Linear Algebra 1

James

November 6, 2018

Let's first examine what linear algebra really even is. Generally, your first linear algebra class illustrates that it's about matrices and row reduction, which is entirely[1] the wrong idea. What you should really think of when you hear linear algebra is **linear functions**. These are extremely nice functions which give us a great deal of versatility to do what we want. I'll start by going through some basic definitions, and then we'll get to the point where we can talk about things.

Definition. (Set) A **set** is just a collection of objects.

To the more astute reader, this should appear to be a very poor definition of sets.[2] However, I'll assume you don't care about the more esoteric side of things, and so I'll breeze through this. Let's just take this at face value. What is a set? Let me give you a few examples.

Example 0.1. The collection $A = \{1, 2, 3\}$ is a set.

This seems pretty dull, and it should. Sets are not super special on their own. However, there are some more notable sets we will be working with.

Example 0.2. The collection of positive integers (that is, numbers greater than or equal to 0) is defined to be $\mathbb{N}$. These are also known as the **natural numbers**.

This should be pretty familiar (although the notation may be scary). These are just the numbers you use to count. Now, if you are a merchant of some kind, you may also want to be able to add and subtract numbers to get estimates of how much things cost, how many things you have, or any other sort of thing. To do that, we need a new set of numbers.

Example 0.3. The collection of postive and negative numbers is $\mathbb{Z}$. These are known as the **integers**.

Now, you may want to repeatedly add a number (and also may want to do the opposite). Multiplication is fine in the integers; if we repeatedly add something, we will still get an integer. However, the other direction is not so fine. For example, say we want to find a $x \in \mathbb{Z}$ (read: x in the integers) so that

$$x \cdot 2 = 3.$$

That is, if I add x to itself twice, I want to get 3. Well, there is in fact no integer that can do that. However, there is something that's sort of close. We call these things the rationals.

Example 0.4. The collection of numbers

$$\frac{p}{q},\ p \in \mathbb{Z},\ q \in \mathbb{Z},\ q \neq 0$$

is denoted by $\mathbb{Q}$, and is referred to as the **rationals**.

[1] Entirely is a strong word here. It is somewhat about this, but to the same extent that calculus is about trigonometry.

[2] Is the set of all sets which do not contain themselves a set?

Going back, I can now find that x:

$$\frac{3}{2} + \frac{3}{2} = \frac{3+3}{2} = \frac{6}{2} = 3.$$

So $x = 3/2$. Now, there's a lot of machinery I'm explicitly skipping over here. We would need to be more careful with our definitions, and need to discuss things like monoids or groups to be entirely precise. However, what I just gave above is really the gist of things.

With multiplication and addition comes things like exponents, and with exponents comes logs, and there are many things we can now do with our set. A long time ago a famous mathematician by the name of Pythagoras discovered the Pythagorean theorem. This says that, for a right triangle with sides a and b and hypotenus c, we have a formula which relates all these values;

$$a^2 + b^2 = c^2.$$

This is fine and dandy, and there is a nice geometric proof of this which I won't include here. However, up until this point we really only knew about the rational numbers. Numbers which were irrational were not really discussed. The Pythagorean theorem gives us our first example of a rational number though; let's say my triangle has sides 1 and 1 and hypotenus c. Then the Pythagorean theorem says

$$1^2 + 1^2 = 2 = c^2.$$

If I take the square root of both sides, this gives me

$$c = \sqrt{2}.$$

This is strange – if $\mathbb{Q}$ really is the biggest universe, this should reduce to some quotient or fraction of two integers. Let's try to figure out what these numbers would be. I'll first just assume that there are two numbers, let's say p and q, so that

$$\frac{p}{q} = \sqrt{2}.$$

I also want this to be the smallest ratio of numbers; that is, I can no longer divide p by q. Alright, if I square both sides, I get

$$\frac{p^2}{q^2} = 2.$$

So if I multiply both sides by q^2, I get

$$p^2 = 2 * q^2.$$

In particular, this says that p^2 is an **even** number, that is, it is divisible by 2. A number is **odd** if it is not an even number. We can generalize this to say that x is an even number if there is an integer k so that $x = 2k$, and it is odd if there is an integer k so that $x = 2k + 1$. If I have an even number multiplied to an even number, then I have

