We will define an object called a vector space. This is a set of objects, called vectors, which have many of the same properties as vectors in $\mathbb{R}^n$. This will allow us to generalize the ideas that have been so successful for $\mathbb{R}^n$ and to better describe the solution sets of linear systems.

1 Definitions and Properties

Definition 1.1. A **vector space** is a nonempty set V whose elements are called **vectors**, and on which are defined two operations, called "addition" and "scalar multiplication" (multiplication by real numbers), which satisfy the properties below for all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$ and all scalars c, d:

1. The sum of $\mathbf{u}$ and $\mathbf{v}$, denoted $\mathbf{u} + \mathbf{v}$, is in V;
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$;
4. There is a zero vector $\mathbf{0}$ such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ (note: for every $\mathbf{u}$);
5. For each $\mathbf{u} \in V$, there is a vector denoted by $-\mathbf{u}$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$ (note: a different $-\mathbf{u}$ for each $\mathbf{u}$! Also note that we do not know yet that $(-1)\mathbf{u} = -\mathbf{u}$, though this will turn out to be true).
6. The scalar multiple of $\mathbf{u}$ by c, denoted $c\mathbf{u}$, is in V;
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$;
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$;
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$;
10. $1\mathbf{u} = \mathbf{u}$.

A couple properties that follow more or less immediately from the definition above:

Proposition 1.2. *There is only one zero vector in V. Also, for each $\mathbf{u}$, the vector $-\mathbf{u}$ is unique. Moreover, for each $\mathbf{u} \in V$ and each scalar c;*

$$0\mathbf{u} = \mathbf{0},$$
$$c\mathbf{0} = \mathbf{0};$$
$$-\mathbf{u} = (-1)\mathbf{u}.$$

The first thing to notice is that $\mathbb{R}^n$ meets the criteria of the definition. The definition is an attempt at distilling the properties of $\mathbb{R}^n$ that are important for what we care about. When in doubt about whether something is true for vector spaces, keep $\mathbb{R}^n$ in mind. There are a few things that are true for $\mathbb{R}^n$ that are not true in other vector spaces, but not so many.

Example 1.3. Let $\mathbb{P}_n$ denote the set of polynomials of degree at most n. The typical element of $\mathbb{P}_n$ looks like

$$a_0 + a_1 t + \ldots + a_n t^n,$$

where a_i is a real number for each i. Scalar multiplication and addition work in the obvious ways:

$$c(a_0 + a_1 t + \ldots + a_n t^n) = ca_0 + ca_1 t + \ldots + ca_n t^n,$$

$$a_0 + a_1 t + \ldots + a_n t^n + b_0 + b_1 t + \ldots + b_n t^n = (a_0 + b_0) + \ldots + (a_n + b_n) t^n.$$

Since addition and scalar multiplication of polynomials produces polynomials of the same degree, $\mathbb{P}_n$ satisfies axioms 1 and 6 in the definition of a vector space. Similarly, axioms 2 and 3 hold by properties of addition of real numbers, and axiom 4 holds with $\mathbf{0} = 0$, the polynomial with all coefficients 0. The remaining properties are also easy to verify, but we will not do this here. You should convince yourself that they hold.

Example 1.4. Let $\mathbb{C}(\mathbb{R})$ be the set of all continuous functions with domain $\mathbb{R}$ (we could consider other domains). A vector in this space is a continuous function on the reals. Addition and scalar multiplication are defined in the usual way for continuous functions. For instance,

$$c(\sin x) = c \sin x.$$

You learned in Calculus class that the sum of two continuous functions is continuous, and so is the product of continuous functions. So Axioms 1 and 6 hold. Again, the other axioms hold by properties of the real numbers. For instance, if f, g, h are continuous functions, then

$$f + (g + h) = (f + g) + h;$$

since what this means is that $f(x) + (g(x) + h(x)) = (f(x) + g(x)) + h(x)$ for all x, and this holds because $f(x), g(x), h(x)$ are just real numbers.

Example 1.5. Let $\mathbb{R}^\infty$ denote the set of all real-valued sequences with only finitely many nonzero elements. That is, a typical element of $\mathbb{R}^\infty$ is an infinite sequence

$$(a_1, a_2, a_3, \ldots)$$

with the property that there is some K such that $a_k = 0$ whenever $k \geq K$. Scalar multiplication is defined by

$$c(a_1, a_2, a_3, \ldots) = (ca_1, ca_2, ca_3, \ldots)$$

and addition by

$$(a_1, a_2, a_3, \ldots) + (b_1, b_2, b_3, \ldots) = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots)$$

Note that in both of these operations, the resulting vector still has only finitely many nonzero entries. So axioms 1 and 6 hold. The others hold by similar observations to the other two examples (check this).

