2 Vectors and Vector Spaces

In section 1.6 we saw that the set of solutions to a homogeneous system of linear equations has the following three properties:

1. It contains the trivial solution of all zeroes;
2. The sum of two solutions is a solution;
3. Any scalar multiple of a solution is a solution.

This set of three properties is very common and very powerful, and the fundamental subject of this course is to study sets that have that list of three properties. When thought of in this abstract way, we call such a set a "vector space".

In this section we will define vector spaces, look at a number of examples, and understand some of their fundamental properties.

However, before we proceed to this abstract idea, we should look at another very concrete example, which is important enough to give its name to the idea as a whole. Thus we will begin this section with geometry, return to algebra, and finally come up with a formal system that ties everything together.

2.1 The Cartesian Plane

We'll start by considering the "Cartesian plane", (named after the French mathematician René Descartes, who is credited with inventing the idea of putting numbered cordinates on the plane).

As probably looks familiar from high school geometry, given two points A and B in the plane, we can write $\overrightarrow{AB}$ for the vector with *initial point* A and *terminal point* B.

Since a vector is just a length and a direction, the vector is "the same" if both the initial and terminal points are shifted by the same amount. If we fix an *origin* point O, then any point A gives us a vector $\overrightarrow{OA}$. Any vector can be shifted until its initial point is O, so each vector corresponds to exactly one point. We call this *standard position.*

We represent points algebraically with pairs of real numbers, since points in the plane are determined by two coordinates. Thus we use $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$ to denote the set of all ordered pairs of real numbers, giving us an algebraic description of the Cartesian plane. We define the origin O to be the "zero" point $(0, 0)$.

Definition 2.1. If $A = (x, y)$ is a point in $\mathbb{R}^2$, then we denote the vector $\overrightarrow{OA}$ by $\begin{bmatrix} x \\ y \end{bmatrix}$.

We can also denote this vector $[x, y]^T$, as we did in section 1.4.3. Poole sometimes just writes $[x, y]$, and when we don't particularly care about the geometric distinction between a point and a vector we will often write (x, y).

However, the vertical orientation is very important for a lot of calculations we will want to do, including the sort of matrix multiplication we used in section 1.4.4. Therefore we will use the vertical form when it isn't terribly inconvenient.

If we want to discuss "a vector" without specifying any coordintes, we will use a single letter, generally either boldface ($\mathbf{v}$) or with an arrow on top ($\vec{v}$).

The vector $\overrightarrow{OO}$ can't really be drawn—it's the vector with zero length—but it is very important. We call it the *zero vector* and write it as $\vec{0}$ or $\mathbf{0}$.

Example 2.2. Suppose $A = (2, 3)$ and $B = (1, 5)$. Then the vector $\overrightarrow{AB}$ has displacement in the x direction of $1 - 2 = -1$, and in the y direction of $5 - 3 = 2$. Thus it is the same as the vector $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ which begins at $(0, 0)$ and ends at $(-1, 2)$.

If we want to take the same vector $\overrightarrow{AB}$ and put its initial point at $(-1, 2)$, then the terminal point will have x coordinate $-1 - 1 = -2$ and y-coordinate $2 + 2 = 4$, and thus be at the point $(-2, 4)$.

2.1.1 Scalar Multiplication

Geometrically, a vector is a direction and a distance. A natural question to ask is "what happens if we go in the same direction, but twice as far?" Or three times, or five times, or π times as far?

Definition 2.3. If $\mathbf{v}$ is a vector and r is a positive real number, we define *scalar multiplication* by setting $r \cdot \mathbf{v}$ to be a vector with the same direction as $\mathbf{v}$, but with its length stretched by a factor of r.

If r is a negative real number then we define $r \cdot \mathbf{v}$ to be the vector with the opposite direction from $\mathbf{v}$, and length equal to $|r|$ times the length of v.

We define $0 \cdot \mathbf{v} = \mathbf{0}$ to be the zero vector.

Remark 2.4. Notice that this means $-1 \cdot \mathbf{v}$ is a vector of the same length, but pointing in the opposite direction. So $(-1) \cdot \overrightarrow{AB} = \overrightarrow{BA}$.

Example 2.5. Let $\mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$. Then we see that $2 \cdot \mathbf{v}$ must go twice as far in the same direction, and thus $2 \cdot \mathbf{v} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$. Similarly, $-2 \cdot \mathbf{v} = \begin{bmatrix} -2 \\ -6 \end{bmatrix}$. Of course, we know that $0 \cdot \mathbf{v} = \mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Looking at these examples suggests an algebraic rule for scalar multiplication:

Definition 2.6. If $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is a vector and r is a real number, then we define *scalar multiplication* by $b \cdot \mathbf{v} = b \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} bv_1 \\ bv_2 \end{bmatrix}$. We sometimes say that scalar multiplication is given by *componentwise* multiplication.

Notice that this is exactly the scalar matrix multiplication of section 1.4.1.