$$x \cdot y = (2k) \cdot (2l) = 4kl,$$

where k and l are integers. If I have an even number multiplied to an odd number, I have

$$x \cdot y = (2k) \cdot (2l + 1) = 4kl + 2k = 2(kl + k).$$

Now, $kl + k$ is an integer, so this says that it is even. If I have an odd number multiplied to an odd number, I get

$$x \cdot y = (2k + 1) \cdot (2l + 1) = 4kl + 2k + 2l + 1 = 2(2kl + k + l) + 1.$$

Notice $2kl + k + l$ is an integer, and so this must mean that it is odd. Going back to the problem, if p^2 is even, then this means that p must be even, since an odd number times an odd number is odd. Now, if p is even, we get that it is of the form $2k$ for some k. So we rewrite this as

$$(2k)^2 = 4k^2 = 2q^2.$$

Dividing by both sides by 2 says

$$2k^2 = q^2.$$

We have k^2 is still an integer, and so this says that q^2 is even. That is, q is even. But if p and q are **both** even, we have an issue, since I originally assumed that p could no longer be divided by q (that is, the fraction is reduced), and yet this says I should be able to take out another 2 from p. I have run into a contradiction, so my original statement must have been false.

This proof perplexed the Greeks, and some downright said it was false. It turns out this is true, and that $\mathbb{Q}$ is not the biggest universe, we can go one step higher. If I take the collection of all numbers which are like $\sqrt{2}$, that is, they are not in $\mathbb{Q}$ but can be approximated by numbers in $\mathbb{Q}$ (this idea will make more sense in real analysis), then I get something called the **real numbers**. I will not give a formal definition, but this is again the gist of things.

Moving forward in this story, the real numbers were great. You could do a lot of things with it, and it's the basis of your high school algebra class. I can't do everything (for example, I can't take the square root of negative numbers), but it's really good enough for me to get by in my everyday life. However, the real world cannot really be classified in terms of just $\mathbb{R}$ (the symbol for the real numbers). There are, for example, many other factors that I need to consider. What if I wanted to study something moving across a plane, relative to where I stand? I would need two bits of information (how far away they are in terms of horizontal to me, and how far away they are in terms of in front or behind of me). Thus, I need to move to a new universe, called $\mathbb{R}^2$. This is the collection of elements (a, b) so that $a, b \in \mathbb{R}$ individually.

How do I define operations here? I can make addition make sense – just do it component wise. And this makes sense in a geometric universe as well; when adding (x, y), I'm moving x units horizontally and y units vertically. But what does multiplication mean? Can I still do it? The answer is no, I can't, and it doesn't really make sense either. But I can do other things, like the dot product and cross product, which we'll explore more so later on. The gist is that I want to classify all of what I can do with $\mathbb{R}^n$, the direct product of $\mathbb{R}$ with itself multiple times. Moreover, I want to also be able to relate sets to other sets, and in particular relate $\mathbb{R}^n$ to other spaces like $\mathbb{R}^m$.

Before jumping into functions, we also need to discuss some operations on sets. There are four things we can do with sets: take the union, intersection, product, and complement.

Definition. (Union) Let A and B be two sets. The **union** of these sets is defined to be

$$A \cup B = \{x \; : \; x \in \text{either A or B}\}.$$

What we used above is called set builder notation: that is, on the left we have the set, on the right is a description of all the elements in that set. The : (sometimes it is a bar $|$) is supposed to be read as "such that." So, translating the notation above, we have that $A \cup B$, or A union B, is the collection of elements x such that x is in either A or B.

Definition. (Intersection) Let A and B be two sets. The **intersection** of these sets is defined to be

$$A \cap B = \{x \; : \; x \in \text{A and B}\}.$$

The product is kind of what we did with $\mathbb{R}^n$. We just keep appending more data on in the form of a tuple.

