Vector Spaces

We have already taken a look at $\mathbb{R}^n$. Now to discuss what is essential about it to make it a proper 'Vector Space.' Well, eventually, first a few examples.

Example:

Consider the set

$$\left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad a, b, c, d \in \mathbb{R} \right\}$$

with the addition

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

and multiplication

$$h \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ah & bh \\ ch & dh \end{bmatrix}.$$

This is a perfectly workable space. It's officially called $\mathcal{M}_{2\times 2}$. You can have any $\mathcal{M}_{n\times m}$ space.

So, how different is this from $\mathbb{R}^4$? Answer: pretty much just format. They are different spaces, but very very similar.

The zero vector is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Example Operation: $$\begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ 1 & 1 \end{bmatrix}.$$

Example: Spaces of Polynomials on $\mathbb{R}$, written $\mathcal{P}_n$. For example, $\mathcal{P}_3$ has elements of the form

$$p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$
$$q(x) = b_3x^3 + b_2x^2 + b_1x + b_0$$

with all a, b terms in $\mathbb{R}$. Addition is easy:

$$(p+q)(x) = p(x) + q(x) = (a_3+b_3)x^3 + (a_2+b_2)x^2 + (a_1+b_1)x + (a_0+b_0)$$

as is scalar multiplication

$$(cp)(x) = cp(x) = ca_3x^3 + ca_2x^2 + ca_1x + ca_0.$$

All very easy. We can create new $\mathcal{P}$ spaces by changing n, so $\mathcal{P}_4$ will have x^4 terms and coefficients, etc.

The general $\mathcal{P}_n$ has polynomials of the form

$$p(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0,$$

$$q(x) = b_nx^n + b_{n-1}x^{n-1} + \cdots + b_1x + b_0$$

with the addition and scalar multiplication exactly what you would expect.

One more space, though: $\mathcal{P}$ is the space of ALL real polynomials. As a result, it includes elements of the form

$$p(x) = a_m x^m + a_{m-1} x^{m-1} + \cdots + a_1 x + a_0 /$$

Notice one thing: there is just the m which has to be 0, 1, 2, and so on. It can be as big as you want. Every single, individual, $p(x) \in \mathcal{P}$ has to have a top x power, but there is no top x power in $\mathcal{P}$. The value m can be as big as necessary.

Now to make things even worse:

Example: The space of *Functions* on any domain $D \subset \mathbb{R}$, $\mathcal{F}(D)$. This space usually includes a domain D. The text writes it $\mathbb{F}[D]$. This covers *all* functions on D:

$$f(x) \in \mathbb{R} \quad \text{for all } x \in D.$$

One thing to note: unlike with polynomials, there is no fullproof way to identify any random function without checking what its values are. With polynomials, you an just look at the coefficients, it is not quite so simple for arbitrary functions. As a result, everything is based on output for all $x \in D$.

$$f = g \qquad \Longleftrightarrow \qquad f(x) = g(x) \text{ for all } x \in D.$$

$$\text{Adding } f, g \in \mathcal{F}(D): \qquad (f+g)(x) = f(x) + g(x)$$

$$\text{Multiplication by } a \in \mathbb{R}: \qquad (af)(x) = af(x).$$

The zero vector for $\mathcal{F}$ is a function that simply outputs zero for all $x \in D$, something like

$$z(x) = 0, \qquad \text{for all } x \in D.$$

We get operations like

$$f(x) = e^x, \quad g(x) = e^{x^2}\sin(x)$$

$$(3f+g): \qquad (3f+g)(x) \quad = \quad 3f(x) + g(x) \quad = \quad 3e^x + e^{x^2}\sin(x).$$

Defining Vector Spaces:

First, lets call our vector space V. The space has an underlying set of vectors, and must be non-empty. Next, there are two operations: vector addition and scalar multiplication. The next ten (yes, ten) axioms are divided into two sets based on each.

First we look at addition.

We have an operation $\boxplus$ which takes any two elements in V, $\mathbf{x}$ and $\mathbf{y}$ and outputs $\mathbf{x} \boxplus \mathbf{y}$, another vector. Notice we use $\boxplus$, not the usual $+$, since *this is not necessarily going to be a familiar operation*. (When the scalar (real) values are added or multiplied, this is the regular operation.)

