Introduction

These are notes for a class on multilinear algebra from Canada/USA Mathcamp 2016. They're meant to accompany that class rather than to be a complete course on the subject on their own — in particular, several of the exercises will be done in the class — but if you're reading them outside of Mathcamp and you have questions, feel free to e-mail them to me at `njmford@gmail.com`.

Throughout these notes, I will assume that the reader is familiar with some basic notions in linear algebra and ring theory. (The ring theory assumed will not be very much; just the definitions of ring, ideal, and quotient ring.) All of our vector spaces will be over an arbitrary field, which we'll call k. Except where I specifically say otherwise, it's fine to assume that $k = \mathbb{R}$ if it helps.

I've marked exercises with a star ($\star$) if I think they're especially important to understanding the rest of the packet.

1 Direct Sums and Universal Properties

We'll begin by introducing the technique we'll be using throughout these notes in a familiar setting. Recall that, given two vector spaces V and W, we can form their *direct sum* $V \oplus W$ by taking the set of ordered pairs $\{(v, w) : v \in V, w \in W\}$ and defining addition and scalar multiplication componentwise, that is,

$$(v, w) + (v', w') = (v + v', w + w')$$

and

$$\lambda(v, w) = (\lambda v, \lambda w).$$

Sometimes people will write $V \times W$ instead of $V \oplus W$, since as a set the direct sum is indeed the Cartesian product of V and W. We'll stick to the $\oplus$ notation, for reasons that we'll discuss shortly.

There is another way to think about the direct sum which is probably less familiar than this one. First, note that $V \oplus W$ comes equipped with two obvious linear maps landing inside it, one which we'll call $i_V : V \to V \oplus W$ and one which we'll call $i_W : W \to V \oplus W$. (We can map any $v \in V$ to $i_V(V) = (v, 0)$ and any $w \in W$ to $i_W(w) = (0, w)$.) We'll draw those inclusions in a diagram like this:

$$\begin{tikzcd} V \arrow[dr, "i_V"] & & W \arrow[dl, "i_W"'] \\ & V \oplus W & \end{tikzcd}$$

These maps capture something essential about $V \oplus W$: we can think of $V \oplus W$ as the vector space which contains both vectors from V and vectors from W, and therefore any linear combination of them.

But the existence of i_V and i_W don't uniquely specify $V \oplus W$; indeed, the zero vector space also has maps like that! If that's too drastic, we could also just use W: pick any linear map from V to W as i_V, and take i_W to be the identity. Something like $V \oplus W \oplus V$ would also work.

Now, there's a sense in which 0 and $V \oplus W \oplus V$ are "missing the point" of this construction. In order to have a vector space that contains both V and W you have to allow linear combinations of v's and w's, but you don't have to have anything else, and you don't have to add on any extra linear relations. What we're aiming for is way to say not just that $V \oplus W$ has these maps from V and W, but that it's somehow the *most general* way to form these maps.

This idea is expressed in the following fact:

Proposition. *Suppose there is some vector space A and linear maps $f : V \to A$ and $g : W \to A$. Then there is a unique linear map $u : V \oplus W \to A$ so that $f = u i_V$ and $g = u i_W$, as in the following diagram:*

$$\begin{tikzcd}
V \arrow[dr, "i_V"] \arrow[ddr, bend right, "f"'] & & W \arrow[dl, "i_W"'] \arrow[ddl, bend left, "g"] \\
& V \oplus W \arrow[d, dashed, "u"] & \\
& A &
\end{tikzcd}$$

Moreover, $V \oplus W$ is the unique vector space with this property. (By "unique" here, we mean that any other vector space with this property has a natural invertible linear map to $V \oplus W$.)

Whenever we draw a diagram like the one in the proposition, we'll often want to say that any two ways of getting from one vector space in the diagram to another are the equal as linear maps. In this example, there are two ways of getting from V to A: you can follow f directly, or you can apply i_V and then u. The conclusion of the theorem is, in part, that there is a u that makes these two paths the same, that is, makes $f = u i_V$. Whenever this happens, we'll say that the diagram *commutes*.

