right]_{\mathcal{E}_2} = \begin{bmatrix}1\\1\end{bmatrix}, \left[\begin{bmatrix}0\\-1\end{bmatrix}\right]_{\mathcal{E}_2} = \begin{bmatrix}0\\-1\end{bmatrix}.$$

Therefore, we have

$$M_{\mathcal{E}_2}^{\mathcal{E}_3} = \begin{bmatrix}1 & 1 & 0\\0 & 1 & -1\end{bmatrix}.$$

Case 2: $\mathcal{B}$ and $\mathcal{C}$ are now given by

$$\mathbf{b}_1 = \begin{bmatrix}1\\0\\1\end{bmatrix},\ \mathbf{b}_2 = \begin{bmatrix}0\\1\\1\end{bmatrix},\ \mathbf{b}_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}.$$

$$\mathbf{c}_1 = \begin{bmatrix}1\\2\end{bmatrix},\ \mathbf{c}_2 = \begin{bmatrix}-1\\1\end{bmatrix}.$$

We calculate, by prescription, $T(\mathbf{b}_i)$, for $i = 1,2,3$:

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

Now, we need to compute $\left[T(\mathbf{b}_1)\right]_{\mathcal{C}}$, $\left[T(\mathbf{b}_2)\right]_{\mathcal{C}}$, and $\left[T(\mathbf{b}_3)\right]_{\mathcal{C}}$. For this, we must solve the three systems of equations whose augmented matrices are:

$$\left[\mathbf{c}_1\ \ \mathbf{c}_2\ \ T(\mathbf{b}_1)\right], \left[\mathbf{c}_1\ \ \mathbf{c}_2\ \ T(\mathbf{b}_2)\right], \left[\mathbf{c}_1\ \ \mathbf{c}_2\ \ T(\mathbf{b}_3)\right].$$

All three systems can be solved simultaneously by setting up the augmented matrix

$$\left[\mathbf{c}_1\ \ \mathbf{c}_2\ \ T(\mathbf{b}_1)\ \ T(\mathbf{b}_2)\ \ T(\mathbf{b}_3)\right],$$

and reducing it to its row reduced echelon form:

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

Therefore,

$$M_{\mathcal{C}}^{\mathcal{B}} = \begin{bmatrix} 0 & \frac{1}{3} & \frac{2}{3} \\ -1 & -\frac{2}{3} & -\frac{4}{3} \end{bmatrix}.$$

As seen in the previous example, while a linear mapping $T : V \to W$ may be independent of the choice of bases for V and W, the explicit expression for its associated matrix will change under different choices of bases.

Our goal here is to start with the associated matrix $M_{\mathcal{C}}^{\mathcal{B}}$ relative to the old bases $\mathcal{B}$ (for V) and $\mathcal{C}$ (for W), and use it to produce the associated matrix relative to new bases $\hat{\mathcal{B}}$ (for V) and $\hat{\mathcal{C}}$ (for W). Let us start by recalling the fundamental property of the associated matrix relative to the old bases $\mathcal{B}$ and $\mathcal{C}$:

$$[T(\mathbf{v})]_{\mathcal{C}} = M_{\mathcal{C}}^{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}}, \text{ for all } \mathbf{v} \in V.$$

Moreover, we can use matrices of change of bases to write the coordinates of the vectors $\mathbf{v} \in V$ and $T(\mathbf{v}) \in W$ relative to the new bases:

$$[\mathbf{v}]_{\mathcal{B}} = P_{\mathcal{B}}^{\hat{\mathcal{B}}}[\mathbf{v}]_{\hat{\mathcal{B}}}, \text{ and } [T(\mathbf{v})]_{\mathcal{C}} = P_{\mathcal{C}}^{\hat{\mathcal{C}}}[T(\mathbf{v})]_{\hat{\mathcal{C}}}.$$

If we insert these two identities in the previous equation, then we obtain

$$P_{\mathcal{C}}^{\hat{\mathcal{C}}}[T(\mathbf{v})]_{\hat{\mathcal{C}}} = M_{\mathcal{C}}^{\mathcal{B}} P_{\mathcal{B}}^{\hat{\mathcal{B}}}[\mathbf{v}]_{\hat{\mathcal{B}}}.$$

Multiplying both sides of the equation on the left by $\left(P_{\mathcal{C}}^{\hat{\mathcal{C}}}\right)^{-1} = P_{\hat{\mathcal{C}}}^{\mathcal{C}}$, we get

$$[T(\mathbf{v})]_{\hat{\mathcal{C}}} = \left[P_{\hat{\mathcal{C}}}^{\mathcal{C}} M_{\mathcal{C}}^{\mathcal{B}} P_{\mathcal{B}}^{\hat{\mathcal{B}}}\right][\mathbf{v}]_{\hat{\mathcal{B}}}.$$

Comparing this equation with the fundamental property satisfied by $M_{\hat{\mathcal{C}}}^{\hat{\mathcal{B}}}$:

$$[T(\mathbf{v})]_{\hat{\mathcal{C}}} = M_{\hat{\mathcal{C}}}^{\hat{\mathcal{B}}}[\mathbf{v}]_{\hat{\mathcal{B}}} \text{ for all } \mathrm{v} \in V,$$

we can finally conclude that

$$M_{\hat{\mathcal{C}}}^{\hat{\mathcal{B}}} = P_{\hat{\mathcal{C}}}^{\mathcal{C}} M_{\mathcal{C}}^{\mathcal{B}} P_{\mathcal{B}}^{\hat{\mathcal{B}}}.$$

Hence, we can obtain the associated matrix for T relative to the new bases $\hat{\mathcal{B}}$ and $\hat{\mathcal{C}}$ through a matrix product of three matrices. Note that, on one hand, $M_{\mathcal{C}}^{\mathcal{B}}$ and $M_{\hat{\mathcal{C}}}^{\mathcal{B}}$ are both $m \times n$ matrices, where $n = \dim V$ and $m = \dim W$. On the other hand, the matrix $P_{\mathcal{C}}^{\mathcal{C}}$ is $m \times m$ and the matrix $P_{\mathcal{B}}^{\hat{\mathcal{B}}}$ is $n \times n$.

We state the above important result as a theorem.

Theorem 3.8

Let $T : V \to W$ be a linear mapping and let $M_{\mathcal{C}}^{\mathcal{B}}$ be its associated matrix relative to the basis $\mathcal{B}$ for V and the basis $\mathcal{C}$ for W. If $\hat{\mathcal{B}}$ and $\hat{\mathcal{C}}$ are new bases for V and W, respectively, then the associate matrix for T relative to the new bases is given by

$$M_{\hat{\mathcal{C}}}^{\hat{\mathcal{B}}} = P_{\hat{\mathcal{C}}}^{\mathcal{C}} M_{\mathcal{C}}^{\mathcal{B}} P_{\mathcal{B}}^{\hat{\mathcal{B}}},$$

where $P_{\mathcal{B}}^{\hat{\mathcal{B}}}$ and $P_{\hat{\mathcal{C}}}^{\mathcal{C}}$ are matrices of change of basis.

Example 3.14

We review the example from above. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear mapping such that

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\2\\3\end{bmatrix},\ T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}2\\1\\0\end{bmatrix}.$$

Its associated matrix relative to the canonical bases is

$$M_{\mathcal{E}_3}^{\mathcal{E}_2} = \begin{bmatrix}1 & 2\\2 & 1\\3 & 0\end{bmatrix}.$$

We wish to determine $M_{\mathcal{C}}^{\mathcal{B}}$ relative to the new bases given by

$$\mathcal{B} = \left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}\right\} \text{ and } \mathcal{C} = \left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\}.$$

By Theorem 3.8, we can obtain $M_{\mathcal{C}}^{\mathcal{B}}$ as follows:

$$M_{\mathcal{C}}^{\mathcal{B}} = P_{\mathcal{C}}^{\mathcal{E}_3} M_{\mathcal{E}_3}^{\mathcal{E}_2} P_{\mathcal{E}_2}^{\mathcal{B}}.$$

Writing the matrix of change of basis from $\mathcal{B}$ and $\mathcal{C}$ to the corresponding canonical bases is straightforward since their columns consist of the basis vectors themselves:

$$P_{\mathcal{E}_2}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \; P_{\mathcal{E}_3}^{\mathcal{C}} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

However, we need to find $P_{\mathcal{C}}^{\mathcal{E}_3} = \left(P_{\mathcal{E}_3}^{\mathcal{C}} \right)^{-1}$:

$$\left[\begin{array}{cccccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cccccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{cccccc} 1 & 0 & 0 & 1 & -1 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right].$$

Therefore,

$$P_{\mathcal{C}}^{\mathcal{E}_3} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

We can finally compute $M_{\mathcal{C}}^{\mathcal{B}}$:

$$M_{\mathcal{C}}^{\mathcal{B}} = P_{\mathcal{C}}^{\mathcal{E}_3} M_{\mathcal{E}_3}^{\mathcal{E}_2} P_{\mathcal{E}_2}^{\mathcal{B}} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ -1 & 0 \\ 3 & 3 \end{bmatrix}.$$

Example 3.15

Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear mapping defined by

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

Its associated matrix relative to the canonical bases is

$$M_{\mathcal{E}_2}^{\mathcal{E}_3} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}.$$

If we wish to determine its associated matrix relative to the bases $\mathcal{B}$ and $\mathcal{C}$ given by

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \text{ and } \mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\},$$

then we need to compute the matrices of change of bases $P_{\mathcal{E}_3}^{\mathcal{B}}$ and $P_{\mathcal{C}}^{\mathcal{E}_2}$. We can write down directly the matrices

$$P_{\mathcal{E}_2}^{\mathcal{C}} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}, P_{\mathcal{E}_3}^{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

We only need to compute $P_{\mathcal{C}}^{\mathcal{E}_2} = \left(P_{\mathcal{E}_2}^{\mathcal{C}} \right)^{-1}$:

$$P_{\mathcal{C}}^{\mathcal{E}_2} = \left(P_{\mathcal{E}_2}^{\mathcal{C}} \right)^{-1} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

By Theorem 3.8, we can obtain $M_{\mathcal{C}}^{\mathcal{B}}$ as follows:

$$M_{\mathcal{C}}^{\mathcal{B}} = P_{\mathcal{C}}^{\mathcal{E}_2} M_{\mathcal{E}_2}^{\mathcal{E}_3} P_{\mathcal{E}_3}^{\mathcal{B}} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -\frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{3} & \frac{2}{3} \\ -1 & -\frac{2}{3} & -\frac{4}{3} \end{bmatrix}.$$

3.2 Eigenvalues, eigenvectors, and diagonalization

In this section, we consider a topic that has a wide range of applications in applied mathematics and by extension in the sciences and engineering: the eigenvectors and eigenvalues of a linear mapping. In the presence of some favourable conditions (namely, when the mapping is *diagonalizable*), they offer a way to understand the structure of a linear mapping that map a vector space to itself and, in turn, understand the structure of the involved vector space. Moreover, the study of diagonalizable linear mappings opens the avenue to more advanced methods (e.g., orthogonal or Fourier decomposition of vectors) that we discuss in the next chapter.

3.2.1 Introduction

For a finite-dimensional vector space V with $\dim V = n$, let $T : V \to V$ be a linear mapping, and let $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ be a basis for V. Since maps V into itself then, for each $\mathbf{x}$ in V, the vector $T(\mathbf{x})$ can also be written in the same basis $\mathcal{B}$. The matrix associated with T relative to $\mathcal{B}$ is given by

$$M_{\mathcal{B}}^{\mathcal{B}} = \left[\left[T(\mathbf{b}_1) \right]_{\mathcal{B}} \quad \left[T(\mathbf{b}_2) \right]_{\mathcal{B}} \quad \cdots \quad \left[T(\mathbf{b}_n) \right]_{\mathcal{B}} \right].$$

For the sake of simplicity, the associated matrix $M_{\mathcal{B}}^{\mathcal{B}}$ can simply be denoted as $M^{\mathcal{B}}$. We must mention that any linear mapping that maps a vector space to itself is usually called a *linear operator*. So, from now on, $M^{\mathcal{B}}$ can be understood as the associated matrix of a linear operator relative to a basis $\mathcal{B}$.

In the previous section, we discussed the behavior of the associated matrices of a linear mapping $T : V \to W$ under a change of bases. In the particular case when T is a linear operator (that is, $W = V$), we can consider the change of basis from $\mathcal{B}$ to some other basis $\mathcal{C}$. In this case, by Theorem 3.8, we have

$$M^{\mathcal{C}} = P_{\mathcal{C}}^{\mathcal{B}} M^{\mathcal{B}} P_{\mathcal{B}}^{\mathcal{C}},$$

where $P_{\mathcal{B}}^{\mathcal{C}}$ is the matrix of change of basis from $\mathcal{C}$ to $\mathcal{B}$ and $P_{\mathcal{C}}^{\mathcal{B}} = \left(P_{\mathcal{B}}^{\mathcal{C}} \right)^{-1}$.

The fundamental goal is to find a basis $\mathcal{C}$ (when possible) such that $M^{\mathcal{C}}$ is a diagonal matrix

$$M^{\mathcal{C}} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

where the empty spaces are understood to be filled with 0 's. If such a basis $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n\}$

exists, then the explicit expression

$$M^{\mathcal{C}} = \left[[T(\mathbf{c}_1)]_{\mathcal{C}} \quad [T(\mathbf{c}_2)]_{\mathcal{C}} \quad \cdots \quad [T(\mathbf{c}_n)]_{\mathcal{C}} \right] = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

shows that

$$T(\mathbf{c}_i) = \lambda_i \mathbf{c}_i, \quad \text{for } i = 1,2,\ldots,n.$$

This means that if the linear mapping is applied to a vector of this particular basis, then we obtain a multiple of the same basis vector! This is quite extraordinary since, in general, the application of a linear mapping to an input vector ***does not*** produce an output vector that is proportional to the input vector. When a non zero vector is mapped by T to a scalar multiple of itself, such a vector is called an ***eigenvector***, which comes from German "eigen," meaning "self." The scalar constant that multiplies and eigenvector is called an ***eigenvalue***. Each eigenvector is associated with an eigenvalue. We elaborate on this in the next section.