Example 2.7. If $\mathbf{v} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ then $7 \cdot \mathbf{v} = \begin{bmatrix} 21 \\ 35 \end{bmatrix}$ and $\pi \cdot \mathbf{v} = \begin{bmatrix} 3\pi \\ 5\pi \end{bmatrix}$.

Remark 2.8. It is very important that scalar multiplication combines two different types of information. We have a real number r, which is a "size" without direction. We also have a vector $\mathbf{v}$ which is a magnitude and direction, and we multiply these two things together.

We cannot multiply two vectors to get another vector (outside of some very specific circumstances like the cross product). We can, of course, multiply two scalars together to get another scalar; you have been doing that since elementary school.

2.1.2 Vector Addition

Another question to ask about geometric vectors is "what happens if we go in this direction for this distance, and then once we get there, go in that direction for that distance?" In our diagram of the plane, this is represented by taking two vectors and placing them "head-to-tail".

Definition 2.9. If $\mathbf{v} = \overrightarrow{AB}$ and $\mathbf{w} = \overrightarrow{BC}$, then we define *vector addition* by $\mathbf{v} + \mathbf{w} = \overrightarrow{AC}$.

Example 2.10. If $A = (1, 2)$, $B = (3, 1)$, $C = (5, -1)$, then we have $\begin{bmatrix} 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 2 \\ -2 \end{bmatrix} = \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$.

Example 2.11. If $\mathbf{v} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -4 \\ 1 \end{bmatrix}$ then we can set $A = (0, 0), B = (5, 2), C = (1, 3)$ and have $\mathbf{v} = \overrightarrow{AB}$ and $\mathbf{w} = \overrightarrow{BC}$. Then $\mathbf{v} + \mathbf{w} = \overline{AC} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Drawing a picture every time we want to add vectors gets tedious very quickly. Fortunately, vector addition is easy algebraically: we can just do *componentwise addition.*

Definition 2.12. Algebraically, we define addition of vectors by $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix}$.

You can see that this gives the same result as the head-to-tail method. And again, it is the same as the matrix addition of section 1.4.1.

Remark 2.13. Given two vectors $\mathbf{u}$ and $\mathbf{v}$, we can form a parallelogram with those vectors as two of its sides. We call this the *parallelogram determined by* $\mathbf{u}$ *and* $\mathbf{v}$. In this case, we see that $\mathbf{u} + \mathbf{v}$ is the vector corresponding to the diagram of the parallelogram.

2.2 Threespace and $\mathbb{R}^n$

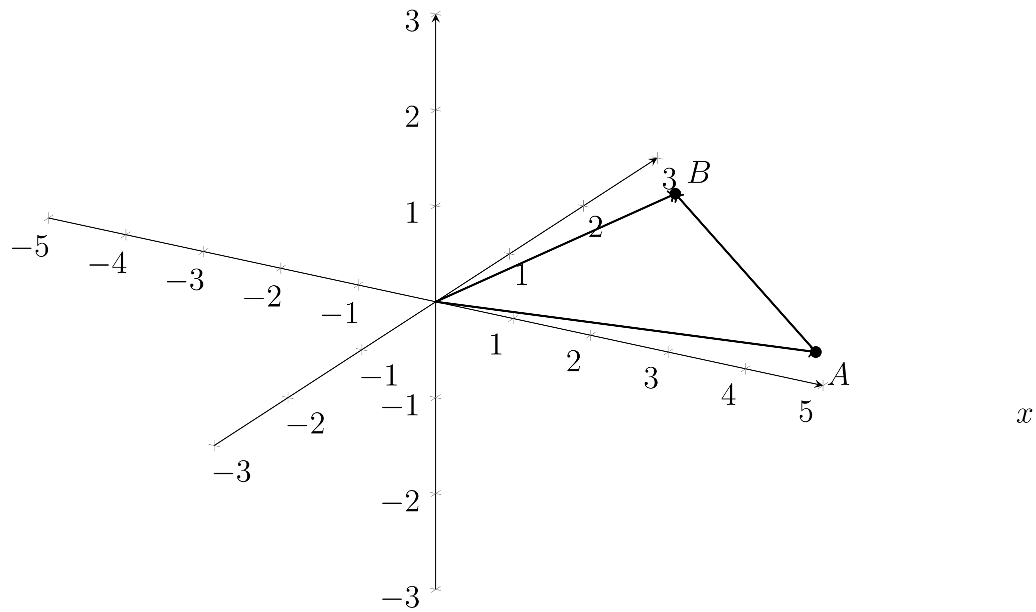
All of the work in section 2.1 took place in the "two-dimensional" plane. We can easily extend this work to three-dimensional space. Where each point in the plane requires two coordinates to express, each point in threespace requires three coordinates.

Definition 2.14. We define *Euclidean threespace* to be the three-dimensional space described by three real coordinates. We notate it $\mathbb{R}^3$. The point $(0, 0, 0)$ is called the *origin* and often notated O.