Proof. Indeed, the definition of u is forced by the conditions we're imposing on it: if $u i_V = f$, then we know that any vector of the form $(v, 0)$ has to map to $f(v)$, since we need $f(v) = u(i_V(v)) = u((v, 0))$. Similarly, any $(0, w)$ has to map to $g(w)$. But then, since u has to be linear, this forces

$$u((v, w)) = u((v, 0) + (0, w)) = u((v, 0)) + u((0, w)) = f(v) + g(w).$$

And it's straightforward to check that this choice of u is indeed linear.

Suppose B were some other vector space with the property we just proved $V \oplus W$ has. Then, since B has maps from V and W, we get a map $u : V \oplus W \to B$. By switching the roles of B and $V \oplus W$ and using the fact that B has the property, we get a map $u' : B \to V \oplus W$.

Now, consider the map $u'u : V \oplus W \to V \oplus W$. By the way we constructed u and u', we have that $u'u i_V = i_V$ and $u'u i_W = i_W$. (Check this if you don't see why!) But the identity map on $V \oplus W$ also has this property, and, by putting $V \oplus W$ in the role of A above, we see there's only one map with this property! So in fact $u'u$ is the identity, and similarly uu' is the identity. That is, $u' = u^{-1}$. So we've built our invertible linear map as desired. □

The uniqueness of $V \oplus W$ is exactly the kind of thing we were hoping for before: it shows that, in a certain sense, $V \oplus W$ really is the "best" vector space with maps from V and W, and that this uniquely picks it out among all the possible choices.

This fact about direct sums is an example of something called a *universal property*, and this is what we're going to spend the rest of these notes studying. Although they can seem very abstract at first, universal properties are a powerful tool for studying constructions like this one.

For example, notice that when proving the uniqueness of $V \oplus W$, we never actually used anything about what u actually looks like; the only fact about u that was relevant to the proof was the universal property itself. This is a typical occurrence: the universal property packs together all of the necessary information about whatever object you've constructed, and you can prove many other facts about it by using the universal property itself rather than bothering with the particulars of how the construction worked.

Exercises

1. ($\star$) There is a stronger sense in which $V \oplus W$ is the unique vector space satisfying its universal property. If B is another, with $i'_V : V \to B$ and $i'_W : W \to B$ the associated maps, prove that there is a unique invertible linear map $\phi : V \oplus W \to B$ so that $i'_V = \phi\, i_V$ and $i'_W = \phi\, i_W$. (One way to think about this is that we have uniqueness not just of the vector space $V \oplus W$ itself, but also the associated maps i_V and i_W.)

2. Suppose we're given a linear map $f : V \to W$. Consider the following universal property: we ask for a vector space K and a linear map $e : K \to V$ so that $f\,e = 0$, and so that for any other linear map $m : M \to V$ with $f\,m = 0$, there is a unique map $u : M \to K$ making the following diagram commute:

$$\begin{tikzcd} K \arrow[r, "e"] & V \arrow[r, bend left, "f"] \arrow[r, bend right, "0"'] & W \\ M \arrow[u, dashed, "u"] \arrow[ur, "m"'] & & \end{tikzcd}$$

   (When we say that this diagram commutes, we aren't saying that $f = 0$, but all other chains of arrows should be equal, including that $e\,u = m$. Note that any map composed with 0 is itself 0, so, for example, the first row commuting is what gives us that $f\,e = 0$.)

   (a) Prove that this universal property uniquely specifies K, just like we did for the direct sum.

   (b) I claim that this universal property is satisfied by an object you're already familiar with. What is it?

3. ($\star$) In this problem, we'll look at the universal property corresponding to the diagram from the previous problem with all the arrows reversed (and some of the objects renamed):

$$\begin{tikzcd} V \arrow[r, bend left, "f"] \arrow[r, bend right, "0"'] & W \arrow[r, "p"] \arrow[dr, "m"'] & Q \arrow[d, dashed, "u"] \\ & & M \end{tikzcd}$$

   We're looking for a vector space Q with a map $p : W \to Q$ so that $p\,f = 0$ and, whenever there's some $m : W \to M$ with $m\,f = 0$, there's a unique $u : Q \to M$ with $u\,p = m$. We'll call Q (together with the associated map p) the *cokernel* of f.

   (a) You can skip this problem if you're already familiar with the concept of quotient vector spaces.

Suppose you're given a vector space A and a subspace $B \subseteq A$. We're going to construct a new vector space, called the *quotient space*, written A/B. The idea is that A/B will look like A, except that every element of B has been identified with 0.