A 1. V is closed under $\boxplus$: for all $\mathbf{x}$, $\mathbf{y} \in V$, $\mathbf{x} \boxplus \mathbf{y}$ is in V

A 2. $\boxplus$ commutes: $\mathbf{x} \boxplus \mathbf{y} = \mathbf{y} \boxplus \mathbf{x}$

A 3. $\boxplus$ is distributive: $\mathbf{x} \boxplus (\mathbf{y} \boxplus \mathbf{z}) = (\mathbf{x} \boxplus \mathbf{y}) \boxplus \mathbf{z}$, for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$

A 4. There exits a zero vector: $\mathbf{0} \in V$ such that $\mathbf{x} \boxplus \mathbf{0} = \mathbf{x}$ for all $\mathbf{x} \in V$

A 5. Every $\mathbf{x} \in V$ has a negative vector (additive inverse) $-\mathbf{x} \in V$ such that $\mathbf{x} \boxplus (-\mathbf{x}) = \mathbf{0}$.

Of those five, four are primarily about $\boxplus$ itself, making sure it has the correct form and so on. Only the first relates primarily to the underlying set. Of the five, the closure property is the one we will use the most.

Next we look to scalar multiplication.

So, an operation $\boxtimes$ which takes a vector $\mathbf{x}$ and a scalar $a \in \mathbb{R}$ and outputs $a \boxtimes \mathbf{x}$, also a vector. Again, $\boxtimes$ is not our usual 'multiply every vector by a' operation, it can, and will, be something weird.

S 1. V is closed under scalar multiplication: for all $a \in \mathbb{R}$ and $\mathbf{x} \in V$, $a \boxtimes \mathbf{x} \in V$

S 2. For all $a \in \mathbb{R}$ and $\mathbf{x}, \mathbf{y} \in V$, $a \boxtimes (\mathbf{x} \boxplus \mathbf{y}) = (a \boxtimes \mathbf{x}) \boxplus (a \boxtimes \mathbf{y})$

S 3. For all $a, b \in \mathbb{R}$ and $\mathbf{x} \in V$ $(a+b) \boxtimes \mathbf{x} = (a \boxtimes \mathbf{x}) \boxplus (b \boxtimes \mathbf{x})$

S 4. For all $a, b \in \mathbb{R}$ and $x \in V$, $(ab) \boxtimes \mathbf{x} = a \boxtimes (b \boxtimes \mathbf{x})$

S 5. $1 \boxtimes \mathbf{x} = \mathbf{x}$

Those are the *Axioms*, that means the absolute, bare bones, necessary, properties. These are what you start with to have a vector space. They are merely stated.

Example: this one is a classic, actually. Well, part of it is. Define the set of vectors

$$\left\{ \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \begin{matrix} x > 0, \in \mathbb{R} \\ y \in \mathbb{R} \end{matrix} \right\}.$$

We have only strictly positive real values on top, all of them on the bottom. We can actually make this the set of vectors for a vector space if we use the following operations:

$$\begin{bmatrix} x \\ y \end{bmatrix} \boxplus \begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} xw \\ y+z \end{bmatrix}$$

and

$$a \boxtimes \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x^a \\ ay \end{bmatrix}$$

so the usual on the bottom, and multiplication and exponentials on top. This actually works. Note that when we don't use $\boxtimes$ we are applying the usual $\mathbb{R}$ multiplication.

The zero vector for this set is $\mathbf{0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Note that this makes $0 \boxtimes \mathbf{x} = \mathbf{0}$.

Example: the space of *Equations*, which we will call $\mathcal{E}$. We will first look at $\mathcal{E}_3$, equations with three variables. The principles will remain the same.

Elements will be of the form

$$E_1 : \quad a_1x_1 + a_2x_2 + a_3x_3 = a_0$$

and

$$E_2 : \quad b_1x_1 + b_2x_2 + b_3x_3 = b_0.$$

We add them by, well, adding on both sides, so

$$E_1 \boxplus E_2 : \quad (a_1+b_1)x_1 + (a_2+b_2)x_2 + (a_3+b_3)x_3 = (a_0+b_0).$$

We multiply by, well, multiplying each term, so

$$d \boxtimes E_1 : \quad da_1x_1 + da_2x_2 + da_3x_3 = da_0.$$

Notice that this arrangement includes the trivial equation

$$E_0 : 0x_1 + 0x_2 + 0x_3 = 0$$

which is the zero vector for all $\mathcal{E}$ spaces. It also includes equations with no answer, like

$$0x_1 + 0x_2 + 0x_3 = 1.$$

Example:

The points in the plane $3x - 2y + 5z = 2$ do *NOT* form a vector space under standard $\mathbb{R}^3$ addition and multiplication. Notice the zero vector is missing from that set.

Oddly enough, the set of points in $3x - 2y + 5z = 0$ *DO* form a vector space under standard $\mathbb{R}^3$ addition and multiplication.