Definition. ([Cartesian] Product) The product of two sets A and B is defined to be

$$A \times B = \{(a, b) \; : \; a \in A, \; b \in B\}.$$

Before getting to complement, we need to discuss a little about subsets. We've been implicitly using this notion, but it is always good to formalize something if you're going to use it a lot.

Definition. (Subset) Let A and B be subsets. We say that A is a **subset** of B if for all $x \in A$, $x \in B$. This is denoted by $A \subset B$. If we have the possibility of equality, we write $A \subseteq B$.

This is kind of similar to your less than and less than or equal to sign. In a sense, you can think of it like that.

Definition. (Complement) Let X be a set, and $A \subset X$ a subset. Then the complement A^c is the set

$$A^c = \{x \; : \; x \in X, \; x \notin A\}.$$

Here, $\notin$ is read "not in."

Notice that for complements to make sense we need a sort of "universal" set. This is what our X is playing the role as. Now, with all of that out of the way, we can move on to functions. Functions are ways of relating sets to one another. In mathematics, and in life, this is really our goal; we want to have ways of understanding things in terms of other things that we are maybe more familiar with. Thus, the definition of a function needs to reflect that. It needs to be a way of pushing one of our sets to another set.

Definition. (Function) For two sets A and B, a **function** $f : A \to B$ takes elements from A and maps them to elements in B. We call A the domain and B the codomain.

Let's look at a silly example.

Example 0.5. Let $A = \{1, 2, 3\}$ and $B = \{1, 2\}$. Then a function $f : A \to B$ can be defined in terms of $f(1) = 1$, $f(2) = 2$, and $f(3) = 2$.

To move forward, we need to put some restriction on our functions. As of now, they are too wild for us to say anything. The first sort of restriction we put is injectivity. This is sort of an embedding restriction; this says that our set can be placed inside of another set. In the context of special kinds of functions, this will also force our set to "play nicely" within another set, although as of now we do not know how these sets are playing, so to speak.

Definition. (Injective) We say a function $f : A \to B$ is **injective** if for all $a_1, a_2 \in A$, $f(a_1) = f(a_2)$ implies $a_1 = a_2$.

Example 0.6. Consider the set $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4\}$. Then the function which send $f(1) = 1$, $f(2) = 2$, $f(3) = 3$ is injective.

Another kind of function is a surjective function. This is a function which projects information from a bigger space to a smaller space. Sometimes we want to squeeze our set into another set. Again, this says nothing of how the set will "play" in the smaller space, rather it will just tell us where things go.

Definition. (Surjective) We say a function $f : A \to B$ is **surjective** if for all $b \in B$, there is an $a \in A$ so that $f(a) = b$.

Example 0.7. Consider the set $A = \{1, 2, 3, 4\}$ and $B = \{1, 2, 3\}$. Then the function which sends $f(1) = 1$, $f(2) = 2$, $f(3) = 3$, $f(4) = 3$ is surjective, since for every element in B I can find an element in A which maps to it. It is not injective; $f(4) = f(3)$ but $3 \neq 4$.

Example 0.8. Consider the set $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4\}$. Then the function which send $f(1) = 1$, $f(2) = 2$, $f(3) = 3$ is not surjective; there is no element $a \in A$ so that $f(a) = 4$.

So we now have seen functions which are not injective but surjective, and not surjective but injective. What happens when they are injective *and* surjective? It turns out there's a special word for this, so I'll quickly throw this definition in.

Definition. (Bijective) A function $f : A \to B$ is **bijective** if it is injective and surjective.

So what can I really say if $f : A \to B$ is bijective? Well, let's think about it for a minute. If there are a certain number of elements in my domain, maybe say n, and my function is injective, then that means intuitively it embeds into a space. So I know that there is at least n elements in my codomain. Now, if my set has n elements and surjects, that means for every element in my codomain I can find an element in my domain which maps to it. So I have that my codomain has at most n. Well, if it's at least n, and at most n, it's gotta be n.

Okay, that's pretty cool. We can say something about the number of elements in a set. There's actually a fancy word for this, and it's called the cardinality, or magnitude.

Definition. (Magnitude) The **magnitude** of a set A is defined to be the number of elements in that set. Notationally we represent this by $|A|$.