We put an equivalence relation on A as follows: say that $a \sim a'$ if there's some $b \in B$ with $a' = a + b$. Then as a set, A/B is the corresponding set of equivalence classes.

Prove that there is a consistent way to define addition and scalar multiplication on A/B, making it into a vector space. Also, show that the function $p : A \to A/B$ taking an element of A to its equivalence class is a surjective linear map and that $\ker p = B$.

(b) Suppose B is a subspace of A and $p : A \to A/B$ is the map taking each element of A to is equivalence class. Show that for any linear map $f : A \to C$ for which $f(B) = 0$, there is a unique linear map $g : A/B \to C$ so that $gp = f$. In this case, we sometimes say that f *descends to* A/B.

(c) Use the preceding exercises to construct a Q and p satisfying the universal property.

(d) Suppose V and W are finite-dimensional. What is the dimension of Q?

2 Bilinear Maps and Tensor Products

We'll move on to the universal property-based construction that will drive the rest of these notes. Again, it's based on an idea that you might have seen in a less general form, although probably not phrased in this way.

Definition. Let V, W, X be vector spaces. A *bilinear map from V and W to X* is a function p from $V \times W$ to X satisfying the following properties:

- For any $v, v' \in V$ and $w \in W$,
$$p(v + v', w) = p(v, w) + p(v', w).$$
- For any $v \in V$ and $w, w' \in W$,
$$p(v, w + w') = p(v, w) + p(v, w').$$
- For any $v \in V$, $w \in W$, and $\lambda \in k$,
$$p(\lambda v, w) = p(v, \lambda w) = \lambda p(v, w).$$

When $V = W$ and $X = k$, so that p goes from $V \times V$ to k, we call p a *bilinear form on V*.

Any inner product on V is a bilinear form, in particular (which you'll check in the exercises) the *dot product* on k^n.

There's something that makes p different from all the other functions you're likely to have encountered in linear algebra up to this point. Even though we can make $V \times W$ into a vector space — by treating it as a direct sum — this doesn't make p a linear map! We can see this even by looking at the dot product on k^2: let $v = (1, 1)$ and $w = (1, 0)$. Then p sends (v, w) to 1, but it sends $2(v, w)$ to 4.

The fact that p isn't linear is unfortunate in a lot of ways. As you've seen in your linear algebra class, the theory of linear maps is very well-developed; you can get information about them by computing their kernels, images, and ranks, and you can compose them to produce new linear maps. It would be very nice if we could do things like this for our bilinear maps.

When confronted with a problem like this, there are usually at least two possible roads to a solution. We could try to redevelop the theory of linear maps, except for bilinear maps instead. It is in fact possible to make a lot of headway with this approach, and it's often what's done in introductory linear algebra courses, usually just for bilinear forms.

But we're going to take a different approach: we'll see that there's a way to represent all the relevant information in a bilinear map with a linear map after all, so we can directly exploit our preexisting linear algebra toolbox to understand them.

This all is meant to suggest that we might like an object with the following universal property: we would like a vector space, which we'll call $V \otimes W$, together with a bilinear map p from V and W to $V \otimes W$, so that for any other bilinear map $q : V \times W \to X$, there is a unique *linear* map u so that $q = up$. The situation is represented by the following diagram:

$$\begin{tikzcd} V \times W \arrow[r, "q"] \arrow[d, "p"'] & X \\ V \otimes W \arrow[ur, dashed, "u"'] & \end{tikzcd}$$

We call $V \otimes W$ the *tensor product* of V and W. Although the definition is kind of a mouthful, the idea it's expressing is simpler than it might sound: just as $V \oplus W$ is the "best" vector space with linear maps from V and W, we can think of $V \otimes W$ has the "best" vector space with a bilinear map from V and W.

There's an important difference between our discussions of the tensor product and the direct sum so far, though: we haven't actually shown that an object satisfying this universal property exists! (We will build such an object soon.) Even so, we'll see that the universal property itself is very powerful. Just as before, it follows formally from the universal property that any object satisfying it is unique. So if you've built some vector space and you want to show that it's equal to $V \otimes W$, it's enough to show that it satisfies that universal property, since $V \otimes W$ is the unique vector space that does so. We'll see this a lot in the exercises.

