CLARK UNIVERSITY

Vector Spaces
Math 130 Linear Algebra
D Joyce, Fall 2015

The abstract concept of vector space. There are a lot of vector spaces besides the plane $\mathbf{R}^2$, space $\mathbf{R}^3$, and higher dimensional analogues $\mathbf{R}^n$. These standard vector spaces are, perhaps, the most used vector spaces, but there are many others, so many that it makes sense to abstract the vector operations of these standard vector spaces and make a general definition. We'll do that.

Real n-space has a lot of structure, some of which we'll require a vector space to have, but some we won't require. We won't require, for instance, that a vector space have a coordinate system. We will require that a vector space have an operation of vector addition and an operation of scalar multiplication.

We won't require that a vector space have a length function that assigns a length $\|\mathbf{v}\|$ to a vector. We'll save that for later when we study inner product spaces.

Thus, for our vector spaces we abstract two of the operations, vector addition and scalar multiplication, from $\mathbf{R}^n$, but we ignore any other structure that $\mathbf{R}^n$ has.

Vector spaces over fields other than the real numbers. Most of the time our scalars will be real numbers. Sometimes, however, we'll use other fields like the complex numbers for our scalar fields.

You can have vector spaces over any field. All that's needed to have a field is a set equipped with operations of addition, subtraction, multiplication, and division with the usual properties. Thus, besides the real field $\mathbf{R}$ and the complex field $\mathbf{C}$, there's the field $\mathbf{Q}$ of rational numbers and many others.

For the time being, think of the scalar field F as being the field $\mathbf{R}$ of real numbers.

The precise definition. A *vector space* over a scalar field F is defined to be a set V, whose elements we will call *vectors*, equipped with two operations, the first called *vector addition*, which takes two vectors $\mathbf{v}$ and $\mathbf{w}$ and yields another vector, usually denoted $\mathbf{v}+\mathbf{w}$, and the second called *scalar multiplication*, which takes a scalar c and a vector $\mathbf{v}$ and returns another vector, usually denoted $c\mathbf{v}$, such that the following properties (called axioms) all hold

1. Vector addition is commutative: $\mathbf{v}+\mathbf{w} = \mathbf{w}+\mathbf{v}$ for all vectors $\mathbf{v}$ and $\mathbf{w}$;
2. Vector addition is associative: $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$ for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$;
3. There is a vector, denoted $\mathbf{0}$ and called the *zero vector*, such that $\mathbf{v}+\mathbf{0} = \mathbf{v} = \mathbf{0}+\mathbf{v}$ for each vector $\mathbf{v}$;
4. For each vector $\mathbf{v}$, there is another vector, denoted $-\mathbf{v}$ and called the *negation* of $\mathbf{v}$, such that $\mathbf{v}+(-\mathbf{v}) = \mathbf{0}$;
5. 1 acts as the identity for scalar multiplication: $1\mathbf{v} = \mathbf{v}$ for each vector $\mathbf{v}$;
6. Multiplication and scalar multiplication associate: $c(d\mathbf{v}) = (cd)\mathbf{v}$ for for all real numbers c and d and each vector $\mathbf{v}$;
7. Scalar multiplication distributes (on the left) over vector addition: $c(\mathbf{v}+\mathbf{w}) = c\mathbf{v}+c\mathbf{w}$ for each real number c and all vectors $\mathbf{v}$ and $\mathbf{w}$; and
8. Scalar multiplication distributes (on the right) over real addition: $(c+d)\mathbf{v} = c\mathbf{v}+d\mathbf{v}$ for all real numbers c and d and each vector $\mathbf{v}$.

That's a long definition, but it has to be long if we want an abstract vector space to have all the properties that $\mathbf{R}^n$ has, at least with respect to vector addition and scalar multiplication.

Lots of other properties follow from these axioms, such as $0\mathbf{v} = \mathbf{0}$, and we'll discuss those properties next time. For now, just note that subtraction can be defined in terms of addition and scalar multiplication by -1 by

$$\mathbf{v} - \mathbf{w} = \mathbf{v} + (-1)\mathbf{w}.$$

Thus, we don't need to have separate axioms to deal with subtraction.

Examples. Of course n-space, $\mathbf{R}^n$, is a vector space. $\mathbf{R}$ itself is a vector space; it's $\mathbf{R}^1$.

But what other vector spaces are there? Quite a few. Most will be finite dimensional like $\mathbf{R}^n$ is, but some will be infinite dimensional. We haven't yet defined dimension, but we will, and we'll define it in such a way that n-space has dimension n. Also, most of the other examples we'll look at have some other structure besides the structure for being vector spaces. That other structure isn't needed for the examples to be vector spaces.