3.2.2 Eigenvalues and eigenvectors

Definition 3.5

Let $T : V \to V$ be a linear mapping. Any **non zero** vector $\mathbf{x} \in V$ such that

$$T(\mathbf{x}) = \lambda \mathbf{x},$$

for some scalar λ is called an eigenvector of T. The scalar λ is called the eigenvalue associated with $\mathbf{x}$. Eigenvalues are allowed to be zero.

Observation 3.1

The zero vector is excluded as eigenvector, because it would satisfy the equation for *any* value of λ. From the above introductory comments, we would expect that there are exactly $n = \dim V$ eigenvectors and eigenvalues. While sometimes this can be indeed the case, however, this topic will be treated with care and detail below.

Example 3.16

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be the linear mapping defined by

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

Recall that $\mathbb{C}^2 = \left\{ \begin{bmatrix} x & y \end{bmatrix}^\top : x, y \in \mathbb{C} \right\}$. Verify that $\lambda_1 = i$ and $\lambda_2 = -i$ are eigenvalues of T. If possible, find a basis for $\mathbb{C}^2$ such that its elements are eigenvectors of T and write its associated matrix relative to such basis.

We consider first $\lambda_1 = i$. If λ_i is an eigenvalue of T, then there is some vector $[xy]^\top \in \mathbb{C}^2$

such that

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = i\begin{bmatrix} x \\ y \end{bmatrix}.$$

However, by the definition of T, the same vector must satisfy

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

This implies that

$$ix = -y,$$
$$iy = x.$$

Solving this system of equations, we obtain

$$\begin{bmatrix} x \\ y \end{bmatrix} = \alpha \begin{bmatrix} i \\ 1 \end{bmatrix}, \ \alpha \in \mathbb{R}.$$

This means that all **non zero** scalar multiples of $\begin{bmatrix} i \\ 1 \end{bmatrix}$ are eigenvectors of T associated with λ_1.

Now, we consider $\lambda_2 = -i$. Then, the following two equations must hold for some $\begin{bmatrix} x & y \end{bmatrix}^\top \in \mathbb{C}^2$:

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = -i\begin{bmatrix} x \\ y \end{bmatrix}, \text{ and } T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} -y \\ x \end{bmatrix}.$$

This implies that

$$-ix = -y,$$
$$-iy = x.$$

The solution of this system of equation is given by

$$\begin{bmatrix} x \\ y \end{bmatrix} = \beta \begin{bmatrix} -i \\ 1 \end{bmatrix}, \ \beta \in \mathbb{R}.$$

This means that all **non zero** scalar multiples of $\begin{bmatrix} -i \\ 1 \end{bmatrix}$ are eigenvectors of T associated with λ_2. We observe that the set

$$\mathcal{C} = \left\{ \begin{bmatrix} i \\ 1 \end{bmatrix}, \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\},$$

formed by one eigenvector associated with λ_1 and one eigenvector associated with λ_2, is a linearly independent set. Moreover, since $\dim \mathbb{C}^2 = 2$, we conclude that $\mathcal{C}$ is a basis for $\mathbb{C}^2$.

It is not hard to verify that the associated matrix of T relative to $\mathcal{C}$ is

$$M^{\mathcal{C}} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

Note that the order that you choose for the elements of the basis is important. In fact, if we change the order of the elements of $\mathcal{C}$, then we obtain a new basis

$$\hat{\mathcal{C}} = \left\{ \begin{bmatrix} -i \\ 1 \end{bmatrix}, \begin{bmatrix} i \\ 1 \end{bmatrix} \right\},$$

and the associate matrix of T relative to $\hat{\mathcal{C}}$ is

$$M^{\hat{\mathcal{C}}} = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \neq M^{\mathcal{C}}.$$

Observation 3.2

The attentive reader would have noticed that in the previous example, the eigenvectors associated with each eigenvalues are not unique: any non zero scalar multiple of the set $\mathcal{C}$ is also an eigevector of T. This means that it is possible to normalize eigenvectors depending on the context, e.g., to obtain an eigenvector of unit length. It also implies that it is allowed to multiply an eigenvector by -1, i.e., changing the sign of all coordinate values at once, if convenient. In fact, we have the following general result.

Theorem 3.9

Let $T : V \to V$ be a linear operator. If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r \in V$ are eigenvectors associated with an eigenvalue λ, then any non zero vector $\mathbf{v} \in \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is also an eigenvector associated with λ.

Proof. Choose any non zero vector $\mathbf{v} \in \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$. Then, $\mathbf{v}$ can be written as a linear combination of the eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$. So, for some scalars $a_1, a_2, \ldots, a_r$, we can write

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r.$$

By the linearity of T and the fact that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$, are eigenvectors associated with the egienvalue λ, we have that

$$\begin{aligned} T(\mathbf{v}) &= T(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r) \\ &= a_1T(\mathbf{v}_1) + a_2T(\mathbf{v}_2) + \cdots + a_rT(\mathbf{v}_r) \\ &= a_1\lambda\mathbf{v}_1 + a_2\lambda\mathbf{v}_2 + \cdots + a_r\lambda\mathbf{v}_r \\ &= \lambda(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r) \\ &= \lambda\mathbf{v}. \end{aligned}$$

Hence, $T(\mathbf{v}) = \lambda\mathbf{v}$, which proves that $\mathbf{v}$ is an eigenvector of T associated with λ. ■

The following important theorem will be needed later.

Theorem 3.10

Let $T : V \to V$ be a linear operator and let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m$ be eigenvectors associated with ***distinct*** eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$, i.e., $\lambda_i \neq \lambda_j$ for $i \neq j$. Then, $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is a linearly independent set.

Proof. We prove this theorem by contradiction. Suppose that $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is a linearly dependent set. Therefore, since $\mathbf{x}_1 \neq 0$, we know that one of the vectors in the set is a linear combination of the preceeding vectors (see Theorem 2.6). Let p be the least index such that $\mathbf{x}_{p+1}$ is a linear combination of the preceeding (linearly independent) vectors. Then there exist scalars $c_1, \cdots, c_p$ such that

$$c_1\mathbf{x}_1 + \cdots + c_p\mathbf{v}_0 = \mathbf{v}_{p+1}.$$

On one hand, applying T to both sides of this equation and using the fact that $T(\mathbf{x}_k) = \lambda_k\mathbf{x}_k$ for each k, we obtain

$$\lambda_1c_1\mathbf{x}_1 + \cdots + \lambda_pc_p\mathbf{x}_p = \lambda_{p+1}\mathbf{v}_{p+1}.$$

On the other hand, if we multiply both sides by λ_{p+1}, we get

$$\lambda_{p+1}c_1\mathbf{x}_1 + \cdots + \lambda_{p+1}c_p\mathbf{x}_p = \lambda_{p+1}\mathbf{v}_{p+1}.$$

Subtracting the previous two equations, we obtain

$$\left(\lambda_1 - \lambda_{p+1}\right)c_1\mathbf{x}_1 + \cdots + \left(\lambda_p - \lambda_{p+1}\right)c_p\mathbf{x}_p = 0.$$

Since $\left\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p\right\}$ is linearly independent, the weights in the equation above are all zero. But none of the factors $\lambda_i - \lambda_{p+1}$ are zero, because the eigenvalues are distinct. Hence $c_i = 0$ for $i = 1, \ldots, p$, which means that $\mathbf{v}_{p+1} = 0$, which is impossible. Hence, $\left\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\right\}$ cannot be linearly dependent and therefore it must be linearly independent. ■

We note that the reciprocal of the previous theorem is not true in general. That is, it is quite possible for a linear operator to have several linearly independent eigenvectors associated with one eigenvalue. What the previous theorem is saying is that an eigenvector of a linear operator can not be associated with two distinct eigenvalues.

The following definition is motivated by Theorem 3.9.

Definition 3.6

Let $T : V \to V$ be a linear operator. For each eigenvalue λ of T, we define the set

$$E_\lambda = \text{Span}\left\{\mathbf{x} : T\left(\mathbf{x}\right) = \lambda\mathbf{x}\right\}.$$

It can easily be deduced from Theorem 3.9 that the set E_λ for each eigenvalue λ of T is a vector subspace of V. So we will call E_λ the *eigenspace associated with* λ. Observe that E_λ is composed of all the (non zero) linear combinations of the eigenvectors associated with λ, together with $0 \in V$ (otherwise, E_λ is not a vector subspace), which is not an eigenvector. By Theorem 3.10 we know that any eigenvector of T cannot belong to two different eigenspaces. Therefore we have that $E_{\lambda_i} \cap E_{\lambda_j} = \left\{0\right\}$ whenever $\lambda_i \neq \lambda_j$.

We are now paving the way for a central theorem for eigenvectors and eigenvalues. Let $T : V \to V$ be a linear operator. Our goal will be to characterize the eigenspaces of T. In order to do this, we search for solutions of the equation

$$T\left(\mathbf{x}\right) = \lambda\mathbf{x},$$

where λ is a scalar and $\mathbf{x}$ is a non zero vector in V. This equation can also be written as

$$\left(T - \lambda Id\right)\left(\mathbf{x}\right) = 0,$$

where Id denotes the identity operator on V defined by $Id(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$. This means that λ is an eigenvalue of T with associated eigenvector $\mathbf{x}$ if and only if the above equation has a non trivial solution. Consequently, we have the following characterization of the eigenspaces of T.

Theorem 3.11

Let $T : V \to V$ be a linear operator. Then, for each eigenvalue λ of T

$$E_\lambda = \mathrm{Ker}(T - \lambda Id).$$

Now, let us fix a basis $\mathcal{B}$ of V and let $M^{\mathcal{B}}$ be the associated matrix of T relative to $\mathcal{B}$. Then the associated matrix of the operator $T - \lambda Id$ relative to $\mathcal{B}$ is given by $M^{\mathcal{B}} - \lambda I_n$, where I_n is the identity matrix and $n = \dim V$. In this way, we can translante the above discussion about eigenspaces to associated matrices and coordinate vectors: the vector $\mathbf{x} \in V$ satisfies $(T - \lambda Id)(\mathbf{x}) = 0$ if and only if

$$\left(M^{\mathcal{B}} - \lambda I_n\right)[\mathbf{x}]_{\mathcal{B}} = 0 \in \mathbb{R}^n.$$

We immediately have the following result.

Theorem 3.12

Let $T : V \to V$ be a linear operator. Let $\mathcal{B}$ be a basis for V and let $M^{\mathcal{B}}$ be the associated matrix of T relative to $\mathcal{B}$. Then, for each eigenvalue λ of T :

(a) $\mathbf{x} \in E_\lambda$ if and only if $[\mathbf{x}]_{\mathcal{B}} \in \mathrm{Nul}\left(M^{\mathcal{B}} - \lambda I_n\right)$, with $n = \dim V$.

(b) $\dim E_\lambda = \dim \mathrm{Nul}\left(M^{\mathcal{B}} - \lambda I_n\right)$.

So far, we have described the eigenspaces of a linear operator assuming that we already know the eigenvalues. However, we still need to describe how to find the eigenvalues, so we make the following observation. The homogeneous equation $\left(M^{\mathcal{B}} - \lambda I_n\right)[\mathbf{x}]_{\mathcal{B}} = 0$ has a non trivial solution if and only if the value of λ is chosen so that the matrix $M^{\mathcal{B}} - \lambda I_n$ is not invertible. Equivalently, the values of λ must be choses so that

$$\det\left(M^{\mathcal{B}} - \lambda I_n\right) = 0.$$

It turns out that $\det\left(M^{\mathcal{B}} - \lambda I_n\right)$ is a polynomial of degree at most n and, therefore, if λ is an eigenvalue of T then it is a root of this polynomial. This means that a linear operator has at most n distinct eigenvalues.

This polynomial is important enough to deserve a special name.

Definition 3.7

Let $T:V\to V$ be a linear operator and let $M^{\mathcal{B}}$ be its associated matrix relative to a basis $\mathcal{B}$.

(a) The polynomial $p(\lambda)=\det\left(M^{\mathcal{B}}-\lambda I_n\right)$ of $\text{degree} n=\dim V$ is called the characteristic polynomial of T.

(b) $p(\lambda)=\det\left(M^{\mathcal{B}}-\lambda I_n\right)=0$ is called the characteristic equation of T.

We must remark that important fact that the characteristic polynomial of a linear operator is independent of the basis that we choose to compute the associated matrix. To show this, suppose that P is a matrix of change of basis from $\mathcal{B}$ to some new basis, and let $M=P^{-1}M^{\mathcal{B}}P$ is the associated matrix relative to the new basis. Then,

$$\det\left(M-\lambda I_n\right)=\det\left[P^{-1}M^{\mathcal{B}}P-\lambda P^{-1}I_nP\right]=\det\left[P^{-1}\left(M^{\mathcal{B}}-\lambda I_n\right)P\right].$$

By the multiplicative property of determinants $\left(\det AB=\left(\det A\right)\left(\det B\right)\right)$, we have

$$\det\left(M-\lambda I_n\right)=\left(\det P^{-1}\right)\left[\det\left(M^{\mathcal{B}}-\lambda I_n\right)\right]\left(\det P\right)=\det\left(M^{\mathcal{B}}-\lambda I_n\right),$$

where we have used the fact that $\left(\det P^{-1}\right)\left(\det P\right)=1$. This means that the characteristic polynomial (and its roots) does not change regardless of its matrix representation.

Theorem 3.13

Let $T:V\to V$ be a linear operator and $n=\dim V$. Then T has at most n linear independent eigenvectors.

Proof. Since the characteristic polynomial has degree n, then T has at most n distinct eigenvalues. It follows from Theorem 3.10 that T has at most n linearly independent eigenvectors. ■

Example 3.17

Find the eigenvalues and eigenvectors of $T:\mathbb{R}^2\to\mathbb{R}^2$ defined by

$$T(\mathbf{x})=A\mathbf{x}\ \text{ with }\ A=\begin{bmatrix}2 & 3\\ 3 & -6\end{bmatrix}.$$

(Although it is not explicitly mentioned, the expressions for $\mathbf{x}\in\mathbb{R}^2$ and T are given in terms of the canonical basis.)

We may proceed as follows.