Exercises

1. Verify that the dot product on k^n is a bilinear form.
2. Construct a bilinear map that isn't a bilinear form on a vector space.
3. Show that for any bilinear map $p : V \times W \to X$ and any $v \in V$, $w \in W$, $p(v, 0) = p(0, w) = 0$.
4. As we did with direct sums, show that tensor products are unique, in the following sense: if $p : V \times W \to T$ and $p' : V \times W \to T'$ are two maps satisfying the universal property defining tensor products, show that there is a unique invertible linear map $\phi : T \to T'$ so that $\phi p = p'$.
5. ($\star$) Assuming tensor products exist, use the universal property to show that $V \otimes W \cong W \otimes V$ for any V, W, and that $U \otimes (V \otimes W) \cong (U \otimes V) \otimes W$.
6. ($\star$) In this problem we're going to show that tensor products exist if one of the two vector spaces is finite-dimensional.
   (a) Prove that for any vector space V, 0 satisfies the universal property defining the tensor product $V \otimes 0$, so the tensor product exists in that case and $V \otimes 0 \cong 0$.
   (b) Similarly, prove that $V \otimes k \cong V$.
   (c) Suppose we already know that the tensor products $V \otimes A$ and $V \otimes B$ exist. Show that $V \otimes (A \oplus B)$ exists and is isomorphic to $(V \otimes A) \oplus (V \otimes B)$. (Try doing this using the universal property of the direct sum.)
   (d) Use the preceding exercises to show that the tensor product $V \otimes W$ exists if W is finite-dimensional.
   (e) If V and W are both finite-dimensional, what is $\dim(V \otimes W)$? If you're given a basis for V and W, find a basis for $V \otimes W$.

3 Constructing Tensor Products

In this short section, we'll go through the general construction of tensor products. Note that in the exercises to the previous section, we built tensor products in the case where at least one of the spaces is finite-dimensional. Still, the construction we present here has a couple advantages. In addition to not requiring any assumptions about dimension, it also self-evidently doesn't depend on any choices. (Although it might not be immediately clear, the one in the exercises requires you to choose a way of writing W as a direct sum of one-dimensional vector spaces and an element of each one, which is the same as choosing a basis for W.)

The construction we present here requires knowing about quotients of vector spaces, as will most of the rest of these notes. If you're not familiar with quotients, go through Problem 3a in Section 1 before going on.

Since the tensor product comes with a bilinear map $p : V \times W \to V \otimes W$, we'll start by building a vector space that has an element for p to hit for every $v \in V$ and $w \in W$. After that, we'll worry about making p bilinear. Specifically, let's define E to be the (gigantic!) vector space with one basis vector, which we'll write $[v, w]$, for every pair of elements $v \in V$, $w \in W$. That is, the elements of E are expressions of the form

$$\sum_{i=1}^{r} \lambda_i [v_i, w_i],$$

and addition and scalar multiplication are defined in the obvious way.

The function sending each $(v, w) \in V \times W$ to $[v, w] \in E$ is very far from being bilinear. To fix this, we'll take a quotient of E to force the required relations to hold. Specifically, let F be the subspace of E spanned by all of the following elements:

- $[v + v', w] - [v, w] - [v', w]$ for each $v, v' \in V$ and $w \in W$,
- $[v, w + w'] - [v, w] - [v, w']$ for each $v \in V$ and $w, w' \in W$,
- $[\lambda v, w] - \lambda[v, w]$ for each $v \in V$, $w \in W$, and $\lambda \in k$,
- $[v, \lambda w] - \lambda[v, w]$ for each $v \in V$, $w \in W$, and $\lambda \in k$.

Finally, we define $V \otimes W = E/F$, and define the map p to take (v, w) to the equivalence class of $[v, w]$ in this quotient. We'll write $v \otimes w$ for this element of $V \otimes W$. Elements of $V \otimes W$ of this form are called *pure tensors*.

Proposition. *This construction of $V \otimes W$ satisfies the universal property.*

Proof. First, we need to see that p is bilinear. But note that the fact that p is bilinear is exactly equivalent to the elements we placed in F being equal to 0. Since we took the quotient by F, this is indeed the case.

Now, suppose we have another bilinear map $q : V \times W \to X$. We'd like to find a linear map $u : V \otimes W \to X$ so that $up = q$ and to show that it's unique. Just as we saw when we were discussing direct sums, our choice of u is completely forced: for every element of the form $v \otimes w$, the fact that $up = q$ forces us to say that $u(v \otimes w) = q(v, w)$. But the elements of the form $v \otimes w$ span $V \otimes W$ (since the $[v, w]$'s span E), so if u exists at all there's only one possible choice.