2 Subspaces

Many subsets of a vector space V are vector spaces themselves. We want to define a subspace to be a vector space which is part of a larger vector space. Rather than going through all the work of checking that a subset is a vector space, though, it turns out that we only need to check a few things. Once we know these, then the rest of the vector space axioms hold via inheritance from the larger space V.

Definition 2.1. A subspace H of a vector space V is a subset of V with the properties:

1. $\mathbf{0} \in H$;
2. H is closed under vector addition–that is, $\mathbf{u} + \mathbf{v}$ is in H whenever $\mathbf{u}$ and $\mathbf{v}$ both are;
3. H is closed under multiplication by scalars. That is, for any number c, $c\mathbf{u}$ is in H whenever $\mathbf{u}$ is.

As long as H has the three properties of the above definition, then H is itself a vector space. This is not complicated to show (think about why)–basically, all the properties of addition and scalar multiplication from V carry over to H. Note that by our definition, V is a subspace of itself.

In the following examples, I will not prove that each H is a subspace. This is straightforward, but make sure you know how to do it; I'll talk about it a bit in class.

Example 2.2. Let V be any vector space. Then the set $H = \{\mathbf{0}\}$ is a subspace, called the "zero subspace". In a sense, it is the smallest possible subspace of V.

Example 2.3. Let $V = \mathbb{R}^3$ and H denote the set of all vectors in V with zero in their middle entry:

$$H = \left\{ \begin{bmatrix} a_1 \\ 0 \\ a_3 \end{bmatrix} : \quad a_1 \in \mathbb{R},\, a_3 \in \mathbb{R} \right\}.$$

Then H is a subspace of V.

Example 2.4. Let $V = \mathbb{P}_4$, the set of polynomials of degree at most 4, and let $H = \{ax^2 : \quad a \in \mathbb{R}\}$ (that is, the set of all multiples of x^2). Then H is a subspace of V.

Here is an example of a set that is not a subspace:

Example 2.5. Let $V = \mathbb{R}^2$, and let $H = \{\mathbf{x} \in \mathbb{R}^2 : \quad x_1 + x_2 = 1\}$ (here we denote the components of $\mathbf{x}$ by x_1 and x_2 as usual). This is not a subspace because it does not contain the zero vector.

2.1 Subspaces as spans

We define linear combinations for general vector spaces analogously to those in $\mathbb{R}^n$. That is, if $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\} \subseteq V$ is some set of vectors, then any vector of the form

$$a_1\mathbf{v}_1 + \ldots + a_p\mathbf{v}_p$$

is said to be a linear combination of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_p$.

Similarly, if S is a set of vectors, then $\operatorname{span} S$ is the set of all linear combinations of the vectors in S.

Example 2.6. Let V be a vector space, and let $S = \{\mathbf{v}_1, \mathbf{v}_2\}$ consist of two vectors of S. Let $H = \operatorname{span} S$. Then H is a subspace. We verify the axioms for a subspace.

$$\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2,$$

so $\mathbf{0} \in H$. Also, if $a_1\mathbf{v}_1 + a_2\mathbf{v}_2$ is in H, then so is

$$c(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) = ca_1\mathbf{v}_1 + a_2\mathbf{v}_2.$$

Finally, if $a_1\mathbf{v}_1 + a_2\mathbf{v}_2$ and $b_1\mathbf{v}_1 + b_2\mathbf{v}_2$ are two elements of H, then so is

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + b_1\mathbf{v}_1 + b_2\mathbf{v}_2 = (a_1 + b_1)\mathbf{v}_1 + (a_2 + b_2)\mathbf{v}_2.$$

In fact, the above example generalizes to arbitrary sets of vectors without much adaptation:

Theorem 2.7. *If* $\mathbf{v}_1, \ldots \mathbf{v}_p$ *are vectors in a vector space* V*, then* $\operatorname{span}\{\mathbf{v}_1, \ldots \mathbf{v}_p\}$ *is a subspace of* V*.*

Example 2.8. Let $V = \mathbb{R}^3$, and let H be the set of all vectors of the form $(0, a+b, b)$ for any real numbers a, b. An arbitrary vector of H has the form

$$\begin{bmatrix} 0 \\ a+b \\ b \end{bmatrix} = a \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

That is,

$$H = \operatorname{span}\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

Since H is the span of a set of vectors in V, it is a subspace of V.