What is a Vector Space?
Geoffrey Scott

These are informal notes designed to motivate the abstract definition of a vector space to my MAT185 students. I had trouble understanding abstract vector spaces when I took linear algebra – I hope these help!

Why we need vector spaces

By now in your education, you've learned to solve problems like the one below.

> **Problem 1:** Find $x_1, x_2, x_3 \in \mathbb{R}$ that solve the system of linear equations
> $$3x_1 + 2x_2 + 0x_3 = 8$$
> $$1x_1 + 0x_2 + 1x_3 = 2$$
> $$2x_1 + 3x_2 + 8x_3 = 7$$

You also learned how to write problem 1 as a matrix equation:

> **Problem 1 (re-worded):** Find $x_1, x_2, x_3 \in \mathbb{R}$ that solve the following matrix equation.
> $$\begin{pmatrix} 3 & 2 & 0 \\ 1 & 0 & 1 \\ 2 & 3 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \\ 7 \end{pmatrix}$$

You can also write problem 1 using just column matrices:

> **Problem 1 (re-worded again):** Find $x_1, x_2, x_3 \in \mathbb{R}$ that solve the equation
> $$x_1 \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 1 \\ 8 \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \\ 7 \end{pmatrix}$$

Using column matrices makes the problem look similar (in a way I describe in the next paragraph) to the next two kinds of problems, both of which also appear in maths and sciences.[1]

> **Problem 2:** Find $x_1, x_2, x_3 \in \mathbb{R}$ such that
> $$x_1(3t^2 + 5t - 2) + x_2(0t^2 - t + 6) + x_3(9t^2 + 0t + 1) = 6t^2 + 9t + 2$$
>
> **Problem 3:** Find $x_1, x_2, x_3, x_4, \cdots \in \mathbb{C}$ such that
> $$x_1(\sin(\pi t)) + x_2(\sin(2\pi t)) + x_3(\sin(3\pi t)) + \cdots = e^{5it}$$

---

[1]Problem 3 is the kind of problem a physicist would solve if she knew the momentum of a particle in a box and wanted to know its energy.

In problems 1, 2 and 3, we have a certain type of mathematical object (column matrices in problem 1, polynomials in problem 2, functions in problem 3) and our goal is to write the object on the right side of the equation as a sum of the objects on the left side by finding the correct values for the x_i coefficients. We can find examples of this type of problem using *any* kind of mathematical object, as long as it's possible to multiply the objects by numbers and add them together.

We already know how to solve problem 1, where the "type of mathematical object" is "column matrices" and the coefficients are real numbers. But how do we solve problems 2 and 3? Should we spend weeks of class time learning how to solve these sorts of problems when the "type of mathematical object" is polynomials instead of column matrices? And again for functions? And what if the coefficients are from a different number system than real numbers, such as complex numbers? Will that take weeks to learn too?

The good news is that we can use the same theory and techniques to solve all problems of this kind. But how, you might ask, can we even *state* theorems and techniques that apply to such disparate types of mathematical objects? Will each of these theorems start with the phrase below?

> "Suppose you have a collection of column vectors *or* polynomials *or* functions *or* any other type of mathematical object that can be multiplied by numbers and added together..."

Actually, yes, this is indeed how many of our theorems will start! But mathematicians like to be concise, so they invented the term *vector space* to mean any type of mathematical object that can be multiplied by numbers and added together. This way, the theorems start with the phrase

> "Let V be a vector space..."

instead of the vague rambling phrase above.

Definition of a Vector Space

Before I give the formal definition of a vector space, I first need to define the concept of a *field* of numbers[2]; these will be the numbers allowed as coefficients ($\mathbb{R}$ in problems 1 and 2 above; $\mathbb{C}$ in problem 3).

Definition: A **field** is a set $\mathbb{F}$ of numbers with the property that if $a, b \in \mathbb{F}$, then $a + b, a - b, ab$ and a/b are also in $\mathbb{F}$ (assuming, of course, that $b \neq 0$ in the expression a/b).