And in fact this choice works. You'll work this out in detail in the exercises, but the sketch is as follows: we can define a linear map $\tilde{u} : E \to X$ by sending $[v, w]$ to $q(v, w)$. But in fact, $\tilde{u}(F) = 0$, so by Problem 3b in Section 1, $\tilde{u}$ descends to E/F. $\square$

Exercises

1. Find vector spaces V and W and an element of $V \otimes W$ which is not a pure tensor.
2. Why can we define the map $\widetilde{u}$ in the way we did in the last paragraph of the proof?
3. Why is $\widetilde{u}(F) = 0$?
4. Why is the fact that $\widetilde{u}$ descends to E/F enough to finish the proof?

4 Tensor Powers

Recall that one of our motivating examples for introducing tensor products in the first place was to study bilinear forms, which we can now recognize as being the same thing as linear maps $V \otimes V \to k$.

Often we insist that a bilinear form be *symmetric*, that is, that we have $f(v, v') = f(v', v)$ for all $v, v' \in V$. (For example, this is one of the conditions defining an inner product.) We can recast this condition in terms of the tensor product: it's the same as asking $v \otimes v'$ and $v' \otimes v$ to map to the same place.

Let $S \subseteq V \otimes V$ be the subspace spanned by all elements of the form

$$v \otimes v' - v' \otimes v.$$

Then what we've just shown is that a symmetric bilinear form is the same as a linear map from $V \otimes V$ to k which takes S to 0. That is, by Problem 3b, it's the same as a linear map $(V \otimes V)/S \to k$. This new vector space — $(V \otimes V)/S$ — serves the same role for symmetric bilinear forms that $V \otimes V$ serves for bilinear forms.

We can repeat this whole discussion in exactly the same way with more than two copies of V. Say we have some vector spaces $V_1, V_2, \ldots, V_k, W$. A *multilinear map* from the V_i's to W is a function $m : V_1 \times \cdots \times V_k \to W$ which is linear in each coordinate. That is, for each i and each $v_1, \ldots, v_k, v_i' \in V$, $\lambda \in k$, we have

- $m(v_1, \ldots, v_i + v_i', \ldots, v_k) = m(v_1, \ldots, v_i, \ldots, v_k) + m(v_1, \ldots, v_i', \ldots, v_k)$, and
- $m(v_1, \ldots, \lambda v_i, \ldots, v_k) = \lambda m(v_1, \ldots, v_i, \ldots, v_k)$.

As you'll verify on the exercises, such multilinear maps are described by linear maps

$$V_1 \otimes \cdots \otimes V_k \to W.$$

(Note that, since tensor products are associative by Problem 5, we don't need to specify how we are grouping the V_i's in that tensor product.) If all the V_i's are equal to the same vector space V, as they will be if we're describing *multilinear forms* on V, we'll sometimes write this tensor product as $V^{\otimes k}$. We call this the *k'th tensor power* of V.

Again, we can characterize the symmetric multilinear forms — the multilinear forms that don't depend on the order in which their arguments appear — in terms of this tensor product. Write $S \subseteq V^{\otimes k}$ for the subspace spanned by all elements of the form

$$(v_1 \otimes \cdots \otimes v_i \otimes v_{i+1} \otimes \cdots \otimes v_k) - (v_1 \otimes \cdots \otimes v_{i+1} \otimes v_i \otimes \cdots \otimes v_k).$$

Proposition. *Under the correspondence between multilinear forms on V and linear maps $V^{\otimes k} \to k$, the symmetric multilinear forms correspond to the maps that take S to 0.*

Proof. First, take some linear map $m : V^{\otimes k} \to k$, and suppose that the multilinear form it corresponds to is symmetric. Then clearly for any elements $v_1, \ldots, v_n \in V$, we have

$$m(v_1 \otimes \cdots \otimes v_i \otimes v_{i+1} \otimes \cdots \otimes v_k) = m(v_1 \otimes \cdots \otimes v_{i+1} \otimes v_i \otimes \cdots \otimes v_k),$$

since these correspond to plugging the same vectors into the multilinear form in a different order. So all the elements that span S map to 0.

