Vectors as algebraic objects

January 26, 2015

The simplest tensors are scalars, which are the measurable quantities of a theory, left invariant by symmetry transformations. By far the most common non-scalars are the vectors, also called rank-1 tensors. Vectors hold a distinguished position among tensors – indeed, tensors must be defined in terms of vectors. The reason for their importance is that, while tensors are those objects that transform linearly and homogeneously under a given set of transformations, we require vectors in order to define the action of the symmetry in the first place. Thus, algebraic vectors cannot be defined in terms of their transformations. In this Note, we provide an axiomatic, algebraic definition of vectors.

1 Vectors as algebraic objects

Alternatively, we can define vectors algebraically. Briefly, a vector space is defined as a set of objects, $V = \{\mathbf{v}\}$, together with a field $\mathcal{F}$ of numbers (general R or C) which form a commutative group under addition and permit scalar multiplication. The scalar multiplication must satisfy distributive laws.

More concretely, being a group under addition guarantees the following:

1. V is closed under addition. If $\mathbf{u}, \mathbf{v}$ are any two elements of V, then $\mathbf{u} + \mathbf{v}$ is also an element of V.
2. There exists an additive identity, which we call the zero vector, $\mathbf{0}$.
3. For each element $\mathbf{v}$ of V there is an additive inverse to $\mathbf{v}$. We call this element $(-\mathbf{v})$.
4. Vector addition is associative, $\mathbf{w} + (\mathbf{u} + \mathbf{v}) = (\mathbf{w} + \mathbf{u}) + \mathbf{v}$

In addition, addition is commutative, $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.

The scalar multiplication satisfies:

1. Closure: $a\mathbf{v}$ is in V whenever $\mathbf{v}$ is in V and a is in $\mathcal{F}$.
2. Scalar identity: $1\mathbf{v} = \mathbf{v}$
3. Scalar and vector zero: $0\mathbf{v} = \mathbf{0}$ for all $\mathbf{v}$ in V and $a\mathbf{0} = \mathbf{0}$ for all a in $\mathcal{F}$.
4. Distributive 1: $(a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$
5. Distributive 2: $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$
6. Associativity: $(ab)\mathbf{v} = a(b\mathbf{v})$

All of the familiar properties of vectors follow from these. An important example is the existence of a basis for any finite dimensional vectors space. We prove this in several steps as follows.

First, define linear dependence. A set of n vectors $\left\{\mathbf{v}_{(i)} \mid i = 1, \ldots, n\right\}$ is *linearly dependent* if there exist numbers $\left\{a^i \mid i = 1, \ldots, n\right\}$, not all of which are zero, such that the sum $a_i \mathbf{v}_i$ vanishes,

$$a^i \mathbf{v}_{(i)} = 0$$

As set of vectors is *linearly independent* if it is not dependent. Now suppose there exists a maximal linearly independent set of vectors. By this we mean that there exists some finite number n, such that we can find one or more linearly independent sets containing n vectors, but there do not exist any linearly independent sets containing $n+1$ vectors. Then we say that n is the *dimension* of the vector space.

In an n-dimensional vector space, and collection of n independent vectors is called a *basis*. Suppose we have a basis,

$$B = \left\{ \mathbf{v}_{(i)} \mid i = 1, \ldots, n \right\}$$

Then, since every set with $n+1$ elements is linearly dependent, the set

$$\{\mathbf{u}\} \cup B = \left\{ \mathbf{u}, \mathbf{v}_{(i)} \mid i = 1, \ldots, n \right\}$$

is dependent, where $\mathbf{u}$ is any nonzero vector in V. Therefore, there exist numbers a_i, b, not all zero, such that

$$b\mathbf{u} + a^i \mathbf{v}_{(i)} = 0$$

Now suppose $b = 0$. Then we have a linear combination of the $\mathbf{v}_i$ that vanishes, $a^i \mathbf{v}_{(i)} = 0$, contrary to our assumption that they form a basis. Therefore, b is nonzero, and we can divide by it. Adding the inverse to the sum $a^i \mathbf{v}_{(i)}$ we can write

$$\mathbf{u} = -\frac{1}{b} a^i \mathbf{v}_{(i)}$$

This shows that every vector in a finite dimensional vector space V can be written as a linear combination of the vectors in any basis. The numbers $u^i = -\frac{a^i}{b}$ are called the components of the vector $\mathbf{u}$ in the basis B.

W1. Prove that two vectors are equal if and only if their components are equal.

Notice that we have chosen to write the labels on the basis vectors as subscripts, while we write the components of a vector as superscripts. This choice is arbitrary, but leads to considerable convenience later. Therefore, we will carefully maintain these positions in what follows.

Often vector spaces are given an inner product. An inner product on a vector space is a symmetric bilinear mapping from pairs of vectors to the relevant field, $\mathcal{F}$,

$$g : V \times V \to \mathcal{F}$$

Here the Cartesian product $V \times V$ means the set of all ordered pairs of vectors, $(\mathbf{u}, \mathbf{v})$, and bilinear means that g is linear in each of its two arguments. Symmetric means that $g(\mathbf{u}, \mathbf{v}) = g(\mathbf{v}, \mathbf{u})$.

There are a number of important consequences of inner products.

Suppose we have an inner product which gives a nonnegative real number whenever the two vectors it acts on are identical:

$$g(\mathbf{v}, \mathbf{v}) = s^2 \geq 0$$

where the equal sign holds if and only if $\mathbf{v}$ is the zero vector. Then g is a *norm* or *metric* on V – it provides a notion of length for each vector. If the inner product satisfies the triangle inequality,

$$g(\mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}) \leq g(\mathbf{u}, \mathbf{u}) + g(\mathbf{v}, \mathbf{v})$$

then we can also define angles between vectors, via

$$\cos\theta = \frac{g(\mathbf{u}, \mathbf{v})}{\sqrt{g(\mathbf{u}, \mathbf{u})\, g(\mathbf{v}, \mathbf{v})}}$$

If the number s is real, but not necessarily positive, then g is called a *pseudo-norm* or a *pseudo-metric*. We use a pseudo-metric when we study relativity.

If $\left\{ \mathbf{v}_{(i)} \right\}$ is a basis, then we can write the inner product of any two vectors as

$$\begin{aligned} g(\mathbf{u}, \mathbf{v}) &= g\left(a^i \mathbf{v}_{(i)}, b^j \mathbf{v}_{(j)} \right) \\ &= a^i b^j g\left(\mathbf{v}_{(i)}, \mathbf{v}_{(j)} \right) \end{aligned}$$

so if we know how g acts on the basis vectors, we know how it acts on any pair of vectors. We can summarize this knowledge by defining the matrix

$$g_{ij} \equiv g(\mathbf{v}_i, \mathbf{v}_j)$$

Now, we can write the inner product of any two vectors as

$$g(\mathbf{u}, \mathbf{v}) = a^i g_{ij} b^j = g_{ij} a^i b^j$$

It's fine to think of this as sandwiching the metric, g_{ij}, between a row vector a^i on the left and a column vector b^j on the right. However, index notation is more powerful than the notions of row and column vectors, and in the long run it is more convenient to just note which sums are required. A great deal of computation can be accomplished without actually carrying out sums.