Examples: We're familiar with how to add, subtract, multiply, and divide the following sets of numbers

$$\mathbb{N} = \{0, 1, 2, 3, \dots\}$$
$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$
$$\mathbb{Q} = \left\{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\right\}$$
$$\mathbb{R} = \text{all real numbers}$$
$$\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\}$$

---

[2]The definition of a field below is wrong. The correct definition uses the concept of a *commutative ring* instead of a *number*, and mathematicians actually study fields of things that don't look like numbers at all! You might learn the definition of a commutative ring if you take more maths courses – in this class, it's okay to think of fields as sets of numbers.

However, not all of these *sets* of numbers are *fields* of numbers. For example, 3 and 5 are in $\mathbb{N}$, but $3 - 5$ is not. Also, 3 and 5 are in $\mathbb{Z}$, but $3/5$ is not. This shows that $\mathbb{N}$ and $\mathbb{Z}$ are not fields of numbers. However, $\mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ are all fields of numbers.

There are other (weird) examples of fields, but for this class you may assume that the word "field" means either $\mathbb{Q}, \mathbb{R}$ or $\mathbb{C}$.

Definition: A **vector space** consists of a set V (elements of V are called **vectors**), a field $\mathbb{F}$ (elements of $\mathbb{F}$ are called **scalars**), and two operations

- An operation called *vector addition* that takes two vectors $v, w \in V$, and produces a third vector, written $v + w \in V$.
- An operation called *scalar multiplication* that takes a scalar $c \in \mathbb{F}$ and a vector $v \in V$, and produces a new vector, written $cv \in V$.

which satisfy the following conditions (called *axioms*).

1. Associativity of vector addition: $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$.
2. Existence of a zero vector: There is a vector in V, written 0 and called the **zero vector**, which has the property that $u + 0 = u$ for all $u \in V$
3. Existence of negatives: For every $u \in V$, there is a vector in V, written $-u$ and called the **negative of** u, which has the property that $u + (-u) = 0$.
4. Associativity of multiplication: $(ab)u = a(bu)$ for any $a, b \in \mathbb{F}$ and $u \in V$.
5. Distributivity: $(a + b)u = au + bu$ and $a(u + v) = au + av$ for all $a, b \in \mathbb{F}$ and $u, v \in V$.
6. Unitarity: $1u = u$ for all $u \in V$.

Different authors write the definition of a vector space differently. For example, the definition from the course notes has two additional axioms: the sum of two vectors must be a vector, and the multiple of a vector by a scalar is a vector. In the definition I wrote above, these axioms are part of the *definition* of the vector addition and scalar multiplication operations. Of course, these are just two ways of writing the same definition: in both cases, the sum of two vectors must be a vector and the scalar multiple of a vector with a scalar must be a vector.

No matter how it's written, the definition of a vector space looks like abstract nonsense the first time you see it. But it turns out that you already know lots of examples of vector spaces; let's start with the most familiar one.

The Familiar Example of a Vector Space: Column Matrices

Let V be the set of n by 1 column matrices of real numbers, let the field of scalars be $\mathbb{R}$, and define vector addition and scalar multiplication by

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \qquad c \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}$$

Let's verify that the conditions for being a vector space over $\mathbb{R}$ are satisfied in this case.

1. Associativity of vector addition:

$$\left(\left(\begin{array}{c}x_1\\x_2\\\vdots\\x_n\end{array}\right)+\left(\begin{array}{c}y_1\\y_2\\\vdots\\y_n\end{array}\right)\right)+\left(\begin{array}{c}z_1\\z_2\\\vdots\\z_n\end{array}\right)=\left(\begin{array}{c}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{array}\right)+\left(\begin{array}{c}z_1\\z_2\\\vdots\\z_n\end{array}\right)$$

$$=\left(\begin{array}{c}(x_1+y_1)+z_1\\(x_2+y_2)+z_2\\\vdots\\(x_n+y_n)+z_n\end{array}\right)$$

$$=\left(\begin{array}{c}x_1+(y_1+z_1)\\x_2+(y_2+z_2)\\\vdots\\x_n+(y_n+z_n)\end{array}\right)$$

$$=\left(\begin{array}{c}x_1\\x_2\\\vdots\\x_n\end{array}\right)+\left(\left(\begin{array}{c}y_1\\y_2\\\vdots\\y_n\end{array}\right)+\left(\begin{array}{c}z_1\\z_2\\\vdots\\z_n\end{array}\right)\right)$$