(1) First, we find the eigenvalues. Recall that the eigenvalues of T are the roots of its characteristic polynomial given by

$$p(\lambda) = \det(A - \lambda I_2) = \det\begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} = \lambda^2 + 4\lambda - 21 = (\lambda + 7)(\lambda - 3).$$

Therefore, the eigenvalues of T are

$$\lambda_1 = -7,\ \lambda_2 = 3.$$

(2) Now, we obtain the eigenvectors associated with each eigenvalue.

(a) Let us start with $\lambda_1 = -7$. The eigenvectors associated with this eigenvalue are vectors in the eigenspace $E_{\lambda_1} = \mathrm{Ker}(T + 7Id)$. Since we have (implicitly) fixed the canonical basis, we have that $\mathbf{x} = [\mathbf{x}]_{\mathcal{E}_2}$ and $M^{\mathcal{E}_2} = A$. Moreover, by Theorem 3.12 we have that $\mathbf{x} \in E_{\lambda_1}$ if and oly if $\mathbf{x} \in \mathrm{Nul}(A + 7I_2)$; that is,

$$(A + 7I_2)\mathbf{x} = 0.$$

Noticing that

$$A + 7I_2 = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \frac{1}{3} \\ 0 & 0 \end{bmatrix},$$

we deduce that

$$\mathrm{Nul}(A + 7I_2) = \mathrm{Span}\left\{\begin{bmatrix} -\frac{1}{3} \\ 1 \end{bmatrix}\right\}.$$

Therefore, $\dim E_{\lambda_1} = 1$ and every eigenvector associated with λ_1 is of the form

$$\mathbf{x} = \mu\begin{bmatrix} -\frac{1}{3} \\ 1 \end{bmatrix}\ \mu \neq 0,$$

(eigenvectors must be non zero vectors).

(b) Similarly, for $\lambda_1 = 3$, we have

$$\mathrm{Nul}(A - 3I_2) = \mathrm{Span}\left\{\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right\}.$$

Therefore, $\dim E_{\lambda_2} = 1$ and every eigenvector associated with λ_2 is of the form

$$\mathbf{x} = \eta \begin{bmatrix} 3 \\ 1 \end{bmatrix} \quad \eta \neq 0.$$

We would like to remark that the two eigenvalues are distinct $\left(\lambda_1 \neq \lambda_2\right)$. Then, by Theorem 3.10, the set $\mathcal{B} = \left\{\mathbf{v}_1, \mathbf{v}_2\right\}$ with

$$\mathbf{v}_1 = \begin{bmatrix} -\frac{1}{3} \\ 1 \end{bmatrix}, \ \mathbf{v}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix},$$

is linearly independent since $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors associated with λ_1 and λ_2, respectively. Furthermore, $\dim \mathbb{R}^2 = 2$, so these two linearly independent eigenvectors form a basis of $\mathbb{R}^2$! It turns out that if we write the matrix representation of T relative to $\mathcal{B}$, we obtain

$$M^{\mathcal{B}} = \begin{bmatrix} -7 & 0 \\ 0 & 3 \end{bmatrix},$$

which, conveniently, is a diagonal matrix whose non zero element are precisely the eigenvalues of T.

Dxample 3.18

Here, we revisit Example 3.16 with the emphasis on showing the systematic procedure of calculating eigenvalues and eigenvectors starting from a linear operator.

Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be a linear operator defined by

$$T\left(\mathbf{x}\right) = A\mathbf{x}, \ \text{ with } A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

We have implicitly fixed the canonical basis for $\mathbb{C}^2$ to represent the operator as a matrix. Find the eigenvalues and eigenvectors of T.

(1) First, let us find the eigenvalues of T. The characteristic polynomial is given by

$$p\left(\lambda\right) = \det\left(A - \lambda I_2\right) = \det \begin{bmatrix} -\lambda & 1 \\ -1 & -\lambda \end{bmatrix} = \lambda^2 + 1.$$

The eigenvalues of T are the roots of the characteristic polynomial. In this case, we have two distinct eigenvalues given by $\lambda_1 = i$ and $\lambda_2 = -i$. Notice that these eigenvalues are complex numbers.

(2) Next, we describe the eigenvectors associated with each of the eigenvalues.

(a) For $\lambda_1 = i, E_{\lambda_1} = \mathrm{Ker}\left(T - iId\right)$. As in Example 3.17, we have fixed the canonical basis and, therefore, E_{λ_1} coincides with $\mathrm{Nul}\left(A - iI_2\right)$. Hence, we compute

$$A - iI_2 = \begin{bmatrix} -i & 1 \\ -1 & -i \end{bmatrix} \sim \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix},$$

and therefore $E_{\lambda_1} = \mathrm{Span}\left\{\begin{bmatrix} -i \\ 1 \end{bmatrix}\right\}$.

(b) Similarly, for $\lambda_2 = -i$, we have $E_{\lambda_2} = \mathrm{Nul}\left(A + iI_2\right)$. Since,

$$A + iI_2 = \begin{bmatrix} i & 1 \\ -1 & i \end{bmatrix} \sim \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix},$$

we obtain $E_{\lambda_2} = \mathrm{Span}\left\{\begin{bmatrix} i \\ 1 \end{bmatrix}\right\}$.

Now, consider the set

$$\mathcal{B} = \left\{\begin{bmatrix} -i \\ 1 \end{bmatrix}, \begin{bmatrix} i \\ 1 \end{bmatrix}\right\},$$

composed of one eigenvector associated with each eigenvalue. Since the eigenvalues are distinct, $\mathcal{B}$ is a linearly independent set in $\mathbb{C}^2$ with $\dim \mathbb{C}^2 = 2$. Therefore, $\mathcal{B}$ is a basis for $\mathbb{C}^2$. It is straightforward to verify that the matrix representation of T in this new basis is

$$M^{\mathcal{B}} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

3.2.3 Multiplicity of eigenvalues

So far, we have encountered examples of linear operators whose eigenvalues are all distinct and, consequently, it was possible to construct a basis for the involved vector space consisting entirely of eigenvectors of the linear

operators. However, if we take into account the multiplicity of the linear factors of the characteristic polynomial, then it is not the case that the eigenvalues of every linear operator are all distinct as we show in the following example.

Example 3.19

Let $T:\mathbb{R}^3 \to \mathbb{R}^3$ be a linear operator defined by

$$T(\mathbf{x}) = A\mathbf{x} \text{ with } A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$$

As usual, we have fixed the canonical basis of $\mathbb{R}^3$ and, thus, A is the matrix associated with T relative to the canonical basis.

Let us compute the characteristic polynomial of T :

$$p(\lambda) = \det(A - \lambda I_3) = \det\begin{bmatrix} 4-\lambda & -1 & 6 \\ 2 & 1-\lambda & 6 \\ 2 & -1 & 8-\lambda \end{bmatrix} = -(\lambda-2)^2(\lambda-9).$$

Notice that $p(\lambda)$ has two distinct real roots $\lambda_1 = 2$ and $\lambda_2 = 9$. However, if we take into account the multiplicity of each linear factor, then $p(\lambda)$ has three roots: one **double** root $\lambda_1 = 2$ and a **single** root $\lambda_2 = 9$. Hence, T does not have three distinct eigenvalues and, unfortunately, this means that the existence of a basis for $\mathbb{R}^3$ consisting entirely of eigenvectors of T is not guaranteed. At this point, we know that we can choose two linearly independent eigenvectors: one from E_{λ_1} and another one from E_{λ_2} because $\lambda_1 \neq \lambda_2$. But it is not clear if a third appropriate eigenvector exists or, even worse, where to find it.

However, if we are lucky enough, one of the eigenspaces could have dimension 2 and, therefore, we could obtain the missing third eigenvector from such eigenspace. In fact, this is precisely what happens in this example. You can verify that

$$E_{\lambda_1} = \text{Span}\left\{\begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}\right\} \text{ and } E_{\lambda_2} = \text{Span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}.$$

Fortunately, $\dim E_{\lambda_1} = 2$ and, therefore, we can choose two linearly independent eigenvectors associated with λ_1; for instance

$$B = \left\{\begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}\right\}.$$

Moreover, the set obtained from B by adding **any** eigenvector associated with λ_2 will be linearly independent (because $\lambda_1 \neq \lambda_2$). In total, we have three linearly independent eigenvectors that can be used to construct a basis for $\mathbb{R}^3$. For example, This means that the set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\},$$

is a basis for $\mathbb{R}^3$ consisting entirely of eigenvectors of T .

The following example shows that not every linear operator $T : V \to V$ admits enough linearly independent eigenvectors to constitute a basis for V .

Example 3.20

Fix the canonical basis of $\mathbb{R}^3$, and let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear operator defined by

$$T(\mathbf{x}) = A\mathbf{x} \text{ with } A = \begin{bmatrix} 2 & 2 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{bmatrix}.$$

Let us examine the eigenvalues and eigenvectors of T .

(1) First, we find the eigenvalues. The characteristic polynomial of T is

$$p(\lambda) = \det(A - \lambda I_3) = \det \begin{bmatrix} 2-\lambda & 2 & 3 \\ 0 & 2-\lambda & 2 \\ 0 & 0 & 2-\lambda \end{bmatrix} = (2-\lambda)^3.$$

Therefore, $\lambda = 2$ is a triple root of $p(\lambda)$ which, in turn, is the only eigenvalue of T .

(2) The corresponding eigenspace for $\lambda = 2$ is $E_2 = \text{Nul}(A - 2I_3)$. We have

$$E_2 = \text{Nul}(A - 2I_3) = \text{Nul} \begin{bmatrix} 0 & 2 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} = \text{Nul} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \text{Span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

In this case, T only admits one linearly independent eigenvector since $\dim E_2 = 1$. Therefore, it is not possible to produce a basis for $\mathbb{R}^3$ consisting entirely of eigevenctors of T .

There are two quantities that play an important role in the general analysis of the eigenspaces of a linear operator: the multiplicity of each eigenvalue as a root of the characteristic polynomial and the dimension of the corresponding eigenspaces. These quantities deserve their own terminology.

Definition 3.8

Let $T:V\to V$ be a linear operator and let $p(x)$ be its characteristic polynomial. For each distinct eigenvalue λ of T, we have

$$p(x)=(x-\lambda)^{m_\lambda}q(\lambda),$$

where $q(x)$ is a non zero polynomial such that $q(\lambda)\neq 0$, and m_λ is a positive integer.

(a) We say that m_λ is the **algebraic multiplicity** of λ.

(b) We say that $\dim E_\lambda$ is the **geometric multiplicity** of λ.

The following theorem states the relationship between the algebraic and geometric multiplicities of the eigenvalues of a linear operator.

Theorem 3.14

Let $T:V\to V$ be a linear operator. Then, for each distinct eigenvalue λ,

$$1\leq \dim E_\lambda \leq m_\lambda.$$

3.2.4 Diagonalization of linear operators

Now, we return to the original motivation of this chapter: given a linear operator $T:V\to V$, find a basis of V such that the matrix associated with T is a diagonal matrix. Let us introduce the following definition.

Definition 3.9

A linear operator $T:V\to V$ is said to be diagonalizable if there is a basis $\mathcal{C}=\{\mathbf{c}_1,\mathbf{c}_2,\ldots,\mathbf{c}_n\}$ of V made up entirely of eigenvectors of T.

If a linear operator is diagonalizable, then its associated matrix relative to the basis of eigenvector $\mathcal{C}$ is of the form

$$M^{\mathcal{C}}=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

where $T(\mathbf{c}_i) = \lambda_i \mathbf{c}_i$, for $i = 1,2,\ldots,n$ (see the introductory comments of this section). As we saw before, not every linear operator is diagonalizable.

An immediate consequence of the above definition is the following theorem.

Theorem 3.15

Let $T : V \to V$ be a linear operator and $n = \dim V$. Then T is diagonalizable if and only if it admits n linear independent eigenvectors.

The previous theorem, together with Theorem 3.10, yields the following result.

Corollary 3.1

Let $T : V \to V$ be a linear operator and $n = \dim V$. If T has n distinct eigenvalues, then T is diagonalizable.

Observe that from Definition 3.8, we have that T has n distinct eigenvalues if and only if $m_\lambda = 1$; that is, each eigenvalue of T has algebraic multiplicity equal to 1. Furthermore, by Theorem 3.14, this means that $\dim E_\lambda = 1$ for each eigenvalue λ.

In general, however, a linear operator may admit eigenvalues with algebraic multiplicity greater than 1. In this general setting, a linear operator is diagonalizable if and only if, for each eigenvalue λ with algebraic multiplicity m_λ, there are m_λ linear independent eigenvectors associated with λ. We formalize this fact in the following theorem.

Theorem 3.16

For a linear operator $T : V \to V$, let $\lambda_1, \lambda_2, \ldots, \lambda_k$, be all the distinct eigenvalues of T. Then T is diagonalizable if and only if

$$\dim E_{\lambda_1} + \dim E_{\lambda_2} + \cdots + \dim E_{\lambda_k} = \dim V$$

Equivalently, T is diagonalizable if and only if

$$\dim E_{\lambda_j} = m_{\lambda_j}, \quad \text{for } j = 1,2,\ldots,k.$$

Now, we turn our attention the the associated matrix of a diagonalizable linear operator.

Theorem 3.17

Let $T : V \to V$ be a diagonalizable linear operator, $n = \dim V$. Then there is a basis for V, $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$, formed by eigenvectors of T; that is, $T(\mathbf{c}_i) = \lambda_i \mathbf{c}_i$ for $i = 1, 2, \ldots, n$. Moreover, the matrix associated with T relative to $\mathcal{C}$ is a diagonal matrix given by

$$M^{\mathcal{C}} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

Some remarks are in order. In Theorem 3.17, the associated matrix $M^{\mathcal{C}}$ is a diagonal matrix with the eigenvalues of T in its main diagonal. If $\mathcal{B}$ is a different basis for V, then

$$M^{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{C}} M^{\mathcal{C}} \left(P_{\mathcal{B}}^{\mathcal{C}}\right)^{-1}$$

where the change of basis matrix is given by

$$P_{\mathcal{B}}^{\mathcal{C}} = \begin{bmatrix} [\mathbf{c}_1]_{\mathcal{B}} & \cdots & [\mathbf{c}_n]_{\mathcal{B}} \end{bmatrix}.$$

Notice that the columns of this change of basis matrix are the coordinates of the eigenvectors of T relative to $\mathcal{B}$. Moreover, the eigenvalues of T appear in the main diagonal of $M^{\mathcal{C}}$ in the same order as the corresponding eigenvectors in the basis $\mathcal{C}$.