Now, this is more of the intuitive definition than the formal definition, but for now this is fine. Let's go back to trying to figure out what the hell a bijection does for us. If we're finite, then this means our sets are the same size. So maybe this is how we will define size properly; that sets are the same size if they are bijective. That is, if I can place or map every element from one set uniquely to an element in another set, then they should be the same size. This seems natural enough, and in fact is kind of how we actually do things in the real world. If I had a bag of things, and another bag of things, and I had no concept of size, one way I could figure out if they are the same size is by placing my things from one bag on top of the things in the other bag. If they line up perfectly, then they are the same size. However, this is going to give us some weird things.

Example 0.9. (Weird things are afoot...) Let's go back to $\mathbb{N}$. In set builder notation, we have

$$\mathbb{N} = \{0, 1, 2, \ldots\},$$

with the trailing dots meaning that it keeps going. So if bijection is going to be our notion of size, what can I say about the set $2\mathbb{N}$? In set builder notation, this is

$$2\mathbb{N} = \{0, 2, 4, \ldots\}.$$

Intuitively (at least, when I was a child first reading about this) I would say this had roughly half the size. So I should not be able to create a bijection between this and $\mathbb{N}$. But, it turns out this is entirely wrong.

Let $f : \mathbb{N} \to 2\mathbb{N}$ be defined by $f(a) = f(2a)$. This seems to be a fine assignment, and is in fact how I would construct the even positive integers. Now recall that for a bijection to happen, we need injectivity and surjectivity. Let's start with injectivity. If $f(a) = f(b)$, then this means that $2a = 2b$. But I can divide both sides by 2 to get that $a = b$. So this is injective. Now, for surjectivity, recall that this says that for all $y \in 2\mathbb{N}$, I should be able to find an $x \in \mathbb{N}$ so that $f(x) = y$. But what if I take $y = x/2$? You may say, "But James!

If I divide by 2, I might no longer be a natural number!" Let me remind you that $y = 2a$ for some $a \in \mathbb{N}$. So what I'm doing is just working backwards, and saying $x = a$. Things work out nicely, and we get that the function is surjective.

This weird paradox turns out to be called the Hilbert Hotel paradox. The idea is that you have a very famous football game going on, and there are infinitely many people at it. At the Hilbert hotel, there is an infinite number of rooms, and Hilbert has filled them all up for the evening. All of the people from the football game come to the hotel and would like to stay the night. What is Hilbert to do? If he says no, he misses out on infinite profit. He can't let them stay though, right?

Wrong. What Hilbert is going to do is he will move all the people currently at the hotel to rooms of even numbers. From our bijection earlier, we know that this is possible, since they have the same size. Now, he can place all of the other people in the odd numbered rooms, which for similar reasons is the same size as $\mathbb{N}$. Thus, he wins and gets all the profit.

We can do more weird things like pushing them to prime numbered rooms (that is, there is a bijection between primes and $\mathbb{N}$), or every tenth room, or whatever thing you wish so long as there is a bijection. This is one of my favorite paradoxes from set theory, as it seems ridiculous until you look at it notationally.[3]

As we can see, there are weird things that happen with sets and functions. I could keep going to describe to you how $\mathbb{R}$ and $\mathbb{N}$ are not in bijection, and so therefore not the same size,[4] but I feel like this is enough for now. We should, after all, move on to actual linear algebra eventually.

We should first talk about **fields** if we're going to talk about linear algebra. A field is the nicest thing you can imagine; its a realm in which you can add, subtract, divide, and multiply. Essentially, it's a world where you can do algebra. Everything has an inverse in terms of multiplication (I just divide by it to get to 1) and everything has an inverse in terms of addition (I subtract by it to get to 0). The key examples you should hold in your head are $\mathbb{Q}$ and $\mathbb{R}$ as you read this definition.

Definition. (Field) A set F equipped with two operations[5] generally denoted by $\cdot$ (for multiplcation) and $+$ (for addition) which satisfies these properties:

(i) There is an identity for both $\cdot$ and $+$. That is, there is an element $a \in F$ so that for all $b \in F$, $a + b = b + a = a$, and there is an element $e \in F$ so that for all $b \in F$, $e \cdot b = b \cdot e = b$.