This describes familiar three-dimensional space, in which we all (apparently) live. Just as in the Cartesian plane $\mathbb{R}^2$, we can think about vectors between points.

Example 2.15. Let $A = (3, 2, -1)$ and $B = (5, -2, 3)$. Then we have

$$\overrightarrow{OA} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}, \quad \overrightarrow{OB} = \begin{bmatrix} 5 \\ -2 \\ 3 \end{bmatrix}, \text{ and } \quad \overrightarrow{AB} = \begin{bmatrix} 2 \\ -4 \\ 4 \end{bmatrix}.$$



We can do vector addition and scalar multiplication as before, too.

Example 2.16. Let $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $w = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}$. Then

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} 5 \\ 0 \\ 6 \end{bmatrix}, \quad 3 \cdot \mathbf{v} = \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}, \text{ and } \quad (-2) \cdot \mathbf{w} = \begin{bmatrix} -8 \\ 4 \\ -6 \end{bmatrix}.$$

We have so far defined two-dimensional space and three-dimensional space. Geometrically it's hard to go farther, since most of us can't visualize a four- or five-dimensional space. (The Greeks actually argued that while you could raise a number to the second power or the third power, it made no sense to talk about 3^4 since there was no reasonable geometric interpretation. This dispute was only finally resolved in 1637 when René Descartes published a geometric method of taking two line segments, and constructing a line segment whose length was the product of the original lengths; this allowed scholars to multiply two distances and obtain a distance, resolving the philosophical concerns).

But algebraically, there's no difficulty in extending our definitions to higher dimensions and more coordinates in our vectors. (This is probably a large portion of why this course is called "linear algebra" and not "linear geometry").

Definition 2.17. We define *real n-dimensional space* to be the set of n-tuples of real numbers, $\mathbb{R}^n = \{(x_1, x_2, \ldots, x_n) : x_i \in \mathbb{R}\}$.

By "abuse of notation" we will also use $\mathbb{R}^n$ to refer to the set of vectors in $\mathbb{R}^n$. We define scalar multiplication and vector addition by

$$r \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} rx_1 \\ rx_2 \\ \vdots \\ rx_n \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}.$$

Example 2.18. Let $\mathbf{v} = (1, 3, 2, 4)$ and $\mathbf{w} = (5, -1, 2, 8)$ be vectors in $\mathbb{R}^4$. Then

$$\mathbf{v} + \mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 5 \\ -1 \\ 2 \\ 8 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \\ 4 \\ 12 \end{bmatrix}, \quad -3 \cdot \mathbf{v} = \begin{bmatrix} -3 \\ -9 \\ -6 \\ -12 \end{bmatrix}.$$

The next question you might ask is "why do we want to talk about $\mathbb{R}^n$?" $\mathbb{R}^2$ and $\mathbb{R}^3$ have obvious geometric interpretations, but it's hard to imagine the geometry of $\mathbb{R}^4$, and far harder to imagine the geometry of $\mathbb{R}^{300}$, or think of what that might describe. I visit very few three hundred dimensional spaces in my life.

And it's true that when we want to talk about "geometry" per se we will find ourselves returning to $\mathbb{R}^2$ and $\mathbb{R}^3$; throughout the course I will be giving low-dimensional examples so you have pictures to mentally reference, and we will do some work on specifically three-dimensional geometry.

But it turns out that a lot of very interesting things we care about "look like" $\mathbb{R}^n$ in a very specific way. We've already seen one example, in the set of solutions to a system of linear equations. Even if a four-dimensional space doesn't make much sense, a set of equations with four variables certainly does!

But there are many other examples, and in section 2.3 we will talk about what it means to look like $\mathbb{R}^n$ in this way.

2.3 Vector Spaces

We will now define the main object we'll be studying in this course. The following definition will look long and cumbersome; it is our first venture into the *formal* perspective we mentioned on the first day of the course.

The important thing to remember is that we're describing things that look like $\mathbb{R}^n$, or like the set of solutions to a homogeneous system of linear equations; so if you get confused, think about those examples for comparison.

Definition 2.19. Let V be a set together with two operations:

- A *vector addition* which allows you to add two elements of V and get a new element of V. If $\mathbf{v}, \mathbf{w} \in V$ then the sum is denoted $\mathbf{v} + \mathbf{w}$ and must also be an element of V.
- A *scalar multiplication* which allows you to multiply an element of V by a real number (or "scalar") and get a new element of V. If $r \in \mathbb{R}$ and $\mathbf{v} \in V$ then the scalar multiple is denoted $r \cdot \mathbf{v}$ and must also be an element of V.