Conversely, suppose that $m(S)=0$. By the preceding logic, this means that if you plug some k-tuple of vectors into the corresponding multilinear form, you get the same result after switching two neighboring arguments. But, by switching neighboring arguments repeatedly, this means that you get the same result after permuting them in any way whatsoever, so the multilinear form is indeed symmetric. □

This object — the quotient space $V^{\otimes k}/S$ — is important enough to have a name. We call in the *k'th symmetric power* of V, and we write it $\mathrm{Sym}^k V$. You'll prove more facts about $\mathrm{Sym}^k V$ in the exercises.

There is another condition besides symmetry that we might put on a multilinear form, and it will turn out to be critical to our upcoming discussion of determinants. We'll start with a special case: take $V=\mathbb{R}^2$, and consider two vectors $v,w\in V$. The *parellelogram defined by v and w* is the set

$$\{av+bw : 0\le a\le 1, 0\le b\le 1\}.$$

(See Figure 1.)

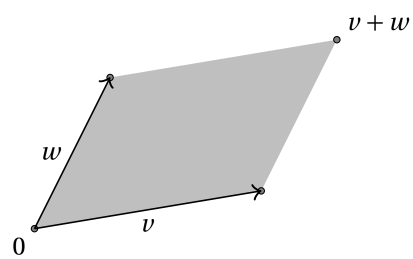

Figure 1: The parallelogram defined by two vectors.

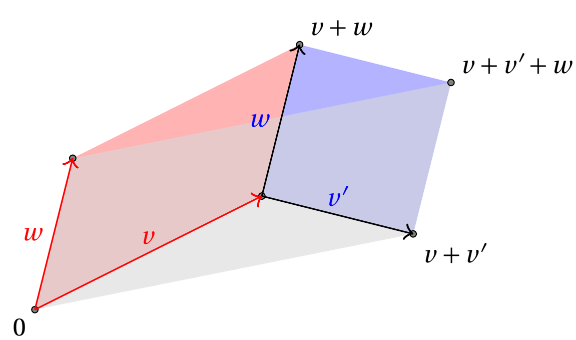

Figure 2: How areas of parallelograms behave when adding vectors: $A(v+v',w)=A(v,w)+A(v'w)$. One way to see this is to notice that the triangle cut out by the top three points is congruent to the one cut out by the bottom three points.

The key thing to notice is that the area of such a parallelogram behaves like a bilinear function of v and w. For example, writing $A(v,w)$ for the area of the parallelogram defined by v and w, Figure 2 shows why we get that

$$A(v+v',w)=A(v,w)+A(v',w).$$

You can draw a very similar picture for multiplying by a positive scalar.

The fact that A is bilinear has a perhaps surprising consequence: if A is always going to be bilinear, then by switching the roles of v and $v+v'$, we also require that

$$A(v,w) = A(v+v',w) + A(-v',w).$$

This forces us to set $A(-v',w) = -A(v',w)$; if they're not both zero, then one of these two numbers has to be negative.

In fact, there's yet another reason that A will sometimes have to take on negative values. If you plug the same vector in for both arguments to A, you should get zero: the resulting "parallelogram" is just a line segment, which has no area. That is, $A(v,v) = 0$ for each v. But if we plug a sum of two vectors into this formula, we see that

$$0 = A(v+w,v+w) = A(v+w,v) + A(v+w,w) = A(v,v) + A(w,v) + A(v,w) + A(w,w).$$

Using this fact again, we can cancel two of the four terms from the right side of this equation, giving us that

$$A(v,w) = -A(w,v).$$

Since this is true for arbitrary vectors $v, w \in V$, we see that, unless $A(v,w) = 0$, one of $A(v,w)$ or $A(w,v)$ has to be negative.

This gives us a way of thinking about what A is telling us. If we want A to be a bilinear form, it's not going to give us the area of the parallelogram, but rather a "signed area"; the absolute value $|A(v,w)|$ is the area we're after, and the sign tells us about the *orientation* of v and w. In the exercises, you'll explore this perspective further.

The condition $A(v,w) = -A(w,v)$ looks a lot like the symmetry we discussed earlier, except with an extra minus sign. When a bilinear form has this property, we call it *alternating*. Note that this is actually equivalent to saying that $A(v,v) = 0$ for each v: we've already shown one direction, and for the other, note that the alternating condition forces $A(v,v) = -A(v,v)$.