Products. Given two vector spaces V and W, their product $V \times W$ is a vector space where the operations are defined coordinatewise. If $(\mathbf{v}, \mathbf{w})$ and $(\mathbf{x}, \mathbf{y})$ are two elements of $V \times W$, then their sum $(\mathbf{v}, \mathbf{w}) + (\mathbf{x}, \mathbf{y})$ is defined in terms of the addition operations on V and W by

$$(\mathbf{v}, \mathbf{w}) + (\mathbf{x}, \mathbf{y}) = (\mathbf{v} + \mathbf{x}, \mathbf{w} + \mathbf{y}),$$

and if c is a scalar, then the scalar product $c(\mathbf{v}, \mathbf{w})$ is defined by

$$c(\mathbf{v}, \mathbf{w}) = (c\mathbf{v}, c\mathbf{w}).$$

Indeed, $\mathbf{R}^2$ is just $\mathbf{R} \times \mathbf{R}$.

Products of any number of vector spaces are defined analogously. $\mathbf{R}^3$ is just $\mathbf{R} \times \mathbf{R} \times \mathbf{R}$.

Some infinite dimensional vector spaces. Consider the set of all infinite sequences $(a_1, a_2, a_3, \ldots)$ of real numbers. They form the vector space $\mathbf{R}^\infty$. Addition and scalar multiplication are performed coordinatewise just like in $\mathbf{R}^n$. In fact, $\mathbf{R}^\infty$ is just an infinite product of $\mathbf{R}$'s.

There are a couple of interesting subspaces of $\mathbf{R}^\infty$. One is where only those sequences which approach 0 are included, $\lim_{n\to\infty} a_n = 0$. These form a vector space because if two sequences approach 0, so does their sum. Also, any constant multiple of a sequence that approaches 0 also approaches 0.

Another subspace consists only of sequences that have a finite number of nonzero elements. So $(a_1, a_2, a_3, \ldots)$ is included only if all the elements are 0 from some point on.

Although they're not finite dimensional, they don't have the same infinite dimensions. A discussion of their transfinite dimensions goes beyond what we can do in this course. If you study set theory, you may see that the first two examples have a dimension which is a cardinality greater than the third.

Polynomials. Let $\mathbf{R}[x]$ be the set of all polynomials with real coefficients in the variable x. (More generally, if F is any field, let $F[x]$ be the set of all polynomials in the variable x with coefficients in the field F. Our text uses the notation $P(F)$, but that's not the usual notation.) Here, x is a formal symbol, so think of a polynomial like $x^2 - 3x + 2$ as being an expression rather than a function.

The set $\mathbf{R}[x]$ has operations of addition, subtraction, and multiplication, but not division. We don't need all those operations to treat $\mathbf{R}[x]$ as a vector space over $\mathbf{R}$. We only need to be able to add polynomials and multiply them by real numbers. Since every one of the required properties hold for these two operations (since they're just addition and multiplication), therefore $\mathbf{R}[x]$ is a vector space.

When we discuss dimension, we'll see that this vector space will not have a finite dimension. It's an infinite dimensional vector space.

Inside it there are a lot of finite dimensional subspaces.

Fix a positive integer n. Let P_n be the subset of $\mathbf{R}[x]$ consisting of all the polynomials of degree n or less. For instance, $5x^4 + 3x^2 - 7$ is a polynomial of degree 4, so it's an element of P_4, and it's an element of all P_n for $n \geq 4$, too. This set P_n is closed under the operation of addition because when we add two polynomials of degree n or less, their sum also has degree n or less. Likewise, the degree doesn't increase when we can multiply a polynomial by a real number. So P_n has the two operations required to be a vector space. Furthermore, every one of the required properties hold for these two operations (since they're just addition and multiplication), so P_n is a vector space.

Note that P_n doesn't have a coordinate system (but it isn't hard to give it one), and P_n doesn't have a length function, but it's still a vector space.

When we finally define the dimension of P_n, we'll see its dimension is $n + 1$. (Why $n + 1$? Why isn't its dimension n?)

Functions. We can generalize the last example to all functions. Consider the set of all real-valued functions with domain some fixed set S. Then, since we can add two functions, and we can multiply a function by a real constant, and all the properties hold, this is another vector space.

There are many modifications you can make to get related vector spaces. For instance, you can consider only continuous functions. Or you could consider only differentiable functions. Each one of these variants is a vector space.

Solutions to homogeneous linear differential equations with constant coefficients. Some of the most useful differential equations are of this type. For example, the exponential differential equation $y' = ky$ has the exponential functions Ae^kt as its solutions, where A is an arbitrary constant. These functions, taken together, form a one-dimensional vector space.

For another example, the second order differential equation $y'' = -y$ has the sine wave functions

$$A \cos t + B \sin t$$

as its solutions, where A and B are arbitrary constants. These functions form a two-dimensional vector space.