Further, suppose the following axioms hold for any $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$, and any $r, s \in \mathbb{R}$:

1. (Closure under addition) $\mathbf{u} + \mathbf{v} \in V$
2. (Additive commutativity) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
3. (Additive associativity) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
4. (Additive identity) There is an element $\mathbf{0} \in V$ called the "zero vector", such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ for every $\mathbf{u}$.
5. (Additive inverses) For each $\mathbf{u} \in V$ there is another element $-\mathbf{u} \in V$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. (Closue under scalar multiplication) $r\mathbf{u} \in V$
7. (Distributivity) $r(\mathbf{u} + \mathbf{v}) = r\mathbf{u} + r\mathbf{v}$
8. (Distributivity) $(r + s)\mathbf{u} = r\mathbf{u} + s\mathbf{u}$
9. (Multiplicative associativity) $r(s\mathbf{u}) = (rs)\mathbf{u}$
10. (Multiplicative Identity) $1\mathbf{u} = \mathbf{u}$.

Then we say V is a *Vector Space*, and we call its elements *vectors*.

Example 2.20. $\mathbb{R}^n$ is a vector space, with the previously defined vector addition and scalar multiplication. We check:

Let $\mathbf{u} = (u_1, \dots, u_n).\mathbf{v} = (v_1, \dots, v_n), \mathbf{w} = (w_1, \dots, w_n) \in \mathbb{R}^n, \quad r, s \in \mathbb{R}$. Then, knowing the usual rules of commutativity and associativity of basic arithmetic, we can compute:

1. $\mathbf{u} + \mathbf{v} = (u_1, \dots, u_n) + (v_1, \dots, v_n) = (u_1 + v_1, \dots, u_n + v_n) \in \mathbb{R}^n$.

2. $$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1, \dots, u_n) + (v_1, \dots, v_n) = (u_1 + v_1, \dots, u_n + v_n) \\ &= (v_1 + u_1, \dots, v_n + u_n) = (v_1, \dots, v_n) + (u_1, \dots, u_n) = \mathbf{v} + \mathbf{u}\end{aligned}$$

3. $$\begin{aligned}(\mathbf{u} + \mathbf{v}) + \mathbf{w} &= (u_1 + v_1, \dots, u_n + v_n) + (w_1, \dots, w_n) = (v_1 + u_1 + w_1, \dots, v_n + u_n + w_n) \\ &= (v_1, \dots, v_n) + (u_1 + w_1, \dots, u_n + w_n) = \mathbf{v} + (\mathbf{u} + \mathbf{w})\end{aligned}$$

4. We have $\mathbf{0} = (0, \dots, 0)$. Then
$$\mathbf{0} + \mathbf{v} = (0 + v_1, \dots, 0 + v_n) = (v_1, \dots, v_n) = \mathbf{v}.$$

5. Set $-\mathbf{u} = (-u_1, \dots, -u_n)$. Then
$$\mathbf{u} + (-\mathbf{u}) = (u_1 + (-u_1), \dots, u_n + (-u_n)) = (0, \dots, 0) = \mathbf{0}.$$

6. $$r\mathbf{u} = r(u_1, \dots, u_n) = (ru_1, \dots, ru_n) \in \mathbb{R}.$$

7. $$\begin{aligned}r(\mathbf{u} + \mathbf{v}) &= r(u_1 + v_1, \dots, u_n + v_n) = (r(u_1 + v_1), \dots, r(u_n + v_n)) \\ &= (ru_1 + rv_1, \dots, ru_n + rv_n) = (ru_1, \dots, ru_n) + (rv_1, \dots, rv_n) = r\mathbf{u} + r\mathbf{v}.\end{aligned}$$

8. $$\begin{aligned}(r + s)\mathbf{u} &= (r + s)(u_1, \dots, u_n) = ((r + s)u_1, \dots, (r + s)u_n) \\ &= (ru_1 + su_1, \dots, ru_n + su_n) = (ru_1, \dots, ru_n) + (su_1, \dots, su_n) = r\mathbf{u} + s\mathbf{u}.\end{aligned}$$

9.
$$r(s\mathbf{u}) = r(su_1, \dots, su_n) = (rsu_1, \dots, rsu_n) = rs(u_1, \dots, u_n).$$

10.
$$1\mathbf{u} = 1(u_1, \dots, u_n) = (1 \cdot u_1, \dots, 1 \cdot u_n) = (u_1, \dots, u_n) = \mathbf{u}.$$

Remark 2.21. That took forever and was incredibly tedious. (It's not actually *difficult*, just extremely annoying). I will ask you to do this exactly once during this entire course.

So what else is a vector space and "looks like $\mathbb{R}^n$"?

Example 2.22. Let $\mathcal{P}(x) = \{a_0 + a_1 x + \cdots + a_n x^n : n \in \mathbb{N}, a_i \in \mathbb{R}\}$ be the set of polynomials with real coefficients. Define addition by

$$(a_0 + a_1 x + \cdots + a_n x^n) + (b_0 + b_1 x + \cdots + b_n x^n) = (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$$

and define scalar multiplication by

$$r(a_0 + a_1 x + \cdots + a_n x^n) = ra_0 + ra_1 x + \cdots + ra_n x^n.$$

Then $\mathcal{P}(x)$ is a vector space.