Most textbooks present the topic of diagonalization of linear operators as a method for factorizing matrices (referring to the associated matrix $M^{\mathcal{B}}$) into a product PDP^{-1}, where P is an invertible matrix and D is a diagonal matrix. However, these notes are intended to approach diagonalization from an operator point of view.

Finally, observe that we have total freedom in choosing the order of the eigenvectors in the basis $\mathcal{C}$. This means that the diagonal associated matrix for T is only uniquely determined up to the order in which the eigenvectors of T appear in $\mathcal{C}$.

Example 3.21

Fix the standard basis of $\mathbb{R}^3$. Let $T:\mathbb{R}^3 \to \mathbb{R}^3$ be a linear operator defined by $T(\mathbf{x}) = A\mathbf{x}$, with

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}.$$

If possible, diagonalize the linear operator T.

(1) First, we find the eigenvalues. The characteristic equation is

$$p(\lambda) = \det(A - \lambda I_3) = \det\begin{bmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{bmatrix} = -(\lambda - 1)(\lambda + 2)^2.$$

Therefore, the eigenvalues of T are $\lambda_1 = 1$ and $\lambda_2 = -2$ with algebraic multiplicities $m_{\lambda_1} = 1$ and $m_{\lambda_2} = 2$, respectively.

(2) Now we find the eigenvectors.

(a) For $\lambda_1 = 1$, we have $m_{\lambda_1} = 1$, so we expect know that $\dim E_{\lambda_1} = 1$. Indeed,

$$E_{\lambda_1} = \text{Nul}(A - I_3) = \text{Nul}\begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} = \text{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right\}.$$

(b) For $\lambda_2 = -2$ with $m_{\lambda_2} = 2$, we have $1 \le \dim E_{\lambda_2} \le 2$. Let us compute

$$E_{\lambda_2} = \text{Nul}(A + 2I_3) = \text{Nul}\begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}.$$

Therefore, $\dim E_{\lambda_2} = 2$.

Since $\dim E_{\lambda_1} + \dim E_{\lambda_2} = \dim \mathbb{R}^3$, by Theorem 3.16, T is diagonalizable.

(3) Let us compute a diagonal associated matrix for T. For this, we must construct a basis for $\mathbb{R}^3$ consisting entirely of eigenvectors. We are free to choose the order of the eigenvectors, but we need to be consistent. Let

$$\mathcal{C} = \left\{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$$

be the basis of eigenvectors. Then, $A = PDP^{-1}$ with

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \; P = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Here, $D \equiv M^{\mathcal{C}}$ and $P \equiv P^{\mathcal{C}}_{\mathcal{E}_3}$ ($\mathcal{E}_3$ denotes the canonical basis).

Example 3.22

Fix the canonical basis of $\mathbb{R}^3$. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $T(\mathbf{x}) = A\mathbf{x}$, with

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

If possible, diagonalize the linear operator T.

(1) First, we find the eigenvalues. The characteristic equation is

$$p(\lambda) = \det(A - \lambda I_3) = \det \begin{bmatrix} 2-\lambda & 4 & 3 \\ -4 & -6-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{bmatrix} = -(\lambda - 1)(\lambda + 2)^2.$$

Therefore, the eigenvalues of T are $\lambda_1 = 1$ and $\lambda_2 = -2$ with algebraic multiplicities $m_{\lambda_1} = 1$ and $m_{\lambda_2} = 2$, respectively. Note that the characteristic polynomial is identical to the one of the previous example, and therefore the eigenvalues are the same as well.

(2) Now we find the eigenvectors. Let us start examining the eigenvectors associated with $\lambda_2 = -2$ with $m_{\lambda_2} = 2$, since there resides the only possiblility for the matrix not being diagonalizable if $\dim E_{\lambda_2} \neq m_{\lambda_2}$. We compute

$$E_{\lambda_2} = \text{Nul}(A + 2I_3) = \text{Nul} \begin{bmatrix} 4 & 4 & 3 \\ -4 & -4 & -3 \\ 3 & 3 & 3 \end{bmatrix} = \text{Nul} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \text{Span} \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}.$$

Since $\dim E_{\lambda_2} \neq m_{\lambda_2}$, it is not possible to find a basis for $\mathbb{R}^3$ consisting entirely of eigenvectors of T and, therefore, T is not diagonalizable.

4. INNER PRODUCTS AND ORTHOGONALITY

In making the definition of a vector space, we generalized the linear structure (addition and scalar multiplication) of $\mathbb{R}^2$ an $\mathbb{R}^3$. We ignored other important features, such as the notions of length and angle. These ideas are embedded in the concept we now investigate, inner products.

4.1 Inner product spaces

We begin with a definition.

Definition 4.1

Let V be a real (or complex) vector space. Suppose that to each pair of vectors $\mathbf{u}, \mathbf{v} \in V$ there is assigned a real number denoted by $\langle \mathbf{u}, \mathbf{v} \rangle$. This function is called an *inner product* on V if it satisfies the following axioms:

(I1) (Linearity in the first entry) $\langle a\mathbf{u}_1 + b\mathbf{u}_2, \mathbf{v} \rangle = a\langle \mathbf{u}_1, \mathbf{v} \rangle + b\langle \mathbf{u}_2, \mathbf{v} \rangle$.

(I2) (Symmetry) $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$.

(I3) (Positive definiteness) $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$; and $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

The vector space V with an inner product is called an *inner product space*.

Using the linearity axiom (I1) and the symmetry axiom (I2), we obtain

$$\langle \mathbf{u}, a\mathbf{v}_1 + b\mathbf{v}_2 \rangle = \langle a\mathbf{v}_1 + b\mathbf{v}_2, \mathbf{u} \rangle = a\langle \mathbf{v}_1, \mathbf{u} \rangle + b\langle \mathbf{v}_2, \mathbf{u} \rangle = a\langle \mathbf{u}, \mathbf{v}_1 \rangle + b\langle \mathbf{u}, \mathbf{v}_2 \rangle.$$

That is, the inner product function is also linear in its second entry. By induction, we obtain

$$\langle a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + \mathbf{a}_r\mathbf{u}_r, \mathbf{v} \rangle = a_1\langle \mathbf{u}_1, \mathbf{v} \rangle + a_2\langle \mathbf{u}_2, \mathbf{v} \rangle + \cdots + a_r\langle \mathbf{u}_r, \mathbf{v} \rangle$$

and

$$\langle \mathbf{u}, b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + \cdots + \mathbf{b}_s\mathbf{v}_s \rangle = b_1\langle \mathbf{u}, \mathbf{v}_1 \rangle + b_2\langle \mathbf{u}, \mathbf{v}_2 \rangle + \cdots + b_s\langle \mathbf{u}, \mathbf{v}_s \rangle.$$

Combining these two properties yields the following general formula:

$$\left\langle \sum_{i=1}^{r} a_i\mathbf{u}_i, \sum_{j=1}^{s} b_j\mathbf{v}_j \right\rangle = \sum_{i=1}^{r}\sum_{j=1}^{s} a_i b_j \langle \mathbf{u}_i, \mathbf{v}_j \rangle$$

$$= \begin{bmatrix} a_1 & a_2 & \cdots & a_r \end{bmatrix} \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{v}_1 \rangle & \langle \mathbf{u}_1, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{u}_1, \mathbf{v}_s \rangle \\ \langle \mathbf{u}_2, \mathbf{v}_1 \rangle & \langle \mathbf{u}_2, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{u}_2, \mathbf{v}_s \rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle \mathbf{u}_r, \mathbf{v}_1 \rangle & \langle \mathbf{u}_r, \mathbf{v}_2 \rangle & \cdots & \langle \mathbf{u}_r, \mathbf{v}_s \rangle \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_s \end{bmatrix}. \tag{4.1}$$

By axiom (I3), $\langle \mathbf{u}, \mathbf{u} \rangle$ is a non negative number and hence its positive real square root exists. We use the notation

$$\| \mathbf{u} \| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}.$$

This non negative real number $\| \mathbf{u} \|$ is called the *norm* of $\mathbf{u}$. The relation $\| \mathbf{u} \|^2 = \langle \mathbf{u}, \mathbf{u} \rangle$ will be used frequently.

Example 4.1

Consider the vector space $\mathbb{R}^n$. The function

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^\top \mathbf{v}$$

defines an inner product on $\mathbb{R}^n$. Notice that this is the usual *dot product* $\mathbf{u} \cdot \mathbf{v}$. The norm $\| \mathbf{u} \|$ of a vector in this space is as follows:

$$\| \mathbf{u} \| = \sqrt{\mathbf{u}^\top \mathbf{v}}$$

Although there are many ways to define an inner product on $\mathbb{R}^n$, we will often use the inner product in $\mathbb{R}^n$ defined by the dot product, unless otherwise stated or implied; we will call it the *usual inner product on* $\mathbb{R}^n$.

Example 4.2

Let V be the vector space of real continuous functions on the interval $a \le t \le b$. Then the following is an inner product on V:

$$\langle f, g \rangle = \int_a^b f(t) g(t)\, dt,$$

where $f(t)$ and $g(t)$ are any continuous functions on $[a, b]$.

Another inner product on V is:

$$\langle f, g \rangle = \int_a^b f(t)\, g(t)\, w(t)\, dt,$$

where $w(t)$ is a given continuous function which is positive on the interval $a \le t \le b$. In this case $w(t)$ is called a *weight function* for the inner product.

Example 4.3

Let V denote the vector space of $m \times n$ matrices over $\mathbb{R}$. The following is an inner product on V:

$$\langle A,B\rangle = Tr\left(B^{\top}A\right)$$

where Tr stands form the trace, the sum of the diagonal elements. If $A=\left(a_{ij}\right)$ and $B=\left(b_{ij}\right)$, then

$$\langle A,B\rangle = \mathrm{Tr}\left(B \quad A\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij}$$

the sum of the products of the corresponding entries. In particular,

$$\| A \|^2 = \langle A,A\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2$$

the sum of the squares of all the elements of A.

4.2 Gram matrices of an inner product

Let V be an inner product space and let $\mathcal{B}=\left\{\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n\right\}$ be a basis for V. Since any vectors $\mathbf{u},\mathbf{v}\in V$ can be written as a linear combination of the basis using their coordinates as weights, we can write the inner product using (4.1) as follows

$$\langle \mathbf{u},\mathbf{v}\rangle = [\mathbf{u}]_{\mathcal{B}}^{\top}\begin{bmatrix} \langle \mathbf{b}_1,\mathbf{b}_1\rangle & \langle \mathbf{b}_1,\mathbf{b}_2\rangle & \cdots & \langle \mathbf{b}_1,\mathbf{b}_n\rangle \\ \langle \mathbf{b}_2,\mathbf{b}_1\rangle & \langle \mathbf{b}_2,\mathbf{b}_2\rangle & \cdots & \langle \mathbf{b}_2,\mathbf{b}_n\rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle \mathbf{b}_n,\mathbf{b}_1\rangle & \langle \mathbf{b}_n,\mathbf{b}_2\rangle & \cdots & \langle \mathbf{b}_n,\mathbf{b}_n\rangle \end{bmatrix}[\mathbf{v}]_{\mathcal{B}}.$$

The $n \times n$ matrix

$$G^{\mathcal{B}} = \begin{bmatrix} \langle \mathbf{b}_1,\mathbf{b}_1\rangle & \langle \mathbf{b}_1,\mathbf{b}_2\rangle & \cdots & \langle \mathbf{b}_1,\mathbf{b}_n\rangle \\ \langle \mathbf{b}_2,\mathbf{b}_1\rangle & \langle \mathbf{b}_2,\mathbf{b}_2\rangle & \cdots & \langle \mathbf{b}_2,\mathbf{b}_n\rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle \mathbf{b}_n,\mathbf{b}_1\rangle & \langle \mathbf{b}_n,\mathbf{b}_2\rangle & \cdots & \langle \mathbf{b}_n,\mathbf{b}_n\rangle \end{bmatrix}$$

is called the *Gram matrix if the inner product relative to the basis* $\mathcal{B}$. Notice that by axioms (I2) and (I3), the Gram matrix must satisfy

$$\left(G^{\mathcal{B}}\right)^{\top} = G^{\mathcal{B}} \quad \text{and} \quad 0 \leq \mathbf{x}^{\top} G^{\mathcal{B}} \mathbf{x}, \quad \text{forall} \mathbf{x} \in \mathbb{R}^{n}.$$

Example 4.4

Recall that if $\mathcal{E}_n$ is the canonical basis in $\mathbb{R}^n$, then for every vector $\mathbf{u} \in \mathbb{R}^n$, we have $[\mathbf{u}]_{\mathcal{E}_n} = \mathbf{u}$. Moreover, the usual inner product is given by $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^{\top} \mathbf{v}$, which can be written as

$$\langle \mathbf{u}, \mathbf{v} \rangle = [\mathbf{u}]_{\mathcal{E}_n}^{\top} I_n [\mathbf{v}]_{\mathcal{E}_n}.$$

So we conclude that the Gram matrix of the usual inner product with respect to the canonical basis is the identity matrix I_n, that is, $G^{\mathcal{E}_n} = I_n$. This is easy to verify using the definition and is left to the reader as a straightforward exercise.

Example 4.5

Consider the usual inner product in $\mathbb{R}^3$ and the basis $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$, where

$$\mathbf{b}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{b}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Then,

$$\begin{aligned} \langle \mathbf{b}_1, \mathbf{b}_1 \rangle = 1, \quad \langle \mathbf{b}_1, \mathbf{b}_2 \rangle = 1, \quad \langle \mathbf{b}_1, \mathbf{b}_3 \rangle = 1, \\ \langle \mathbf{b}_2, \mathbf{b}_2 \rangle = 2, \quad \langle \mathbf{b}_2, \mathbf{b}_3 \rangle = 2, \quad \langle \mathbf{b}_3, \mathbf{b}_3 \rangle = 2. \end{aligned}$$

So the Gram matrix of the usual inner product relative to $\mathcal{B}$ is

$$G^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Notice that the Gram matrix of the usual inner product in $\mathbb{R}^3$ relative to the canonical basis is I_3. In fact, the Gram matrix of an inner product is, in general, different for each choice of basis, that is, the Gram matrix of an inner product depends on the basis. This is true for any inner product space.