Properties of Vector Spaces:

These are essential properties that are derivable from the axioms. Vector spaces can't work without them, but they are not in the axiom set, because they are implied. Remember, axioms are *supposed* to be absolutely bare bones requirements.

First: what the text calls the 'Cancelation Property':

$$\mathbf{x} \boxplus \mathbf{z} = \mathbf{y} \boxplus \mathbf{z} \qquad \text{means that} \qquad \mathbf{x} = \mathbf{y}.$$

This is easy to see, and it's also easy to see how essential this is to the functioning of a vector space (and to basic sanity). We will now prove it in a manner called *From First Principles* which means using only axioms and definitions. In these cases, it is particularly useful to reference which axioms are used, either by name or the convenient numbering system we used earlier (A1, A2, S1 etc).

We start with

$$\mathbf{x} \boxplus \mathbf{z} = \mathbf{y} \boxplus \mathbf{z}.$$

The first step is to take $-\mathbf{z}$, which we know exists due to axiom A5. We add it to each side

$$(\mathbf{x} \boxplus \mathbf{z}) \boxplus -\mathbf{z} = (\mathbf{y} \boxplus \mathbf{z}) \boxplus -\mathbf{z}.$$

Next, we use the distributive property of $\boxplus$, axiom A3 for

$$\mathbf{x} \boxplus (\mathbf{z} \boxplus -\mathbf{z}) = \mathbf{y} \boxplus (\mathbf{z} \boxplus -\mathbf{z}).$$

Now to use axiom A5 again, and use the negatives for

$$\mathbf{x} \boxplus \mathbf{0} = \mathbf{y} \boxplus \mathbf{0}$$

which we convert using the zero vector related axiom, A4, for

$$\mathbf{x} = \mathbf{y}.$$

That took a while, but, as you will find, proving the blindingly obvious can be difficult in mathematics. Now to look at some essential properties.

B 1. $0 \boxtimes \mathbf{x} = \mathbf{0}$

B 2. $a \boxtimes \mathbf{0} = \mathbf{0}$

B 3. $a \boxtimes \mathbf{x} = \mathbf{0}$ implies that either $a = 0$ or $\mathbf{x} = \mathbf{0}$, or possibly both.

B 4. $(-1) \boxtimes \mathbf{x} = -\mathbf{x}$

B 5. $(-a) \boxtimes x = -(a \boxtimes \mathbf{x}) = a \boxtimes (-\mathbf{x})$

All of these can be proven, in a few steps (mostly) using just the axioms (and the cancelation property).

For instance: $0 \boxtimes \mathbf{x} = \mathbf{0}$.

$$(0 \boxtimes \mathbf{x}) \boxplus \mathbf{x} = (0 \boxtimes \mathbf{x}) \boxplus (1 \boxtimes \mathbf{x})$$

using axiom S5. Next, we use axiom S3 for

$$(0 \boxtimes \mathbf{x}) \boxplus (1 \boxtimes \mathbf{x}) = (0+1) \boxtimes \mathbf{x} = 1 \boxtimes \mathbf{x} = \mathbf{x}$$

using standard $\mathbb{R}$ addition then axiom S5, again. Next we use axiom A4 to get

$$(0 \boxtimes \mathbf{x}) + \mathbf{x} = \mathbf{0} \boxplus \mathbf{x}$$

and the cancelation property gives us $0 \boxtimes \mathbf{x} = \mathbf{0}$.

A Warning

This format, with the $\boxplus$ and $\boxtimes$ is intended for clarity. The idea is to make it abundantly clear that these are NOT necessarily familiar operations, and to distinguish them from the usual addition and multiplication used for the scalars. The $\boxplus$, $\boxtimes$ format will probably not be used on the midterm, exam, or assignments. It is not standard, though it is not incorrect or anything. The textbook writes

$$a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$$

and not the more explicit

$$a \boxtimes (\mathbf{x} \boxplus \mathbf{y}) = (a \boxtimes \mathbf{x}) \boxplus (a \boxtimes \mathbf{y}).$$

When working with these types of operations, switching to odd names for the vector addition and scalar multiplication can be useful, and you are free to do so.

Exercises

Some questions from section 5.1

- 4. I will try to avoid asking you to test all 10 axioms on an exam or assignment, but this is neither. Have fun.
- 5. Something similar is on the assignment. Possible source of confusion: yes, the vectors are just real numbers. The difference between '$+$' and '$\boxplus$' becomes particularly important.
- 14, 21, 22. Bunch of questions that should be done using the axioms, though feel free to use the cancelation property.