Example 2.23. Let $\mathcal{F}(\mathbb{R}, \mathbb{R}) = \mathcal{F}$ be the set of functions from $\mathbb{R}$ to $\mathbb{R}$—that is, functions that take in a real number and return a real number, the vanilla functions of single-variable calculus. Define addition by $(f+g)(x) = f(x) + g(x)$ and define scalar multiplication by $(rf)(x) = r \cdot f(x)$. Then $\mathcal{F}$ is a vector space.

1. We have vector addition defined by $(f+g)(x) = f(x)+g(x)$. This does give a function, so the vector space is closed under addition.

2. $(f+g)(x) = f(x) + g(x) = g(x) + f(x) = (g+f)(x)$.

3. $((f+g)+h)(x) = (f+g)(x) + h(x) = f(x) + g(x) + h(x) = f(x) + (g+h)(x) = (f + (g+h))(x)$.

4. Let $\mathbf{0}$ be the zero function defined by $\mathbf{0}(x) = 0$. Then we see that $(f + \mathbf{0})(x) = f(x) + \mathbf{0}(x) = f(x) + 0 = f(x)$.

5. Define $(-f)(x)$ by $(-f)(x) = -f(x)$. Then $(f + (-f))(x) = f(x) + (-f)(x) = f(x) - f(x) = 0 = \mathbf{0}(x)$.

6. Define scalar multiplication by $(rf)(x) = rf(x)$. This does give a function, so the vector space is closed under multiplication.

7. $(r(f+g))(x) = r(f+g)(x) = r(f(x)+g(x)) = rf(x)+rg(x) = (rf)(x)+(rg)(x)$.

8. $((r+s)f)(x) = (r+s)f(x) = rf(x)+sf(x) = (rf)(x)+(sf)(x)$.

9. $(r(sf))(x) = r(sf)(x) = rsf(x) = (rs)f(x) = ((rs)f)(x)$.

10. $(1 \cdot f)(x) = 1f(x) = f(x)$.

Thus $\mathcal{F}(\mathbb{R}, \mathbb{R})$ is a vector space.

Example 2.24. If $A\mathbf{x} = \mathbf{0}$ is a homogeneous system of linear equations, then the set of solutions $N(A)$ is a vector space. I won't prove this now because we will shortly develop techniques to make proving this much faster and less irritating in section 2.5.

Example 2.25. The set $M_{m\times n}$ of $m \times n$ matrices is a vector space under the addition and scalar multiplication defined in section 1.4.1, with zero vector given by

$$\mathbf{0} = (0) = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

I'm not going to prove this, but you can see that it should be true for the same reason $\mathbb{R}^{mn}$ is a vector space: they're both just lists of real numbers, but one is arranged in a column and the other in a rectangle. The operations are the same.

Example 2.26. The integers $\mathbb{Z}$ are *not* a vector space (under the usual definitions of addition and multiplication). For instance, $1 \in \mathbb{Z}$ but $.5 \cdot 1 = .5 \notin \mathbb{Z}$.

(We only need to find one axiom that doesn't hold to show that a set is not a vector space, since a vector space must satisfy all the axioms).

Example 2.27. The closed interval $[0, 5]$ is not a vector space (under the usual operations) , since $3, 4 \in [0, 5]$ but $3 + 4 = 7 \notin [0, 5]$.

Example 2.28. Let $V = \mathbb{R}$ with scalar multiplication given by $r \cdot x = rx$ and addition given by $x \oplus y = 2x + y$. Then V is not a vector space, since $x \oplus y = 2x + y \neq 2y + x = y \oplus x$; in particular, we see that $3 \oplus 5 = 11$ but $5 \oplus 3 = 13$.

There are many more examples of vector spaces, but as you can see it's fairly tedious to prove that any particular thing is a vector space. In section 2.5 we'll develop a *much* easier way to establish that something is a vector space, so we won't develop any more examples now.

2.4 Properties of Vector Spaces

The great thing about the formal approach is that we can show that anything that satisfies the axioms of a vector space must aso follow some other rules. We'll establish a few of those rules here, and you will establish a few more in your homework. Of course, there's a sense in which the entire rest of this course will be spent establishing those rules.

Proposition 2.29 (Cancellation)**.** *Let V be a vector space and suppose $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ are vectors. If $\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}$, then $\mathbf{u} = \mathbf{v}$.*

Proof. By axiom we know that $\mathbf{w}$ has an additive inverse $-\mathbf{w}$. Then we have

$$\begin{aligned}
\mathbf{u} + \mathbf{w} &= \mathbf{v} + \mathbf{w} \\
(\mathbf{u} + \mathbf{w}) + (-\mathbf{w}) &= (\mathbf{v} + \mathbf{w}) + (-\mathbf{w}) \\
\mathbf{u} + (\mathbf{w} + (-\mathbf{w})) &= \mathbf{v} + (\mathbf{w} + (-\mathbf{w})) && \text{Additive associativity} \\
\mathbf{u} + \mathbf{0} &= \mathbf{v} + \mathbf{0} && \text{Additive inverses} \\
\mathbf{u} &= \mathbf{v} && \text{Additive identity.}
\end{aligned}$$