Example 4.6

Consider the vector space $\mathbb{P}_2$ endowed with the inner product defined by

$$\langle f, g\rangle = \int_0^1 f(t)\, g(t)\, dt\,.$$

For the standard basis $\mathcal{S}_2 = \left\{1, t, t^2\right\}$, we have

$$\langle 1,1\rangle = \int_0^1 dt = 1, \quad \langle 1,t\rangle = \int_0^1 t\, dt = \frac{1}{2}, \quad \left\langle 1,t^2\right\rangle = \int_0^1 t^2 dt = \frac{1}{3}$$
$$\langle t,t\rangle = \int_0^1 t^2 dt = \frac{1}{3}, \quad \left\langle t,t^2\right\rangle = \int_0^1 t^3 dt = \frac{1}{4}, \quad \left\langle t^2,t^2\right\rangle = \int_0^1 t^4 dt = \frac{1}{5}\,.$$

Then the Gram matrix of the inner product relative to the standard basis is

$$G^{\mathcal{S}_2} = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}.$$

We can use the Gram matrix to compute the inner product between to elements of $\mathbb{P}_2$. If $p(t) = 1 - t$ and $q(t) = 1 - t + t^2$, then

$$\langle p(t), q(t)\rangle = [p(t)]_{\mathcal{S}_2}^{\top} G^{\mathcal{S}_2} [q(t)]_{\mathcal{S}_2} = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \frac{5}{12}$$

The definition of an inner product is independent of the choice of basis. Nevertheless, the Gram matrices of an inner product relative to an old basis $\mathcal{B}$ and a new basis $\mathcal{C}$ are, in general, not equal; that is, $G^{\mathcal{B}} \neq G^{\mathcal{C}}$. So it is sensible to conclude that the two Gram matrices are related in some way. We explore this relation here.

First, using the Gram matrix $G^{\mathcal{B}}$, we can write

$$\langle \mathbf{u}, \mathbf{v}\rangle = [\mathbf{u}]_{\mathcal{B}}^{\top} G^{\mathcal{B}} [\mathbf{v}]_{\mathcal{B}}.$$

Recall that for any vector $\mathbf{x} \in V$, the coordinate vector $[\mathbf{x}]_{\mathcal{B}}$ is related to the coordinate vector $[\mathbf{x}]_{\mathcal{C}}$ by means of the change of basis matrix $P_{\mathcal{B}}^{\mathcal{C}}$:

$$[\mathbf{x}]_{\mathcal{B}} = P_{\mathcal{B}}^{\mathcal{C}} [\mathbf{x}]_{\mathcal{C}}.$$

Therefore, we can use the change of basis matrix to change the coordinate vectors relative to the old basis to coordinate vectors relative to the new basis:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \left(P_{\mathcal{B}}^{\mathcal{C}} [\mathbf{u}]_{\mathcal{C}} \right)^{\top} G^{\mathcal{B}} \left(P_{\mathcal{B}}^{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}} \right) = [\mathbf{u}]_{\mathcal{C}}^{\top} \left(\left(P_{\mathcal{B}}^{\mathcal{C}} \right)^{\top} G^{\mathcal{B}} P_{\mathcal{B}}^{\mathcal{C}} \right) [\mathbf{v}]_{\mathcal{C}}.$$

Comparing the last equation with

$$\langle \mathbf{u}, \mathbf{v} \rangle = [\mathbf{u}]_{\mathcal{C}}^{\top} G^{\mathcal{C}} [\mathbf{v}]_{\mathcal{C}},$$

we conclude that

$$G^{\mathcal{C}} = \left(P_{\mathcal{B}}^{\mathcal{C}} \right)^{\top} G^{\mathcal{B}} P_{\mathcal{B}}^{\mathcal{C}}.$$

Example 4.7

Consider the usual inner product in $\mathbb{R}^3$. Its Gram matrices relative to two different bases were computed in Example 4.4 and Example 4.5. The Gram matrix relative to the canonical basis is $G^{\mathcal{E}_3} = I_3$, where as the Gram matrix relative to the basis $\mathcal{B}$ given in Example 4.5 is

$$G^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

The matrix $G^{\mathcal{B}}$ can be obtained from $G^{\mathcal{E}_3} = I_3$ by means of the change of basis matrix

$$P_{\mathcal{E}_3}^{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Indeed, we compute

$$\left(P_{\mathcal{E}_3}^{\mathcal{B}} \right)^{\top} G^{\mathcal{E}_3} P_{\mathcal{E}_3}^{\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix} = G^{\mathcal{B}}.$$

4.3 Cauchy-Schwarz inequality

The following inequality is called the *Cauchy-Schwarz inequality*; it is used in many branches of mathematics.

Theorem 4.1

For any vectors $\mathbf{u},\mathbf{v} \in V$,

$$|\langle \mathbf{u},\mathbf{v}\rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|.$$

This inequality is an equality if and only if one of $\mathbf{u},\mathbf{v}$ is a scalar multiple of the other.

Proof. Let $\mathbf{u},\mathbf{v} \in V$. If $\mathbf{v} = \mathbf{0}$, then both sides of the inequality are equal to 0 and the desired result holds. Thus we can assume that $\mathbf{v} \neq \mathbf{0}$.

By axiom (I3), for all $\lambda \in \mathbb{R}$, we get:

$$\langle \mathbf{u} - \lambda\mathbf{v}, \mathbf{u} - \lambda\mathbf{v}\rangle \geq 0.$$

Consequently,

$$\|\mathbf{u}\|^2 - 2\lambda\langle \mathbf{u},\mathbf{v}\rangle + \lambda^2 \|\mathbf{v}\|^2 \geq 0.$$

Taking $\lambda = \dfrac{\langle \mathbf{u},\mathbf{v}\rangle}{\|\mathbf{v}\|^2}$, we obtain

$$\|\mathbf{u}\|^2 - 2\frac{\langle \mathbf{u},\mathbf{v}\rangle^2}{\|\mathbf{v}\|^2} + \frac{\langle \mathbf{u},\mathbf{v}\rangle^2}{\|\mathbf{v}\|^2} \geq 0.$$

That is,

$$\|\mathbf{u}\|^2 - \frac{\langle \mathbf{u},\mathbf{v}\rangle^2}{\|\mathbf{v}\|^2} \geq 0.$$

Multiplying both sides by $\|\mathbf{v}\|^2$,

$$\|\mathbf{u}\|^2\|\mathbf{v}\|^2 - \langle \mathbf{u},\mathbf{v}\rangle^2 \geq 0s,$$

which implies

$$\langle \mathbf{u},\mathbf{v}\rangle^2 \leq \|\mathbf{u}\|^2\|\mathbf{v}\|^2.$$

Taking the positive square root of both sides, we finally get

$$|\langle \mathbf{u},\mathbf{v}\rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|.$$

Observe that if $\mathbf{u} = \lambda\mathbf{v}$, then

$$\langle \mathbf{u} - \lambda\mathbf{v}, \mathbf{u} - \lambda\mathbf{v} \rangle = 0.$$

In this case, each step of the proof is valid with $\leq$ and $\geq$ replaced with equality signs. Moreover, each step (with equality replacing inequalities) is reversible. Therefore, the Cauchy-Schwarz inequality is an equality if and only if $\mathbf{u}$ or $\mathbf{v}$ is a scalar multiple of the other. ■

Example 4.8

Consider any real numbers $a_1,\ldots,a_n,b_1,\ldots,b_n$. Then the Cauchy-Schwarz inequality applied to the usual inner product in $\mathbb{R}^n$ implies

$$\left(a_1b_1 + a_2b_2 + \cdots + a_nb_n\right)^2 \leq \left(a_1^2 + a_2^2 + \cdots + a_n^2\right)\left(b_1^2 + b_2^2 + \cdots + b_n^2\right).$$

Example 4.9

Let f and g be any real continuous functions defined on the interval $a \leq t \leq b$. Then, applying the Cauchy-Schwarz inequality to the first inner product introduced in Example 4.2, we obtain

$$\left(\int_a^b f(t)g(t)\,dt\right)^2 \leq \left(\int_a^b f(t)^2\,dt\right)\left(\int_a^b g(t)^2\,dt\right).$$

The next theorem gives basic properties of the norm; the proof of the third property requires the Cauchy-Schwarz inequality.

Theorem 4.2

Let V be an inner product space. Then the norm in V satisfies the following properties:

(N1) $\|\mathbf{v}\| \geq 0$; and $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.

(N2) $\|k\mathbf{v}\| = |k|\,\|\mathbf{v}\|$.

(N3) (Triangle inequality) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$; this inequality is an equality if and only if one of $\mathbf{u},\mathbf{v}$ is a non negative multiple of the other.

Proof. Property (N1) is a direct consequence of axiom (I3). Property (N2) is a direct consequence of the definition of norm and the linearity of the inner product in both entries:

$$\| k\mathbf{v} \| = \sqrt{\langle k\mathbf{v}, k\mathbf{v}\rangle} = \sqrt{k^2\langle \mathbf{v}, \mathbf{v}\rangle} = |k| \| \mathbf{v} \|.$$

Most of our effort is in proving property (N3). For $\mathbf{u}, \mathbf{v} \in V$,

$$\begin{aligned} \| \mathbf{u} + \mathbf{v} \|^2 &= \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}\rangle \\ &= \| \mathbf{u} \|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle + \| \mathbf{v} \|^2 \\ &\le \| \mathbf{u} \|^2 + 2|\langle \mathbf{u}, \mathbf{v}\rangle| + \| \mathbf{v} \|^2 \\ &\le \| \mathbf{u} \|^2 + 2\| \mathbf{u} \| \| \mathbf{v} \| + \| \mathbf{v} \|^2 \quad \text{(by Cauchy} - \text{Schwartz)} \\ &= (\| \mathbf{u} \| + \| \mathbf{v} \|)^2 \end{aligned}$$

Taking square roots of both sides of the inequality above gives the triangle inequality. The proof above shows that the triangle inequality is an equality if and only if $\langle \mathbf{u}, \mathbf{v}\rangle = \| \mathbf{u} \| \| \mathbf{v} \|$. If one $\mathbf{u}, \mathbf{v}$ is a non negative multiple of the other, then the above equality holds. Conversely, if the above equality holds, then the condition for the Cauchy-Schwarz inequality implies that one of $\mathbf{u}, \mathbf{v}$ must be a scalar multiple of the other. If $\mathbf{u} = \lambda\mathbf{v}$, then bye (N2), $\langle \mathbf{u}, \mathbf{v}\rangle = \| \mathbf{u} \| \| \mathbf{v} \|$ implies $\lambda = |\lambda|$, which forces the scalar in question to be non negative, as desired. ■

If $\| \mathbf{u} \| = 1$, or, equivalently, if $\langle \mathbf{u}, \mathbf{u}\rangle = 1$, then $\mathbf{u}$ is called a *unit vector* and is said to be *normalized*. Every non zero vector $\mathbf{v} \in V$ can be multiplied by the reciprocal of its norm to obtain a unit vector

$$\hat{\mathbf{v}} = \frac{\mathbf{v}}{\| \mathbf{v} \|}$$

which is a positive multiple of $\mathbf{v}$. This process is called *normalizing* $\mathbf{v}$.

The non negative real number

$$d(\mathbf{u}, \mathbf{v}) = ||\mathbf{u} - \mathbf{v}||$$

is called the *distance* between $\mathbf{u}$ and $\mathbf{v}$. For any non zero vectors $\mathbf{u}, \mathbf{v} \in V$, the angle between $\mathbf{u}$ and $\mathbf{v}$ is defined to be the angle θ such that $0 \le \theta \le \pi$ and

$$\cos\theta = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\| \mathbf{u} \| \| \mathbf{v} \|}.$$

By the Cauchy-Schwartz inequality, $-1 \le \cos\theta \le 1$ and so the angle always exists and is unique.

4.4 Orthogonality

Let V be an inner product space. The vectors $\mathbf{u},\mathbf{v} \in V$ are said to be *orthogonal* and $\mathbf{u}$ is said to be *orthogonal* to $\mathbf{v}$ if

$$\langle \mathbf{u},\mathbf{v} \rangle = 0.$$

The relation is clearly symmetric; that is, if $\mathbf{u}$ is orthogonal to $\mathbf{v}$, then $\langle \mathbf{v},\mathbf{u} \rangle = 0$ and so $\mathbf{v}$ is orthogonal to $\mathbf{u}$. We note that $\mathbf{0} \in V$ is orthogonal to every vector $\mathbf{v} \in V$ for

$$\langle \mathbf{0},\mathbf{v} \rangle = \langle 0\mathbf{v},\mathbf{v} \rangle = 0\langle \mathbf{v},\mathbf{v} \rangle = 0.$$

Conversely, if $\mathbf{u}$ is orthogonal to every $\mathbf{v} \in V$, then $\langle \mathbf{u},\mathbf{u} \rangle = 0$ and hence $\mathbf{u} = \mathbf{0}$ by axiom (I3). Observe that $\mathbf{u}$ and $\mathbf{v}$ are orthogonal if and only if $\cos\theta = 0$ where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$, and this is true if and only if $\theta = \pi/2$.

Example 4.10

Consider an arbitrary vector $\mathbf{u} = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}^\top$ in $\mathbb{R}^n$. Then a vector $\mathbf{v} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^\top$ is orthogonal to $\mathbf{u}$ if

$$\langle \mathbf{u},\mathbf{v} \rangle = a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0.$$

In other words, $\mathbf{v}$ is orthogonal to $\mathbf{u}$ if $\mathbf{v}$ satisfies a homogeneous equation whose coefficients are the entries of $\mathbf{u}$.