2. Existence of a zero vector is proven by showing that the all-zero column matrix satisfies the conditions for being a zero vector:

$$\left(\begin{array}{c}x_1\\x_2\\\vdots\\x_n\end{array}\right)+\left(\begin{array}{c}0\\0\\\vdots\\0\end{array}\right)=\left(\begin{array}{c}x_1\\x_2\\\vdots\\x_n\end{array}\right)$$

3. Existence of negatives is proven by showing that for any column matrix v, the new column matrix $-v$ constructed by multiplying every entry of v by -1 satisfies the condition for being the negative of v. That is, $v+(-v)=0$:

$$\left(\begin{array}{c}x_1\\x_2\\\vdots\\x_n\end{array}\right)+\left(\begin{array}{c}-x_1\\-x_2\\\vdots\\-x_n\end{array}\right)=\left(\begin{array}{c}0\\0\\\vdots\\0\end{array}\right)$$

4. Associativity of multiplication:

$$(ab)\left(\begin{array}{c}x_1\\x_2\\\vdots\\x_n\end{array}\right)=\left(\begin{array}{c}(ab)x_1\\(ab)x_2\\\vdots\\(ab)x_n\end{array}\right)=\left(\begin{array}{c}a(bx_1)\\a(bx_2)\\\vdots\\a(bx_n)\end{array}\right)=a\left(b\left(\begin{array}{c}x_1\\x_2\\\vdots\\x_n\end{array}\right)\right)$$

5. Distributivity:

$$a\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}\right) = \begin{pmatrix} a(x_1 + y_1) \\ a(x_2 + y_2) \\ \vdots \\ a(x_n + y_n) \end{pmatrix}$$

$$= \begin{pmatrix} ax_1 + ay_1 \\ ax_2 + ay_2 \\ \vdots \\ ax_n + ay_n \end{pmatrix} = a\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + a\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

and

$$(a+b)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} (a+b)x_1 \\ (a+b)x_2 \\ \vdots \\ (a+b)x_n \end{pmatrix}$$

$$= \begin{pmatrix} ax_1 + bx_1 \\ ax_2 + bx_2 \\ \vdots \\ ax_n + bx_n \end{pmatrix} = a\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + b\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

6. Unitarity:

$$1\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1x_1 \\ 1x_2 \\ \vdots \\ 1x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

More Examples of Vector Spaces

1. Let V be the set of n by 1 column matrices of complex numbers, let the field of scalars be $\mathbb{C}$, and define vector addition and scalar multiplication by

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} \qquad c\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}$$

just like in the previous example. The verification that this defines a vector space is the same as the real case.

2. Let V be the set of all polynomials of degree $\leq n$ with real coefficients, let the field of scalars be $\mathbb{R}$, and define vector addition and scalar multiplication in the way you remember from primary school:

$$(a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n) + (b_0 + b_1 t + b_2 t^2 + \cdots + b_n t^n)$$
$$= (a_0 + b_0) + (a_1 + b_1)t + (a_2 + b_2)t^2 + \cdots + (a_n + b_n)t^n$$

and

$$c(a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n) = ca_0 + ca_1 t + \cdots + ca_n t^n$$

Verifying that this is a vector space is straightforward.

3. Let V be the set of infinite sequences of real numbers $(x_1, x_2, x_3, \ldots)$, let the field of scalars be $\mathbb{R}$, and define vector addition and scalar multiplication by

$$(x_1, x_2, x_3, \ldots) + (y_1, y_2, y_3, \ldots) = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \ldots)$$
$$c(x_1, x_2, x_3, \ldots) = (cx_1, cx_2, cx_3, \ldots)$$

Verifying that this is a vector space is straightforward.