□

Proposition 2.30. *The additive inverse $-\mathbf{v}$ of a vector $\mathbf{v}$ is unique. That is, if $\mathbf{v} + \mathbf{u} = \mathbf{0}$, then $\mathbf{u} = -\mathbf{v}$.*

Proof. Suppose $\mathbf{v} + \mathbf{u} = \mathbf{0}$. By the additive inverses property we know that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$, and thus $\mathbf{v} + \mathbf{u} = \mathbf{v} + (-\mathbf{v})$. By cancellation we have $\mathbf{u} = -\mathbf{v}$. □

Remark 2.31. In our axioms we asserted that every vector *has* an inverse, but didn't require that there be only one.

Proposition 2.32. *Suppose V is a vector space with $\mathbf{u} \in V$ a vector and $r \in \mathbb{R}$ a scalar. Then:*

1. $0\mathbf{u} = \mathbf{0}$
2. $r\mathbf{0} = \mathbf{0}$

3. $(-1)\mathbf{u} = -\mathbf{u}$.

Remark 2.33. We would actually be pretty sad if any of those statements were false, since it would make our notation look very strange. (Especially the last statement). The fact that these statements *are* true justifies us using the notation we use.

Proof. 1.

$$\begin{aligned} \mathbf{u} = 1 \cdot \mathbf{u} &= (0+1)\mathbf{u} && \text{Multiplicative identity} \\ &= 0\mathbf{u} + 1\mathbf{u} && \text{Distributivity} \\ &= 0\mathbf{u} + \mathbf{u} && \text{Multiplicative identity} \\ \mathbf{0} + \mathbf{u} &= 0\mathbf{u} + \mathbf{u} && \text{Additive identity} \\ \mathbf{0} &= 0\mathbf{u} && \text{Cancellation} \end{aligned}$$

2. We know that $\mathbf{0} = \mathbf{0}+\mathbf{0}$ by additive identity, so $r\mathbf{0} = r(\mathbf{0}+\mathbf{0}) = r\mathbf{0}+r\mathbf{0}$ by distributivity. Then we have

$$\begin{aligned} \mathbf{0} + r\mathbf{0} &= r\mathbf{0} + r\mathbf{0} && \text{additive identity} \\ \mathbf{0} &= r\mathbf{0} && \text{cancellation.} \end{aligned}$$

3. We have

$$\begin{aligned} \mathbf{v} + (-1)\mathbf{v} &= 1\mathbf{1} + (-1)\mathbf{v} && \text{multiplicative inverses} \\ &= (1 + (-1))\mathbf{v} && \text{distributivity} \\ &= 0\mathbf{v} = \mathbf{0}. \end{aligned}$$

Then by uniqueness of additive inverses, we have $(-1)\mathbf{v} = -\mathbf{v}$.

□

Example 2.34. We'll give one last example of a vector space, which is both important and silly.

We define the *zero vector space* to be the set $\{\mathbf{0}\}$ with addition given by $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and scalar multiplication given by $r \cdot \mathbf{0} = \mathbf{0}$. It's easy to check that this is in fact a vector space.

Notice that we didn't ask what "kind" of object this is; we just said it has the zero vector and nothing else. As such, this could be the zero vector of any vector space at all. In section 2.5 we will talk about vector spaces that fit inside other vector spaces, like this one.

2.5 Subspaces

Our very first two examples of a vector space were the Cartesian plane and Euclidean three-space. But we see that while we can think of them as totally distinct vector spaces, the plane sits *inside* threespace, as a subset. In fact it sits inside it in a number of different ways; we can start by taking the xy plane, the xz plane, or the yz plane.

Similarly, though I haven't proven it yet, we also said that the set of solutions to a homogeneous system of linear equations forms a vector space. But if our system has three variables, then the set of solutions is contained in $\mathbb{R}^3$.

Every vector space has a number of "smaller" vector spaces sitting inside of it. In this section we will study "subspaces", which are vector spaces that are subsets of another vector space. They will be helpful in a number of ways; among these, the easiest way to show that a new object is a vector space is to show that it is a subspace of a vector space we already understand.

Definition 2.35. Let V be a vector space. A subset $W \subseteq V$ is a *subspace* of V if W is also a vector space with the same operations as V.

Example 2.36. The Cartesian plane $\mathbb{R}^2$ is a subset of threespace $\mathbb{R}^3$. Similarly the line $\mathbb{R}^1$ is a subset of the plane $\mathbb{R}^2$. (And we can stack this up as high as we want; $\mathbb{R}^7 \subseteq \mathbb{R}^8$.