Example 4.11

Suppose that we want a non zero vector which is orthogonal to $\mathbf{v}_1 = \begin{bmatrix} 1 & 3 & 5 \end{bmatrix}^\top$ and $\mathbf{v}_2 = \begin{bmatrix} 0 & 1 & 4 \end{bmatrix}^\top$ in $\mathbb{R}^3$. Let $\mathbf{w} = \begin{bmatrix} x & y & z \end{bmatrix}^\top$. We want

$$0 = \langle \mathbf{v}_1,\mathbf{w} \rangle = x + 3y + 5z \quad \text{and} \quad 0 = \langle \mathbf{v}_2,\mathbf{w} \rangle = y + 4z.$$

Thus we obtain the homogeneous system

$$\begin{aligned} x + 3y + 5z &= 0, \\ y + 4z &= 0. \end{aligned}$$

with general solution

$$\mathbf{w} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} 7 \\ -4 \\ 1 \end{bmatrix}, \quad \lambda \in \mathbb{R}.$$

Then, $\mathbf{w}$ with $\lambda \neq 0$ is orthogonal to $\mathbf{v}_1$ and $\mathbf{v}_2$.

Let S be a subset of an inner product space V. The *orthogonal complement* of S, denoted by $S^{\perp}$ (read "S perp") consists of those vectors in V which are orthogonal to every vector $\mathbf{u} \in S$:

$$S^{\perp} = \{\mathbf{v} \in V : \ \mathbf{v},\mathbf{u} \ = 0 \text{ for every } \mathbf{u} \in S\}.$$

We show that $S^{\perp}$ is a subspace of V. Clearly $\mathbf{0} \in S^{\perp}$ since $\mathbf{0}$ is orthogonal to every vector in V. Now suppose that $\mathbf{u},\mathbf{w} \in S^{\perp}$. Then, for any scalars a and b and any vector $\mathbf{u} \in S$, we have

$$\langle a\mathbf{v} + b\mathbf{w},\mathbf{u}\rangle = a\langle \mathbf{v},\mathbf{u}\rangle + b\langle \mathbf{w},\mathbf{u}\rangle = a0 + b0 = 0.$$

Thus $a\mathbf{v} + b\mathbf{w} \in S^{\perp}$ and therefore $S^{\perp}$ is a subspace of V.

Example 4.12

Suppose we want to find a basis for the subspace S = Span{$\mathbf{v}$} in $\mathbb{R}^3$ where $\mathbf{u} = [1\,3 - 4]^{\top}$. Note that $S^{\perp}$ consists of all vectors $\mathbf{w}$ = $[x\ y\ z]^{\top}$ such that

$$\langle \mathbf{w},\mathbf{u}\rangle = x + 3y - 4z = 0.$$

The general solution of this homogeneous equation is

$$\mathbf{w} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}, \quad \lambda,\mu \in \mathbb{R}.$$

Therefore, a basis for $S^{\perp}$ is

$$\left\{ \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

The next theorem is a basic result in linear algebra.

Theorem 4.3

Let S = Span{$\mathbf{v}$} where $\mathbf{v} \in V$ and $\mathbf{v} \neq \mathbf{0}$. Then for every $\mathbf{x} \in V$ there is a unique $\mathbf{u} \in S$ and a unique $\mathbf{w} \in S^{\perp}$ such that

$$\mathbf{x} = \mathbf{u} + \mathbf{w}.$$

Proof. Let λ be a scalar. Then

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

Notice that $\lambda \in S$. Thus we need to choose λ so that $\mathbf{v}$ is orthogonal to $(\mathbf{x} - \lambda\mathbf{v})$. In other words, we want

$$0 = \langle \mathbf{x} - \lambda\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{x}, \mathbf{v} \rangle - \lambda \| \mathbf{v} \|^2 .$$

The equation above shows that we should choose λ to $\langle \mathbf{x}, \mathbf{v} \rangle / \| \mathbf{v} \|^2$. Making this choice of λ, we can write

$$\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\| \mathbf{v} \|^2}\mathbf{v} + \left(\mathbf{x} - \frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\| \mathbf{v} \|^2}\mathbf{v} \right).$$

■

Example 4.13

Consider the vectors $\mathbf{x} = \begin{bmatrix} 0 & 1 \end{bmatrix}^\top$ and $\mathbf{v} = \begin{bmatrix} 1 & 1 \end{bmatrix}^\top$ in $\mathbb{R}^2$ with the usual inner product. Then $\langle \mathbf{x}, \mathbf{v} \rangle = 1, \| \mathbf{v} \|^2 = 2$, and

$$\mathbf{w} = \mathbf{x} - \frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\| \mathbf{v} \|^2}\mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}.$$

Then $\mathbf{x}$ can be written uniquely as

$$\mathbf{x} = \frac{1}{2}\mathbf{v} + \mathbf{w},$$

where $\langle \mathbf{v}, \mathbf{w} \rangle = 0$.

Example 4.14

Consider the vector space $\mathbb{P}_2$ endowed with the inner product defined by

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt.$$

Relative to the standard basis $\mathcal{S}_2 = \left\{ 1, t, t^2 \right\}$, we have the Gram matrix

$$G^{\mathcal{S}_2} = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}.$$

If $p(t) = 1 - t$ and $q(t) = 1 - t + t^2$, then

$$\langle p(t), q(t) \rangle = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \frac{5}{12}$$

and

$$\langle q(t), q(t) \rangle = \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \frac{7}{10}$$

Then

$$p(t) = \frac{25}{42} q(t) + \left(p(t) - \frac{25}{42} q(t) \right) = \frac{25}{42} q(t) + \left(\frac{17}{42} - \frac{17}{42} t - \frac{25}{42} t^2 \right)$$

where

$$\left\langle q(t), \frac{17}{42} - \frac{17}{42} t - \frac{25}{42} t^2 \right\rangle = \begin{bmatrix} \frac{17}{42} & -\frac{17}{42} & -\frac{25}{42} \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 0.$$

4.5 Orthogonal sets and bases

A set S of vectors in V is called *orthogonal* if each pair of vectors in S are orthogonal, and S is called *orthonormal* if S is orthogonal and each vector in S has unit norm. In other words, $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ is orthogonal if

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0 \quad \text{for} \quad i \neq j$$

and is *orthonormal* if

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{i,j} = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}$$

Normalizing an orthogonal set S refers to the process of multiplying each vector in S by the reciprocal of its norm in order to transform S into an orthonormal set of vectors. The following theorems apply.

Theorem 4.4

Suppose that S is an orthogonal set of non zero vectors. Then S is linearly independent.

Proof. Let $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ be an orthogonal set. Suppose that the scalars $a_1, a_2, \ldots, a_r$ satisfy

$$a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_r\mathbf{u}_r = \mathbf{0}.$$

Computing the inner product of both sides of the above equation with $\mathbf{u}_i$ for $i = 1, 2, \ldots, r$, by the orthogonality of S we obtain

$$\begin{aligned} 0 &= \langle \mathbf{u}_i, \mathbf{0} \rangle = \langle \mathbf{u}_i, a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_r\mathbf{u}_r \rangle \\ &= a_1 \langle \mathbf{u}_i, \mathbf{u}_1 \rangle + \cdots + a_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle + \cdots + a_r \langle \mathbf{u}_i, \mathbf{u}_r \rangle = a_i \|\mathbf{u}_i\|^2, \quad \text{for } i = 1, 2, \ldots, r. \end{aligned}$$

This implies that $a_i = 0$ for $i = 1, 2, \ldots, r$ is the unique solution of the homogeneous equation. Therefore S is a linearly independent set. ■

Theorem 4.5: Pythagoras

Suppose that $S = \{\mathbf{u}, \mathbf{v}\}$ is an orthogonal set of two vectors. Then

$$\| \mathbf{u} + \mathbf{v} \|^2 = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 .$$

Proof. Suppose that $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. Then

$$\| \mathbf{u} + \mathbf{v} \|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle = \| \mathbf{u} \|^2 + \| \mathbf{v} \|^2$$

which gives the desired result. ■

Example 4.15

Consider the canonical basis $\mathcal{E}_3 = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ in $\mathbb{R}^3$ with the usual inner product. It is clear that

$$\langle \mathbf{e}_1, \mathbf{e}_2 \rangle = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle = \langle \mathbf{e}_2, \mathbf{e}_3 \rangle = 0 \quad \text{and} \quad \langle \mathbf{e}_1, \mathbf{e}_1 \rangle = \langle \mathbf{e}_2, \mathbf{e}_2 \rangle = \langle \mathbf{e}_3, \mathbf{e}_3 \rangle = 1.$$

Thus $\mathcal{E}_3$ is an orthonormal basis of $\mathbb{R}^3$. More generally, the canonical basis in $\mathbb{R}^n$ with the usual inner product is orthonormal for every n.

Example 4.16

Consider the vector space $\mathbb{P}_2$ endowed with the inner product defined by

$$\langle f, g \rangle = \int_0^1 f(t)\, g(t)\, dt.$$

Relative to the standard basis $\mathcal{S}_2 = \{1, t, t^2\}$, we have the Gram matrix

$$G^{\mathcal{S}_2} = \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix}.$$

Let $S = \{1, 1-2t, 2-12t+12t^2\}$ is an orthogonal set since

$$\langle 1, 1-2t \rangle = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} = 0$$

$$\left\langle 1, 2-12t+12t^2 \right\rangle = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} \begin{bmatrix} 2 \\ -12 \\ 12 \end{bmatrix} = 0$$

$$\left\langle 1-2t, 2-12t+12t^2 \right\rangle = \begin{bmatrix} 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{bmatrix} \begin{bmatrix} 2 \\ -12 \\ 12 \end{bmatrix} = 0$$

The most useful property of orthogonal bases is that the coordinates of any vector relative to an orthogonal basis can be computed directly one by one, without the need to solve a system of linear equations.

Theorem 4.6

Suppose that $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthogonal basis for V. Then, for any $\mathbf{v} \in V$,

$$\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} \mathbf{u}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{u}_n \rangle}{\langle \mathbf{u}_n, \mathbf{u}_n \rangle} \mathbf{u}_n.$$

Proof. Since $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is a basis for V, any vector $\mathbf{v} \in V$ can be written as a linear combination of the basis; that is,

$$\mathbf{v} = a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \cdots + a_n \mathbf{u}_n$$

for suitable scalars $a_1, a_2, \ldots, a_n$. Using the orthogonality of the basis, we compute

$$\begin{aligned} \langle \mathbf{v}, \mathbf{u}_i \rangle &= \langle a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \cdots + a_n \mathbf{u}_n, \mathbf{u}_i \rangle \\ &= a_1 \langle \mathbf{u}_1, \mathbf{u}_i \rangle + \cdots + a_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle + \cdots + a_n \langle \mathbf{u}_n, \mathbf{u}_i \rangle = a_i \langle \mathbf{u}_i, \mathbf{u}_i \rangle, \quad \text{for } i = 1, 2, \ldots, n. \end{aligned}$$

This implies that $a_i = \langle \mathbf{v}, \mathbf{u}_i \rangle / \langle \mathbf{u}_i, \mathbf{u}_i \rangle$ for $i = 1, 2, \ldots, n$, which is the desired result. ■

The above scalars

$$a_i = \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} = \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\|\mathbf{u}_i\|^2}, \quad \text{for } i = 1, 2, \ldots, n,$$

are called the *Fourier coefficients* of $\mathbf{v}$ with respect to $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$.

Consider a non zero vector $\mathbf{w}$ in an inner product space V. For any $\mathbf{v} \in V$, we have showed that

$$\lambda = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\| \mathbf{w} \|^2}$$

is the unique scalar such that $\hat{\mathbf{v}} = \mathbf{v} - \lambda \mathbf{w}$ is orthogonal to $\mathbf{w}$. The *projection* of $\mathbf{v}$ along $\mathbf{w}$, denoted by $proj_{\mathbf{w}} \mathbf{v}$, is defined as

$$\text{proj}_{\mathbf{w}} \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \mathbf{w} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\| \mathbf{w} \|^2} \mathbf{w}.$$

The above notion is generalized as follows.

Theorem 4.7

Suppose that $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ form an orthogonal set of non zero vectors in V. Let $\mathbf{v}$ be any vector in V. Define

$$\hat{\mathbf{v}} = \mathbf{v} - \left(c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \cdots + c_r \mathbf{u}_r \right)$$

where

$$c_1 = \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}, \quad c_2 = \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}, \quad \ldots, \quad c_r = \frac{\langle \mathbf{v}, \mathbf{u}_r \rangle}{\langle \mathbf{u}_r, \mathbf{u}_r \rangle}.$$

Then $\hat{\mathbf{v}}$ is orthogonal to $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$.

Proof. We verify that $\hat{\mathbf{v}}$ is orthogonal to $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$. For $i = 1, 2, \ldots, r$, compute using the orthogonality of $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$,

$$\begin{aligned}\langle \hat{\mathbf{v}}, \mathbf{u}_i \rangle &= \langle \mathbf{v} - (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_r\mathbf{u}_r), \mathbf{u}_i \rangle \\ &= \langle \mathbf{v}, \mathbf{u}_i \rangle - c_1\langle \mathbf{u}_1, \mathbf{u}_i \rangle - c_2\langle \mathbf{u}_2, \mathbf{u}_i \rangle - \cdots - c_i\langle \mathbf{u}_i, \mathbf{u}_i \rangle - \cdots - c_r\langle \mathbf{u}_r, \mathbf{u}_i \rangle \\ &= \langle \mathbf{v}, \mathbf{u}_i \rangle - c_i\langle \mathbf{u}_i, \mathbf{u}_i \rangle \\ &= \langle \mathbf{v}, \mathbf{u}_i \rangle - \frac{\langle \mathbf{v}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle}\langle \mathbf{u}_i, \mathbf{u}_i \rangle = \langle \mathbf{v}, \mathbf{u}_i \rangle - \langle \mathbf{v}, \mathbf{u}_i \rangle = 0.\end{aligned}$$

Therefore $\hat{\mathbf{v}}$ is orthogonal to $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$. ■

Notice that each c_i in the theorem above is the Fourier coefficient of $\mathbf{v}$ along the corresponding vector $\mathbf{u}_i$.