Just as we did for symmetry, it's possible to generalize this notion to multilinear forms. We'll say that a multilinear form m on a vector space V is alternating if, for every $v_1,\ldots,v_k \in V$, we have

$$m(v_1,\ldots,v_i,v_{i+1},\ldots,v_k) = -m(v_1,\ldots,v_{i+1},v_i,\ldots,v_k).$$

As in the $k=2$ case from before, it would be equivalent to say that m is zero on any k-tuple of vectors where the same vector appears twice in a row.

Define $D \subseteq V^{\otimes k}$ to be the subspace cut out by all elements of the form

$$(v_1 \otimes \cdots \otimes v_i \otimes v_{i+1} \otimes \cdots \otimes v_k) + (v_1 \otimes \cdots \otimes v_{i+1} \otimes v_i \otimes \cdots \otimes v_k).$$

For exactly the same reason as for symmetric multilinear forms, the alternating multilinear forms correspond exactly to the maps $V^{\otimes k}$ taking D to zero. So we should give a name to the quotient $(V^{\otimes k})/D$: we'll call it the *n'th exterior power of* V, and write it $\bigwedge^k V$.

There's one annoying detail that shows up with alternating multilinear forms that doesn't appear in the symmetric case. When discussing symmetry, we could simply say that it meant that reordering the arguments to the multilinear form doesn't affect the result. For alternating forms, this isn't true: reordering the arguments can either introduce a minus sign or not, depending on whether you needed to perform an even or odd number of switches. It's not immediately clear that this is even well-defined: why can't you get the same reordering in two different ways, one using an odd number or switches and another using an even number? For this and a couple other reasons, it will be helpful to be able to figure out this sign simply by looking at the reordering itself.

A *permutation of* k is a way of reordering the numbers $1, 2, \ldots, k$, that is, a bijection from the set $\{1, 2, \ldots, k\}$ to itself. Given such a permutation π, we'll say an *inversion of* π is any pair of numbers whose order is reversed by π. That is, any pair $i < j$ where $\pi(i) > \pi(j)$. If π has an odd number of inversions, we'll say that π is an *odd permutation*; otherwise, we'll call it *even.* The oddness or evenness of a permutation is usually referred to as its *sign.*

In the exercises, you'll show that if you take an odd permutation and switch two elements, you get an even permutation, and vice versa. You'll also show that this gives us the answer to our question from before: if you reorder the arguments to an alternating multilinear form, you get have to multiply by -1 if the permutation you used is odd and you get the same result if it's even.

We'll close this section by computing the dimension of the exterior power. In addition to being interesting in its own right, this computation is necessary for the discussion of determinants in the final section.

Proposition. *If* $\dim V = n$, *then* $\dim(\bigwedge^k V) = \binom{n}{k}$. *In particular, if* $k > n$, *then* $\dim(\bigwedge^k V) = 0$.

Proof. For both this proof and the rest of these notes, given vectors $v_1, \ldots, v_k \in V$, we'll write $v_1 \wedge \cdots \wedge v_k$ for the image of $v_1 \otimes \cdots \otimes v_k$ in $\bigwedge^k V$.

Pick a basis $e_1, \ldots, e_n$. Then I claim we get a basis for $\bigwedge^k V$ by taking all vectors of the form

$$e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}$$

for sequences $1 \le i_1 < i_2 < \cdots < i_k \le n$. Since there are $\binom{n}{k}$ such sequences, this will prove the result.

By applying Problem 6e from Section 2, we see that all the k-fold tensor products of the e_i's form a basis for $V^{\otimes k}$, and since $\bigwedge^k V$ is a quotient of $V^{\otimes k}$, their images span $\bigwedge^k V$. Moreover, given such an element

$$e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k},$$

you may reorder the $e_{i_j}' s$ until they appear in increasing order; doing so only involves adding and subtracting elements of D, so you get the same element of $\bigwedge^k V$ (possibly with a minus sign). Finally, if after doing this, two neighboring i_j's are equal, then the result is zero in $\bigwedge^k V$. (This is an exercise.)