Example 2.37. Let $V = \mathbb{R}^3$ and let $W = \{(x, y, x + y) \in \mathbb{R}^3\}$. Geometrically, this is a plane (given by $z = x + y$). We could in fact write $W = \{(x, y, z) : z = x + y\}$; this is a more useful way to write it for multivariable calculus, but less useful for lienar algebra. W is certainly a subset of V, so we just need to figure out if W is a subspace.

We could do this by checking all ten axioms, but that would take a very long time; we want a better tool. And it seems like we should be able to avoid a lot of that work since we *already* know many of the axioms hold in $\mathbb{R}^3$.

Proposition 2.38. *Let V be a vector space and $W \subseteq V$. Then W is a subspace of V if and only if the following three "subspace" conditions hold:*

1. $\mathbf{0} \in W$ *(zero vector);*
2. *Whenever $\mathbf{u}, \mathbf{v} \in W$ then $\mathbf{u} + \mathbf{v} \in W$ (Closed under addition); and*
3. *Whenever $r \in \mathbb{R}$ and $\mathbf{u} \in W$ then $r\mathbf{u} \in W$ (Closed under scalar multiplication).*

Proof. Suppose W is a subspace of V. Then W is a vector space, so it contains a zero vector and is closed under addition and multiplication by the definition of vector spaces.

Conversely, suppose $W \subseteq V$ and the three subspace conditions hold. We need to check the ten axioms of a vector space. But most of these properties are inherited from the fact that any element of W is also an element of V, and W has the same operations as V.

Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in W$ (and thus $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$), and $r, s \in \mathbb{R}$.

1. W is closed under addition by hypothesis.
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ since V is a vector space.
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ since V is a vector space.
4. $\mathbf{0} \in W$ by hypothesis, and $\mathbf{u} + \mathbf{0} = \mathbf{u}$ since V is a vector space.
5. $-\mathbf{u} = (-1)\mathbf{u} \in W$ by closure under scalar multiplication.
6. W is closed under scalar multiplication by hypothesis.
7. $r(\mathbf{u} + \mathbf{v}) = r\mathbf{u} + r\mathbf{v}$ since V is a vector space.
8. $(r + s)\mathbf{u} = r\mathbf{u} + s\mathbf{u}$ since V is a vector space.
9. $(rs)\mathbf{u} = r(s\mathbf{u})$ since V is a vector space.
10. $1\mathbf{u} = \mathbf{u}$ since V is a vector space.

Thus W satisfies the axioms of a vector space, and is itself a vector space. □

Example 2.39. To continue our earlier example of $W = \{(x, y, x + y)\}$, we only need to check three things. If $(x_1, y_1, x_1 + y_1), (x_2, y_2, x_2 + y_2) \in W$ then

$$\begin{bmatrix} x_1 \\ y_1 \\ x_1 + y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ x_2 + y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ (x_1 + x_2) + (y_1 + y_2) \end{bmatrix} \in W.$$

If $r \in \mathbb{R}$, then

$$r \begin{bmatrix} x \\ y \\ x + y \end{bmatrix} = \begin{bmatrix} rx \\ ry \\ (rx) + (ry) \end{bmatrix} \in W.$$

And the zero vector is

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0+0 \end{bmatrix} \in W.$$

Thus W is a subspace of V.

Example 2.40. If V is a vector space, then 0 and V are both subspaces of V. We don't actually need to check anything here, since both are clearly subsets of V, and both are already known to be vector spaces.

(When we want to ignore this possibility we will refer to "proper" subspaces, which are neither the trivial space nor the entire space).

Example 2.41. Let $V = \mathbb{R}^2$ and let $W = \{(x, x^2)\} = \{(x, y) : y = x^2\} \subseteq V$. Then W is *not* a subspace (and thus not a vector space):

W does in fact contain the zero vector $(0, 0) = (0, 0^2)$. But we see that $(1, 1) \in W$, and $(1, 1) + (1, 1) = (2, 2) \notin W$. Thus W is not a subspace.

Example 2.42. Let $V = \mathbb{R}^3$ and let $W = \{(x, 0, x) \in \mathbb{R}^3\}$. Is W a subspace of $\mathbb{R}^3$?

We need to check three things.

1. $(0, 0, 0) \in W$ ("by inspection", which basically means "look at it and see that this is true").
2. If $(x, 0, x), (y, 0, y) \in W$, then $(x, 0, x) + (y, 0, y) = (x + y, 0, x + y) \in W$.
3. If $r \in \mathbb{R}$ and $(x, 0, x) \in W$ then $r(x, 0, x) = (rx, 0, rx) \in W$.

Example 2.43. Now let $V = \mathbb{R}^3$ and let $W = \{(x, 1, x) \in \mathbb{R}^3\}$. Is W a subspace of $\mathbb{R}^3$?

We need to check the three properties. But we see in fact that $(0, 0, 0) \notin W$ so this is not a subspace.