In general, the solutions to differential equations of this type form a vector space.

A homogeneous linear differential equations with constant coefficients of degree n is a differential equation of the form

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \cdots + a_2 y'' + a_1 y' + a_0 y = 0.$$

If two functions f and g both satisfy this differential equation, then so does $f + g$, and if c is a constant then the function cf also satisfies it. In a course in differential equations, you'll learn how to solve the equation and see that its solutions form an n-dimensional vector space.

Matrices. Before long we'll be using matrices in this course. For now, a matrix is just a rectangular arrangement of scalars. For example,

$$\begin{bmatrix} 9 & 0 & 3 \\ 4 & -1 & 2 \end{bmatrix}$$

is a matrix with two rows and three columns.

There's some standard terminology and notation for matrices. When a single symbol is used to denote an entire matrix, it's usually a capital letter, like A or B.

Usually m is used for the number of rows and n for the number of columns. Such a matrix is called an $m \times n$ matrix.

When $m = n$ we say the matrix is a *square matrix*.

When $m = 1$, the matrix only has one row, and it's called a *row vector*. Likewise, when $n = 1$, the matrix has only one column, and it's called a *column vector*.

When symbols are used for the elements (that is, entries) of a matrix, they are often doubly indexed (that is, subscripted), and the indices indicate where the entry is located. For instance, a_{34} indicates the element in the 3rd row and 4th column. Note that the first index gives the row number and the second index gives the column number. When a generic row is needed, usually i is used, and when a generic column is needed, usually j is used. So a_{ij} is the element in the ith row and jth column.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

The name of the matrix is sometimes subscripted to access its entries, for instance, for the matrix A above, A_{23} denotes the entry a_{23}.

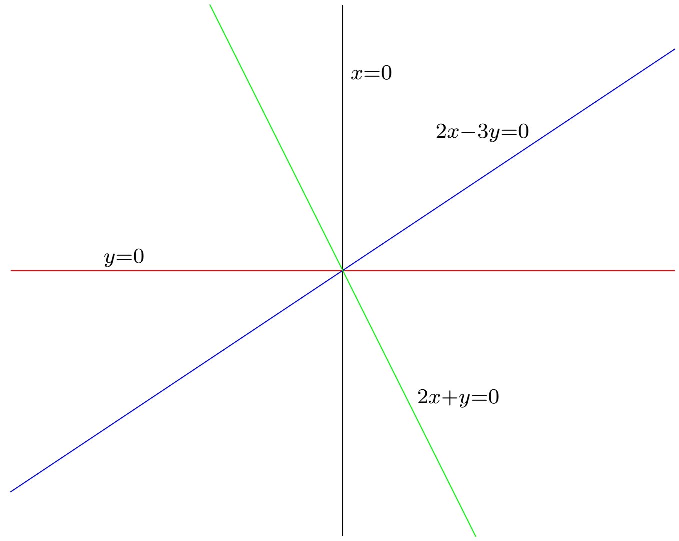
Fix m and n and consider the set $M_{mn}(\mathbf{R})$ of all $m \times n$ matrices with real entries. We'll turn $M_{mn}(\mathbf{R})$ into a vector space over $\mathbf{R}$ by defining addition and scalar multiplication coordinatewise. So the sum of two $m \times n$ matrices $A+B$ is defined by saying $(A+B)_{ij} = A_{ij}+B_{ij}$, and if c is a scalar, then cA is defined by $(cA)_{ij} = cA_{ij}$, where i denotes any row number and j denotes any column number. All the axioms for a vector space automatically follow.

In fact, the vector space $M_{mn}(\mathbf{R})$ is the same as the vector space $\mathbf{R}^{mn}$ except the scalars are arranged in a rectangular array instead of being listed in an mn-tuple. The way to express that is to say that $M_{mn}(\mathbf{R})$ and $\mathbf{R}^{mn}$ are isomorphic vector spaces. We'll have more to say about isomorphisms later.

Subspaces. We will be particularly interested in subspaces of vector spaces. A subspace of a vector space is just a subset of the vector space that is itself a vector space with the same operations. We've seen some of those above.

Let's look at some subspaces of the plane $\mathbf{R}^2$.

Consider the set of vectors (that is, points) of the form $(3t, 2t)$ in $\mathbf{R}^2$.



It's the line $2x - 3y = 0$ through the origin. It's a vector space because the sum of any two vectors in it is another vector in it, and a scalar multiple of any vector in it is another vector in it. The diagram shows four 1-dimensional subspaces of the plane. There's the line $2x - 3y = 0$, the line $2x + y = 0$, the x-axis $y = 0$, and the y-axis $x = 0$.

There's much more on subspaces to come later.

Math 130 Home Page at
`http://math.clarku.edu/~ma130/`