(ii) Every element has in inverse for both $\cdot$ and $+$. That is, for every element $b \in F$, there is a $-b$ so that $b + (-b) = (-b) + b = a$ (where a is as above), and for every $b \in F$ there is a b^{-1} so that $b \cdot b^{-1} = b^{-1} \cdot b = e$.

[3]Even then, it might still seem a little ridiculous. But it's how things go.

[4]See: Cantor's diagonalization argument

[5]I didn't really discuss operations, so let me do it in a footnote. Let A be a set. An operation is a function which goes from $A \times A$ into A which is closed; that is, it doesn't leave A.

It may be good for your soul to check that $\mathbb{R}$ and $\mathbb{Q}$ are both fields. I will not do that, as it's tedious and I've done it many times already. However, fields are important, as again they let us do things like algebra. Linear algebra is more about things called **vector spaces** than it is about fields though.

Definition. (Vector Spaces) A vector space V over a field F is a set V with two operations $+$ and $\cdot$ such that it satisfies these properties:

(i) For all $u, v, w \in V$, $u + (v + w) = (u + v) + w$.

(ii) For all $u, v \in V$, $u + v = v + u$.

(iii) There is an element $0 \in V$ so that for all $u \in V$ $0 + u = u + 0 = u$.

(iv) For every $u \in V$ there is a $(-u)$ so that $u + (-u) = (-u) + u = 0$.

(v) For $a, b \in F$ and $u \in V$, we have $a(bu) = (ab)u$.

(vi) If 1 is the multiplicative identity in F, then for all $u \in V$ we have $1u = u$.

(vii) For all $a \in F$, $u, v \in V$, $a(u + v) = au + av$.

(viii) For all $a, b \in F$, $u \in V$, $(a + b)u = au + bu$.

This is a complicated definition with a lot to unpackage. Maybe it's best to go back to our example $\mathbb{R}^n$. There, we wanted to do things kind of like $\mathbb{R}$; that is, we want to do algebra. But if we imagine things geometrically, we see that things like multiplication don't quite make sense. However, if I defined multiplication by an element in $\mathbb{R}$ to be just multiplying every component by that number, then that makes a lot of sense; it's simply just making your vector longer. The vector space definition is a geometric one, and it has a lot to do with understanding how these vectors interact and play with one another. However, in mathematics we always want to get more out of the little we have, and so we noticed that if we define vector spaces as above, we could do everything we were doing in $\mathbb{R}^n$ and $\mathbb{R}$ in full generality. For example, we will use this to do things in differential equations in a bit.

Let's (painfully) try showing that $\mathbb{R}^n$ (as we've defined it) over $\mathbb{R}$ is a vector space.

Example 0.10. Let $(a_1, \ldots, a_n), (b_1, \ldots, b_n), (c_1, \ldots, c_n) \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$ throughout.

(i) We have

$$(a_1, \ldots, a_n)+((b_1, \ldots, b_n)+(c_1, \ldots, c_n)) = (a_1, \ldots, a_n)+(b_1+c_1, \ldots, b_n+c_n)$$

$$= (a_1 + b_1 + c_1, \ldots, a_n + b_n + c_n) = (a_1 + b_1, \ldots, a_n + b_n) + (c_1, \ldots, c_n)$$

$$= ((a_1, \ldots, a_n) + (b_1, \ldots, b_n)) + (c_1, \ldots, c_n).$$

So this checks out.

(ii) We have

$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n)$$
$$= (b_1 + a_1, \dots, b_n + a_n) = (b_1, \dots, b_n) + (a_1, \dots, a_n).$$

(iii) The 0 here is just $(0, \dots, 0)$, since

$$(a_1, \dots, a_n) + (0, \dots, 0) = (a_1 + 0, \dots, a_n + 0) = (a_1, \dots, a_n).$$

(iv) It is just the negative. So

$$(a_1, \dots, a_n) + (-a_1, \dots, a_n) = (a_1 - a_1, \dots, a_n - a_n) = (0, \dots, 0).$$