Corollary 2.44. *If $A\mathbf{x} = \mathbf{0}$ is a homogeneous system of linear equations, and U is the set of solutions to this system, then U is a subspace of $\mathbb{R}^n$.*

Proof. This follows from proposition 1.49. □

Remark 2.45. The converse is also true: every subspace of $\mathbb{R}^n$ is the set of solutions to some homogeneous system of linear equations. We won't prove this until later.

Example 2.46. Let's look at our earlier examples again. We took $W \subset \mathbb{R}^3$ defined by $W = \{x, y, x + y\}$. This is precisely the set of solutions to the linear equation $x + y - z = 0$.

We also had the subspace given by $W = \{x, 0, x)\}$. This is the solution to the system of linear equations

$$x - z = 0$$
$$y = 0.$$

In contrast, we can see that $\{(x, 1, x)\}$ is the solution to the system

$$x - z = 0$$
$$y = 1$$

which is not homogeneous. And $W = \{(x, x^2)\}$ is the solution to the equation $x^2 - y = 0$, which *is* homogeneous, but isn't *linear*. We saw that neither of these is a vector space.

Example 2.47. Let $V = \mathcal{P}(x)$ and let $W = \{a_1 x + \cdots + a_n x^n\} = x\mathcal{P}(x)$ be the set of polynomials with zero constant term. Is W a subspace of V?

1. The zero polynomial $0 + 0x + \cdots + 0x^n = 0$ certainly has zero constant term, so is in W.
2. If $a_1 x + \cdots + a_n x^n$ and $b_1 x + \cdots + b_n x^n \in W$, then

   $$(a_1 x + \cdots + a_n x^n) + (b_1 x + \cdots + b_n x^n) = (a_1 + b_1)x + \cdots + (a_n + b_n)x^n \in W.$$

   Alternatively, we can say that if we add two polynomials with zero constant term, their sum will have zero constant term.
3. If $r \in \mathbb{R}$ and $a_1 x + \cdots + a_n x^n \in W$, then

   $$r\,(a_1 x + \cdots + a_n x^n) = (ra_1)x + \cdots + (ra_n)x^n$$

   has zero constant term and is in W.

Thus W is a subspace of V.

Example 2.48. Let $V = \mathcal{P}(x)$ and let $W = \{a_0 + a_1 x\}$ be the space of linear polynomials. Then W is a subspace of V.

1. The zero polynomial $0 + 0x \in W$.

2. If $a_0 + a_1x, b_0 + b_1x \in W$, then $(a_0 + a_1x) + (b_0 + b_1x) = (a_0 + b_0) + (a_1 + b_1)x \in W$.

3. If $r \in \mathbb{R}$ and $a_0 + a_1x \in W$, then $r(a_0 + a_1x) = ra_0 + (ra_1)x \in W$.

Example 2.49. Let $V = \mathcal{P}(x)$ and let $W = \{1 + ax\}$ be the space of linear polynomials with constant term 1. Is W a subspace of V?

No, because $0 = 0 + 0x \notin W$.

Exercise 2.50. *Fix a natural number $n \geq 0$. Let $V = \mathcal{P}(x)$ and let $W = \mathcal{P}_n(x) = \{a_0 + a_1x + \cdots + a_nx^n\}$ be the set of polynomials with degree at most n. Then $\mathcal{P}_n(x)$ is a subspace of $\mathcal{P}(x)$.*

Example 2.51. Let $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ be the space of functions of one real variable, and let $W = \mathcal{D}(\mathbb{R}, \mathbb{R})$ be the space of differentiable functions from $\mathbb{R}$ to $\mathbb{R}$. Is W a subspace of V?

1. The zero function is differentiable, so the zero vector is in W.

2. From calculus we know that the derivative of the sums is the sum of the derivatives; thus the sum of differentiable functions is differentiable. That is, $(f + g)'(x) = f'(x) + g'(x)$.

   So if $f, g \in W$, then f and g are differentiable, and thus $f + g$ is differentiable and thus in W.

3. Again we know that $(rf)'(x) = rf'(x)$. If f is in W, then f is differentiable. Thus rf is differentiable and therefore in W.

Example 2.52. Let $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ and let $W = \mathcal{F}([a, b], \mathbb{R})$ be the space of functions from the closed interval $[a, b]$ to $\mathbb{R}$. We can view W as a subset of V by, say, looking at all the functions that are zero outside of $[a, b]$. Is W a subspace of V?

1. The zero function is in W.

2. If f and g are functions from $[a, b] \to \mathbb{R}$, then $(f + g)$ is as well.

3. If f is a function from $[a, b] \to \mathbb{R}$, then rf is as well.

Example 2.53. Let $V = \mathcal{F}(\mathbb{R}, \mathbb{R})$ and let $W = \mathcal{F}(\mathbb{R}, [a, b])$ be the space of functions from $\mathbb{R}$ to the closed interval $[a, b]$. Is W a subspace of V?

No! The simplest condition to check is scalar multiplication. Let $f(x) = b$ be a function in V. Let $r = (b + 1)/b$. Then $(rf)(x) = fb = b + 1$ and thus $rf \notin W$.