The notion of the projection of a vector $\mathbf{v} \in V$ along a subspace W of V is defined as follows. If $W = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ where the $\mathbf{w}_i$ form an orthogonal set, then the *projection* of $\mathbf{v}$ along W, denoted by $proj_W\mathbf{v}$, is defined as

$$\text{proj}_W\mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{w}_1 \rangle}{\langle \mathbf{w}_1, \mathbf{w}_1 \rangle}\mathbf{w}_1 + \frac{\langle \mathbf{v}, \mathbf{w}_2 \rangle}{\langle \mathbf{w}_2, \mathbf{w}_2 \rangle}\mathbf{w}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{w}_r \rangle}{\langle \mathbf{w}_r, \mathbf{w}_r \rangle}\mathbf{w}_r.$$

4.6 Gram-Schmidt orthogonalization

Suppose that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis of an inner product space V. One can use this to construct an orthogonal basis $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ of V as follows. Set

$$\begin{aligned}&\mathbf{w}_1 = \mathbf{v}_1, \\ &\mathbf{w}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{w}_1}\mathbf{v}_2, \\ &\mathbf{w}_3 = \mathbf{v}_3 - \text{proj}_{\mathbf{w}_1}\mathbf{v}_3 - \text{proj}_{\mathbf{w}_2}\mathbf{v}_3, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\mathbf{w}_n = \mathbf{v}_n - \text{proj}_{\mathbf{w}_1}\mathbf{v}_n - \text{proj}_{\mathbf{w}_2}\mathbf{v}_n - \cdots - \text{proj}_{\mathbf{w}_{n-1}}\mathbf{v}_n.\end{aligned}$$

In other words, for $k = 2, 3, \ldots, n$, we define

$$\mathbf{w}_k = \mathbf{v}_k - proj_{W_{k-1}}\mathbf{v}_k, \quad \text{where} \quad W_{k-1} = Span\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{k-1}\}.$$

By Theorem 4.7, each $\mathbf{w}_k$ is orthogonal to the preceding $\mathbf{w}$'s. Thus, $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ forms an orthogonal basis for V as claimed. Normalizing will then yield an orthonormal basis for V.

The above construction is known as the *Gram-Schmidt orthogonalization process*. Each vector $\mathbf{w}_k$ is a linear combination of $\mathbf{v}_k$ and the preceding

$\mathbf{w}$'s. Hence, it can easily be shown, by induction, that each $\mathbf{w}_k$ is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Because taking multiples of vectors does not affect orthogonality, it may be simpler in hand calculations to clear fractions in any new $\mathbf{w}_k$, by multiplying $\mathbf{w}_k$ by an appropriate scalar, before obtaining the next $\mathbf{w}_{k+1}$.

Suppose $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r$ are linearly independent, and so they form a basis for

$$U = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}.$$

Applying Gram-Schmidt orthogonalization to the u's yields an orthogonal basis for U. The following theorems use the above algorithm and remarks.

Theorem 4.8

Suppose that $S = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_r\}$ is an orthogonal basis for a subspace W of a vector space V. Then one may extend S to an orthogonal basis for V; that is, one may find vectors $\mathbf{w}_{r+1}, \ldots, \mathbf{w}_n$ such that $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ is an orthogonal basis for V.

Proof. Since S is a linearly independent set in V, we can extend S to a basis of V. That is, we can find vectors $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_n$ such that $\{\mathbf{w}_1, \ldots, \mathbf{w}_r, \mathbf{u}_{r+1}, \ldots \mathbf{u}_n\}$ is a basis for V. For $k = r+1, \ldots, n$, we define

$$\mathbf{w}_k = \mathbf{u}_k - \text{proj}_{W_{k-1}} \mathbf{u}_k, \quad \text{where} \quad W_{k-1} = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{k-1}\}.$$

This is just the Gram-Schmidt orthogonalization process applied to $\{\mathbf{w}_1, \ldots, \mathbf{w}_r, \mathbf{u}_{r+1}, \ldots \mathbf{u}_n\}$, which yields an orthogonal basis for V. ■

Theorem 4.9

Let W be a subspace of an inner product space V. Then each vector $\mathbf{v} \in V$ can be expressed uniquely in the form

$$\mathbf{v} = \mathbf{w} + \widehat{\mathbf{w}}, \quad \mathbf{w} \in W, \quad \widehat{\mathbf{w}} \in W^{\perp}$$

Proof. Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ be a basis for W. Applying the Gram-Schmidt orthogonalization process, we obtain an orthogonal basis for $W : S = \{\mathbf{w}_1, \ldots, \mathbf{w}_r\}$. Using S, for every $\mathbf{v} \in V$, we can write

$$\mathbf{v} = \text{proj}_W \mathbf{v} + (\mathbf{v} - \text{proj}_W \mathbf{v})$$

Clearly, $\text{proj}_W\mathbf{v} \in W$. By Theorem 4.7, $\left(\mathbf{v} - \text{proj}_W\mathbf{v}\right)$ is orthogonal to every vector in S and, thus, to every vector in W. Therefore, $\left(\mathbf{v} - \text{proj}_W\mathbf{v}\right) \in W^{\perp}$.

Suppose that we can write

$$\mathbf{v} = \mathbf{w} + \widehat{\mathbf{w}}, \quad \mathbf{w} \in W, \quad \widehat{\mathbf{w}} \in W^{\perp}$$

Then subtracting the above expression from $\mathbf{v} = \text{proj}_W\mathbf{v} + \left(\mathbf{v} - \text{proj}_W\mathbf{v}\right)$, we get

$$\mathbf{0} = \left(\text{proj}_W\mathbf{v} - \mathbf{w}\right) + \left(\mathbf{v} - \text{proj}_W\mathbf{v} - \widehat{\mathbf{w}}\right).$$

Observe that $\left(\text{proj}_W\mathbf{v} - \mathbf{w}\right) \in W$ and $\left(\mathbf{v} - \text{proj}_W\mathbf{v} - \widehat{\mathbf{w}}\right) \in W^{\perp}$. Therefore,

$$\begin{aligned} 0 = \left\langle \mathbf{0}, \text{proj}_W\mathbf{v} - \mathbf{w} \right\rangle &= \left\langle \left(\text{proj}_W\mathbf{v} - \mathbf{w}\right) + \left(\mathbf{v} - \text{proj}_W\mathbf{v} - \widehat{\mathbf{w}}\right), \text{proj}_W\mathbf{v} - \mathbf{w} \right\rangle \\ &= \left\langle \text{proj}_W\mathbf{v} - \mathbf{w}, \text{proj}_W\mathbf{v} - \mathbf{w} \right\rangle. \end{aligned}$$

This implies that $proj_W\mathbf{v} - \mathbf{w} = \mathbf{0}$ or, equivalently, $\mathbf{w} = proj_W\mathbf{v}$. Similarly,

$$\begin{aligned} 0 = \left\langle \mathbf{0}, \mathbf{v} - \text{proj}_W\mathbf{v} - \widehat{\mathbf{w}} \right\rangle &= \left\langle \left(\text{proj}_W\mathbf{v} - \mathbf{w}\right) + \left(\mathbf{v} - \text{proj}_W\mathbf{v} - \widehat{\mathbf{w}}\right), \mathbf{v} - \text{proj}_W\mathbf{v} - \widehat{\mathbf{w}} \right\rangle \\ &= \left\langle \mathbf{v} - \text{proj}_W\mathbf{v} - \widehat{\mathbf{w}}, \mathbf{v} - \text{proj}_W\mathbf{v} - \widehat{\mathbf{w}} \right\rangle. \end{aligned}$$

Therefore, $\mathbf{v} - \text{proj}_W\mathbf{v} - \widehat{\mathbf{w}} = \mathbf{0}$ or, equivalently, $\widehat{\mathbf{w}} = \mathbf{v} - \text{proj}_W\mathbf{v}$. This proves that each $\mathbf{v} \in V$ is expressed uniquely as $\mathbf{v} = \text{proj}_W\mathbf{v} + \left(\mathbf{v} - \text{proj}_W\mathbf{v}\right)$. ■

Example 4.17

Apply the Gram-Schmidt orthogonalization process to find an orthogonal basis for the subspace U of $\mathbb{R}^4$ spanned by

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

Therefore, we first set $\mathbf{w}_1 = \mathbf{v}_1$. Then we set

$$\mathbf{w}_2 = \mathbf{v}_2 - \text{proj}_{\mathbf{w}_1}\mathbf{v}_2 = \mathbf{v}_2 - \frac{12}{4}\mathbf{w}_1 = \begin{bmatrix} -2 \\ -1 \\ 1 \\ 2 \end{bmatrix}.$$

Finally, compute

$$\mathbf{v}_3 - \text{proj}_{\mathbf{w}_1}\mathbf{v}_3 - \text{proj}_{\mathbf{w}_2}\mathbf{v}_3 = \mathbf{v}_3 - \frac{(-8)}{4}\mathbf{w}_1 - \frac{(-7)}{10}\mathbf{w}_2 = \begin{bmatrix} \frac{8}{5} \\ -\frac{17}{10} \\ -\frac{13}{10} \\ \frac{7}{5} \end{bmatrix}.$$

Multiplying the last vector by 10 to clear fractions, we set

$$\mathbf{w}_3 = \begin{bmatrix} 16 \\ -17 \\ -13 \\ 14 \end{bmatrix}.$$

Thus, $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ form an orthogonal basis for U.

Example 4.18

Let V be the vector space of polynomials with inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(t)\, g(t)\, dt$$

Apply the Gram-Schmidt orthogonalization process to $\{1, t, t^2, t^3\}$ to find an orthogonal basis for $\mathbf{P}_3$.

Here we use the fact that, for $r + s = n$,

$$\langle t^r, t^s \rangle = \int_{-1}^{1} t^n dt = \frac{t^{n+1}}{n+1}\Big|_{-1}^{1} = \begin{cases} 2/(n+1), & \text{when } n \text{ is even} \\ 0, & \text{when } n \text{ is odd} \end{cases}$$

We compute

$$f_0(t) = 1$$
$$f_1(t) = t - proj_{f_0}\, t = t - 0 = t$$
$$f_2(t) = t^2 - proj_{f_0}\, t^2 - proj_{f_1}\, t^2 = t^2 - \frac{1}{3}$$
$$f_3(t) = t^3 - proj_{f_0}\, t^3 - proj_{f_1}\, t^3 - proj_{f_2}\, t^3 = t^3 - \frac{3}{5}t.$$

Thus, $\left\{1, t, t^2 - \frac{1}{3}, t^3 - \frac{3}{5}t\right\}$ is the required orthogonal basis.

4.7 Orthogonal projections and minimization

Suppose that U is a subspace of an inner product space V. Then each vector $\mathbf{v} \in V$ can be written uniquely in the form

$$\mathbf{v} = \mathbf{u} + \mathbf{w},$$

where $\mathbf{u} \in U$ and $\mathbf{w} \in U^{\perp}$. We use this decomposition to define a linear mapping on V, denoted by $proj_U$, called the *orthogonal projection* of V along U. For $\mathbf{v} \in V$, we define $proj_U \mathbf{v}$ to be the vector $\mathbf{u}$ in the decomposition above.

The orthogonal projection along a subspace satisfies the following properties.

Theorem 4.10

Let U is a subspace of an inner product space V. Then,

1. range $\text{proj}_U = U$
2. ker $\text{proj}_U = U^{\perp}$
3. $\mathbf{v} - \text{proj}_U \mathbf{v} \in U^{\perp}$ for every $\mathbf{v} \in V$.
4. $\text{proj}_U^2 = \text{proj}_U$
5. $\|\text{proj}_U \mathbf{v}\| \leqslant \|\mathbf{v}\|$ for every $\mathbf{v} \in V$.

Proof. 1. If $\mathbf{u} \in U$, then $\mathbf{u} = \text{proj}_U \mathbf{u}$. Therefore, $\mathbf{u} \in$ range proj_U. Conversely, if $\mathbf{u} \in$ range proj_U, then by definition, $\mathbf{u} \in U$. Altogether, we have $U \subseteq$ range proj_U and range $\text{proj}_U \subseteq U$, which implies range $\text{proj}_U = U$.

2. If $\mathbf{u} \in U^{\perp}$, then it is expressed uniquely as $\mathbf{u} = \mathbf{0} + \mathbf{u}$ since it has no component in U other than the zero vector. Then $\text{proj}_U \mathbf{u} = \mathbf{0}$ which means that $\mathbf{u} \in$ ker proj_U. Conversely, if $\mathbf{u} \in$ ker proj_U then $\text{proj}_U \mathbf{u} = \mathbf{0}$. But $\mathbf{u}$ is expressed uniquely as $\mathbf{u} = \mathbf{0} + \mathbf{w}$ where $\mathbf{w} \in U^{\perp}$. It follows that $\mathbf{u} \in U^{\perp}$. Altogether, we have $U^{\perp} \subseteq$ ker proj_U and ker $\text{proj}_U \subseteq U^{\perp}$, which implies ker $\text{proj}_U = U^{\perp}$.

3. Since $\mathbf{v}$ can be expressed as $\mathbf{v} = \mathbf{u} + \mathbf{w}$ where $\mathbf{w} \in U^{\perp}$ and, by definition, $\mathbf{u} = \text{proj}_U \mathbf{v}$. It follows that $\mathbf{w} = \mathbf{v} - \text{proj}_U \mathbf{v}$ and, thus, $\mathbf{v} - \text{proj}_U \mathbf{v} \in U^{\perp}$.

4. Consider any vector $\mathbf{v} \in V$. Then, by the first and second statements above, it can be written as $\mathbf{v} = \text{proj}_U \mathbf{v} + \mathbf{w}$ where $\mathbf{w} \in \text{ker proj}_U$. Therefore, applying the orthogonal projection to both sides of this equation, we get

$$\text{proj}_U \mathbf{v} = \text{proj}_U^2 \mathbf{v} + \text{proj}_U \mathbf{w} = \text{proj}_U^2 \mathbf{v}$$

But this is true for any $\mathbf{v} \in V$. Then, $\text{proj}_U^2 = \text{proj}_U$.