4. Let V be the set of all continuous functions $f : \mathbb{R} \to \mathbb{R}$, let the field of scalars be $\mathbb{R}$, and define vector addition and scalar multiplication by

$f + g$ is the continuous function defined by $(f + g)(a) = f(a) + g(a)$

cf is the continuous function defined by $(cf)(x) = cf(x)$

To show that these vector addition and scalar multiplication operations make sense, we must show that adding two elements of V yields another element in V, and multiplying an element of V by a scalar yields another element of V. This is true because of the deep fact from calculus that the sum of two continuous functions is continuous, and multiplying a continuous function by a real number gives another continuous function.

Verifying that the axioms for being a vector space are satisfied is straightforward.

5. Let V be the set of all functions $f : \mathbb{R} \to \mathbb{R}$ that satisfy the equation

$$f'' = -f$$

For example, the functions defined by $\sin(x)$ and $\cos(x)$ are both elements of V. Let the field of scalars be $\mathbb{R}$, and define vector addition and scalar multiplication in the same way as the previous example.

To show that vector addition and scalar multiplication operations make sense, we must show that adding two elements of V yields another element in V, and multiplying an element of V by a scalar yields another element of V. To check this, suppose $f_1, f_2 \in V$ and $c \in \mathbb{R}$. Then

$$(f_1 + f_2)'' = f_1'' + f_2'' = -f_1 + -f_2 = -(f_1 + f_2)$$

and

$$(cf_1)'' = c(f_1'') = c(-f_1) = -(cf_1)$$

so the operations of vector addition and scalar multiplication make sense.

Verifying that the axioms for being a vector space are satisfied is straightforward.

Frequently Asked Questions

Aren't vectors "arrows" that have a direction and magnitude?

The most frequently-encountered vector spaces have, as their set of vectors, column matrices of real numbers like the one below.

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

One popular way to illustrate this column matrix is to draw an arrow that starts at the origin in the plane and ends at the point $(2, 1)$. The "vectors are arrows" idea comes from this specific way to visualize the vectors of one specific example of a vector space. For other vector spaces, like ones where the vectors are functions or infinite sequences, you can't visualize vectors as arrows.

The definition of a vector space has *four* parts to it – a set, a field, and two operations. I still don't understand what *is* a vector space?

Think of a recipe that you find in a cookbook. A mathematician might define the concept of a *recipe* like this:

Definition: A **recipe** consists of

    1. A set I (items in this set are called **ingredients**)

    2. Instructions for turning the elements of I into the final dish.

A recipe isn't complete if you only know the set of ingredients, or if you only know the instructions: if someone asked you "does 1 kilo dry spaghetti, 1 jar pasta sauce, and 4 quarts of water constitute a recipe," you'd say "you haven't given me all the information yet – what are the instructions for turning these ingredients into a dish?" Likewise, if someone asks "Is the space of 3 by 1 column matrices with real entries a vector space?" the precise answer is "You haven't given me all the information yet – what is the scalar field and how you define vector addition and scalar multiplication?"

Of course, you'll hear mathematicians say "the vector space of 3 by 1 column matrices with real entries" all the time without specifying this additional information, because there's such an *obvious* choice of field of scalars, vector addition, and scalar multiplication, and it would be tedious if we had to say all these details each time we wanted to talk about any vector space. In the same way, if your dad hands you 1 kilo of spaghetti, 1 jar of pasta sauce, and 4 quarts of water and tells you to make dinner, you don't *need* to ask for the "instructions" part of the recipe – it's obvious what he wants you to do.

Why are the "vector addition" and "scalar multiplication" operations *part* of the definition of a vector space? I already know how to add and multiply!

Consider how many different kinds of mathematical objects can be "added" together. In primary school you learned what it means to add numbers together. Then in secondary school you learned what it means to add polynomials together. Then you learned what it means to add functions together. Then you learned what it means to add matrices together. You needed to learn a new definition of "addition" for each of these types of mathematical object.

Evidently, the word "addition" (and the plus-sign symbol) describes many different operations on many different mathematical objects. We want to make sure that there's some appropriate definition of "addition" for whatever kind of mathematical objects our vectors might be, and also that there's an appropriate definition of "scalar multiplication" for whatever kind of vectors and scalars we have. Otherwise, equations like the ones from page 1 wouldn't even make sense.