5. For every $\mathbf{v} \in V$, by Pythagoras,

$$\left\|\text{proj}_U \mathbf{v}\right\|^2 \leqslant \left\|\text{proj}_U \mathbf{v}\right\|^2 + \left\|\mathbf{v} - \text{proj}_U \mathbf{v}\right\|^2 = \| \mathbf{v} \|^2$$

Taking the positive square root of both sides, we get $\left\|\text{proj}_U \mathbf{v}\right\| \leqslant \| \mathbf{v} \|$. ■

By the definition of orthogonal projection, if $U = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ where the $\mathbf{u}_i$ form an orthogonal basis, then

$$\text{proj}_U \mathbf{v} = \frac{\langle \mathbf{v}, \mathbf{u}_1 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle} \mathbf{u}_1 + \frac{\langle \mathbf{v}, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} \mathbf{u}_2 + \cdots + \frac{\langle \mathbf{v}, \mathbf{u}_r \rangle}{\langle \mathbf{u}_r, \mathbf{u}_r \rangle} \mathbf{u}_r.$$

The following problem often arises: given a subspace U of V and a vector $\mathbf{v} \in V$, find a vector $\mathbf{u} \in U$ such that $\| \mathbf{v} - \mathbf{u} \|$ is as small as possible. The next theorem shows that this minimization problem is solved by taking $\mathbf{u} = \text{proj}_U \mathbf{v}$.

Theorem 4.11

Suppose that U is a subspace of V and $\mathbf{v} \in V$. Then

$$\left\|\mathbf{v} - \text{proj}_U \mathbf{v}\right\| \leq \| \mathbf{v} - \mathbf{u} \|$$

for every $\mathbf{u} \in U$. Furthermore, if $\mathbf{u} \in U$ and the inequality above is an equality, then $\mathbf{u} =$ projv.

Proof. Suppose that $\mathbf{u} \in U$. Then

$$\begin{aligned} \left\|\mathbf{v} - \text{proj}_U \mathbf{v}\right\|^2 &\leq \left\|\mathbf{v} - \text{proj}_U \mathbf{v}\right\|^2 + \left\|\text{proj}_U \mathbf{v} - \mathbf{u}\right\|^2 \\ &= \left\|\left(\mathbf{v} - \text{proj}_U \mathbf{v}\right) + \left(\text{proj}_U \mathbf{v} - \mathbf{u}\right)\right\|^2 \\ &= \| \mathbf{v} - \mathbf{u} \|^2, \end{aligned}$$

where the first equality comes from Pythagoras, which applies because $\left(\mathbf{v} - \text{proj}_U \mathbf{v}\right) \in U^{\perp}$ and $\left(\text{proj}_U \mathbf{v} - \mathbf{u}\right) \in U$. Taking the square roots gives the desired inequality, which is an equality if and only if

$$\left\|\mathbf{v} - \text{proj}_U \mathbf{v}\right\|^2 = \left\|\mathbf{v} - \text{proj}_U \mathbf{v}\right\|^2 + \left\|\text{proj}_U \mathbf{v} - \mathbf{u}\right\|^2.$$

This happens if and only if $\left\|\text{proj}_U \mathbf{v} - \mathbf{u}\right\| = 0$, which happens if and only if $\mathbf{u} =$ projv. ■

Example 4.19

Let V be the vector space of polynomials with inner product

$$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)dt$$

Then, $\{1, t\}$ is an orthogonal basis of $\mathbf{P}_1$. The polynomial $q(t) \in \mathbf{P}_1$ such that $\left\| t^{2k} - q(t) \right\|$ is as small as possible is

$$q(t) = \text{proj}_{\mathbf{P}_1} t^{2k} = \frac{1}{2k+1},$$

and the polynomial $r(t) \in \mathbf{P}_1$ such that $\left\| t^{2k+1} - r(t) \right\|$ is as small as possible is

$$r(t) = \text{proj}_{\mathbf{P}_1} t^{2k+1} = \frac{3t}{2k+3}.$$

4.8 Least-squares problems

Inconsistent systems of equations $A\mathbf{x} = \mathbf{b}$ arise often in applications. When a solution is demanded and none exists, the best one can do is to find an $\mathbf{x}$ that makes $A\mathbf{x}$ as close as possible to $\mathbf{b}$. Think of $A\mathbf{x}$ as an *approximation* to $\mathbf{b}$. The smaller the distance between $\mathbf{b}$ and $A\mathbf{x}$, given by $\| \mathbf{b} - A\mathbf{x} \|$, the better the approximation. The *general least-squares problem* is to find an $\mathbf{x}$ that makes $\| \mathbf{b} - A\mathbf{x} \|$ as small as possible.

Definition 4.2

If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a *least-squares solution* of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$\| \mathbf{b} - A\hat{\mathbf{x}} \| \leq \| \mathbf{b} - A\mathbf{x} \|$$

for all $\mathbf{x}$ in $\mathbb{R}^n$.

The most important aspect of the least-squares problem is that no matter what $\mathbf{x}$ we select, the vector $A\mathbf{x}$ will necessarily be in the column space, Col$A$. So we seek an $\mathbf{x}$ that makes $A\mathbf{x}$ the closest point in Col$A$ to $\mathbf{b}$. Of course, if $\mathbf{b}$ happens to be in Col$A$, then $\mathbf{b}$ *is* $A\mathbf{x}$ for some $\mathbf{x}$, and such an $\mathbf{x}$ is a "least-squares solution."

Given A and $\mathbf{b}$ as above, let

$$\hat{\mathbf{b}} = \text{proj}_{\text{Col } A} \mathbf{b}$$

where the orthogonal projection of $\mathbf{b}$ along Col$A$ is defined with respect to a given inner product. Because $\hat{\mathbf{b}}$ is in the column space of A, the equation $A\mathbf{x} = \mathbf{b}$ *is* consistent, and there is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

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

Since $\hat{\mathbf{b}}$ is the closest point in $ColA$ to $\mathbf{b}$, a vector $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if $\hat{\mathbf{x}}$ satisfies $A\mathbf{x} = \mathbf{b}$, which may have many solutions if the equation has free variables.

Suppose that $\hat{\mathbf{x}}$ satisfies $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. The projection $\hat{\mathbf{b}}$ has the property that $\mathbf{b} - \hat{\mathbf{b}}$ is orthogonal to $ColA$, so $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to each column of A. If $\mathbf{a}_j$ is any column of A, then $\left\langle \mathbf{a}_j, \mathbf{b} - A\hat{\mathbf{x}} \right\rangle = 0$, and $\mathbf{a}_j^\top G(\mathbf{b} - A\hat{\mathbf{x}}) = 0$, where G is the Gram matrix relative to an appropriate basis. Since $\mathbf{a}_j^\top$ is a row of $A^\top$,

$$A^\top G(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$$

These calculations show that each least-squares solution of $A\mathbf{x} = \mathbf{b}$ satisfies the equation

$$A^\top GA\hat{\mathbf{x}} = A^\top G\mathbf{b}$$

This matrix equation represents a system of equations called the *normal equations* for $A\mathbf{x} = \mathbf{b}$.

Theorem 4.12

The set of least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the non empty set of solutions of the normal equations $A^\top GA\hat{\mathbf{x}} = A^\top G\mathbf{b}$.

Proof. As shown above, the set of least-squares solutions is non empty and each least-squares solution $\hat{\mathbf{x}}$ satisfies the normal equations. Conversely, suppose that $\hat{\mathbf{x}}$ satisfies $A^\top GA\hat{\mathbf{x}} = A^\top G\mathbf{b}$. Then $\hat{\mathbf{x}}$ satisfies $A^\top G(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$, which shows that $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to the

rows of $A^\top$ and hence is orthogonal to the columns of A. Since the columns of A span $ColA$, the vector $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to all of $ColA$. Hence the equation

$$\mathbf{b} = A\hat{\mathbf{x}} + (\mathbf{b} - A\hat{\mathbf{x}})$$

is the decomposition of $\mathbf{b}$ into the sum of a vector in $ColA$ and a vector in $(ColA)^\perp$. By the uniqueness of the orthogonal decomposition, $A\hat{\mathbf{x}}$ must

be the orthogonal projection of $\mathbf{b}$ along $Col A$. That is, $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ and $\hat{\mathbf{x}}$ is a least-squares solution.

Example 4.20

Find the least-squares solution with respect to the usual inner product in $\mathbb{R}^2$ of the inconsistent system $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}.$$

In this case, we may assume that everything is expressed relative to the canonical basis, thus the Gram matrix is the identity matrix I_2, so the normal equations read $A^\top A\mathbf{x} = A^\top \mathbf{b}$. To use the normal equations, we compute:

$$A^\top A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^\top \mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}.$$

The equation $A^\top A\mathbf{x} = A^\top \mathbf{b}$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}.$$

Row operations can be used to solve this system, but since $A^\top A$ is invertible and 2×2, it is probably faster to compute

$$\left(A^\top A\right)^{-1} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

and then to solve $A^\top A\mathbf{x} = A^\top \mathbf{b}$ as

$$\hat{\mathbf{x}} = \left(A^\top A\right)^{-1} A^\top \mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

In many calculations $A^\top A$ is invertible, but this is not always the case as in the following example.

Example 4.21

Find a least-squares solution relative to the usual inner product of $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

As in the previous example, the Gram matrix of the inner product relative to the standard basis is the identity matrix. Compute

$$A^{\top}A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}$$

$$A^{\top}\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}.$$

The augmented matrix for $A^{\top}A\mathbf{x} = A^{\top}\mathbf{b}$ is

$$\begin{bmatrix} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

So the general least-squares solution of $\mathbf{x} = \mathbf{b}$ has the form

$$\hat{\mathbf{x}} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + \lambda \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

4.9 Orthogonal diagonalization

Let V be an inner product space. A *symmetric* linear mapping $T : V \to V$ relative to an inner product is a linear mapping such that $\langle T(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, T(\mathbf{v}) \rangle$ for all $\mathbf{u}, \mathbf{v} \in V$.

Theorem 4.13

Suppose that $T : V \to V$ is a symmetric linear mapping. If $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is an orthonormal basis for V and $M = (m_{j,k})$ is the matrix representation of T relative to $\mathcal{B}$, then $M^\top = M$.

Proof. Recall that we obtain the k-th column of M by writing $T(\mathbf{b}_k)$ as a linear combination of the basis $\mathcal{B}$; the scalars used in this linear combination then become the k-th column of M. Since $\mathcal{B}$ is orthonormal, we know how to write $T(\mathbf{b}_k)$ as a linear combination of $\mathcal{B}$:

$$T(\mathbf{b}_k) = \langle T(\mathbf{b}_k), \mathbf{b}_1 \rangle \mathbf{b}_1 + \cdots + \langle T(\mathbf{b}_k), \mathbf{b}_n \rangle \mathbf{b}_n$$

Thus $m_{j,k} = \langle T(\mathbf{b}_k), \mathbf{b}_j \rangle$. Using the symmetry of T and the symmetry of the inner product, we get

$$m_{j,k} = \langle T(\mathbf{b}_k), \mathbf{b}_j \rangle = \langle \mathbf{b}_k, T(\mathbf{b}_j) \rangle = \langle T(\mathbf{b}_j), \mathbf{b}_k \rangle = m_{k,j}$$

Therefore, $M = (m_{j,k}) = (m_{k,j}) = M^\top$. ■

One of the most important properties of symmetric linear mappings is stated in the following theorem.

Theorem 4.14

If $T : V \to V$ is a symmetric linear mapping, then any two eigenvectors from different eigenspaces are orthogonal.

Proof. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors that correspond to distinct eigenvalues, say, λ_1 and λ_2. To show that $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$, compute

$$\lambda_1 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \lambda_1 \mathbf{v}_1, \mathbf{v}_2 \rangle = \langle T(\mathbf{v}_1), \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, T(\mathbf{v}_2) \rangle = \langle \mathbf{v}_1, \lambda_2 \mathbf{v}_2 \rangle = \lambda_2 \langle \mathbf{v}_1, \mathbf{v}_2 \rangle$$

Hence $(\lambda_1 - \lambda_2) \langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$. But $\lambda_1 - \lambda_2 \neq 0$ so $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$. ■

A linear mapping $T : V \to V$ is said to be orthogonally diagonalizable if V has an orthonormal basis consisting of eigenvectors of T. The proof of the following theorem is quite involved and is ommitted.

Theorem 4.15

Let V be an inner product space. A linear mapping $T : V \to V$ is orthogonally diagonalizable if and only if T is symmetric relative to the inner product.

This theorem is rather amazing, because the work in the previous chapter would suggest that it is usually impossible to tell when a linear mapping is diagonalizable. But this is not the case for symmetric matrices.

The next example treats a linear mapping whose eigenvalues are not all distinct.

Example 4.22

Orthogonally diagonalize the linear mapping $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined relative to the canonical basis by

$$T(\mathbf{x}) = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix} \mathbf{x}.$$

The usual calculations produce the characteristic equation

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2).$$

A basis for the eigenspace associated with $\lambda_1 = 7$ is

$$\left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} \right\}$$

and a basis for $\lambda_2 = -2$ is

$$\left\{ \mathbf{v}_3 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix} \right\}$$

Although $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent, they are not orthogonal. Applying the Gram-Schmidt orthogonalization process and normalization, we obtain the following orthonormal basis for the eigenspace associated with λ_1 :

$$\left\{ \mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix} \right\}.$$

An orthonormal basis for the eigenspace associated with λ_2:

$$\left\{\mathbf{u}_3 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}\right\}.$$

Hence, $\mathcal{C} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for $\mathbb{R}^3$ and the matrix representation of T relative to $\mathcal{C}$ is

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

An interesting observation is that if $P = [\mathbf{u}_1 \ldots \mathbf{u}_n]$ is a square matrix whose columns satisfy $\mathbf{u}_i^\top \mathbf{u}_j = 0$ for $i \neq j$ and $\mathbf{u}_j^\top \mathbf{u}_j = 1$ (as in the example above), then $P^\top P = I_n$. This implies that $P^{-1} = P^\top$. Moreover, it is not difficult to verify that $P^{-1} = P^\top$ if and only if the columns of P satisfy $\mathbf{u}_i^\top \mathbf{u}_j = 0$ for $i \neq j$ and $\mathbf{u}_j^\top \mathbf{u}_j = 1$.

Universidad
Rey Juan Carlos

Material docente

Misael E. Marriaga
Michael Stich

Universidad
Rey Juan Carlos

Material docente

Editorial Academia Abierta