So the vectors in our purported basis of $\bigwedge^k V$ span it, and it suffices to show that they're linearly independent. For a set of indices $I = \{i_1, \ldots, i_k\}$ with $i_1 < \cdots < i_k$, write

$$e_I = e_{i_1} \wedge e_{i_2} \wedge \cdots \wedge e_{i_k}.$$

Suppose we have some linear combination of these that's equal to zero, say

$$\sum_I \lambda_I e_I = 0.$$

Our goal is to show that each $\lambda_I = 0$.

Our method will be as follows: for each I, we're going to define an alternating multilinear form on V so that the corresponding linear map $\bigwedge^k V \to k$ takes e_I to 1 and all the other $e_{I'}$'s to 0. Once we have this, we can apply it to the above sum and see directly that $\lambda_I = 0$.

So it suffices to construct this alternating multilinear form. Suppose $I = \{i_1, \dots, i_k\}$, and we're given k vectors from V, say $v_1, \dots, v_k$. We'll expand each of these vectors in terms of our basis, writing

$$v_j = \alpha_{j1} e_1 + \alpha_{j2} e_2 + \cdots + \alpha_{jn} e_n.$$

One seemingly good choice would be to set

$$m_I(v_1, \dots, v_k) = \prod_{j=1}^{k} \alpha_{j i_j}.$$

This is multilinear in the v's (which you should check if you don't believe) and when you plug in basis vectors for the v's, you get 1 if they were exactly the ones corresponding to the indices in I and 0 otherwise. Sadly, m_I is not alternating, so it doesn't define a map on $\bigwedge^k V$.

Luckily, this can be fixed with the following trick. Given a permutation π, we'll write $\operatorname{sgn}(\pi)$ to mean 1 if π is even and -1 if π is odd. Given some sequence of indices I, we'll write $\pi(I)$ for the sequence you get by reordering the indices according to π. That is, the j'th index in $\pi(I)$ is the $\pi(j)$'th index in I. Then we'll define

$$a_I = \sum_{\pi} \operatorname{sgn}(\pi) m_{\pi(I)}.$$

Suppose we switch two of the arguments to a_I, say the i'th and $(i+1)$'st. For a permutation π, write $s_i\pi$ for the permutation you get by switching i and $i+1$ in π. Note that π and $s_i\pi$ always have opposite signs. Then switching these two arguments is the same as replacing $m_{\pi(I)}$ with $m_{s_i\pi(I)}$. So switching these two arguments has the effect of reordering the terms of the sum, except that each term is assigned the opposite sign. This proves that a_I is alternating.

If you plug the vectors corresponding to the indices in I into a_I, you'll get 0 except when π is the identity: all of the indices are distinct, so you get a distinct sequence whenever you reorder them. And if I' is some ordered set of indices other than I, corresponding to a different one of our basis vectors, then we can't get I' just by reordering I, so every term of a_I will be zero. Therefore, a_I has all the properties we wanted, completing the proof. □

Exercises

1. Formulate and prove a universal property expressing the relationship between the k-fold tensor product $V_1 \otimes \cdots \otimes V_k$ and multilinear maps.

2. Express the relationship between Sym^k and symmetric multilinear forms, and the relationship between $\bigwedge^k$ and alternating multilinear forms, in terms of a universal property. (You might want to consider symmetric or alternating multilinear *maps*, that is, maps from $V \times \cdots \times V$ to some other vector space W that satisfy the properties in question.)

3. In this section, we discussed how the bilinear form A we defined on $\mathbb{R}^2$ should be thought of as providing information about the "orientation" of the two vectors plugged into it. Make this idea precise; what should it mean for a pair of vectors to be positively or negatively oriented?

4. ($\star$) Prove that if you take an odd permutation and switch two elements, you get an even permutation, and vice versa.

5. ($\star$) Use this to show that if m is an alternating multilinear form on V, then reordering the arguments to m introduces a minus sign if the permutation is odd and leaves the result the same if it's even.

6. Prove that if $\dim V = n$, there are no nonzero k-fold alternating multilinear forms on V for $k > n$.

7. ($\star$) In this problem, we fill in a missing step from the computation of $\dim(\bigwedge^k V)$ from this section.

   (a) Show that if $v_1, \dots, v_k \in V$ and $v_i = v_{i+1}$ for some i, then $v_1 \wedge \cdots \wedge v_k = 0$.

   (b) Show that if $v_1, \dots, v_k \in V$, then the element $v_1 \wedge \cdots \wedge v_k$ is 0 in $\bigwedge^k V$ if and only if the v_i's are linearly dependent. [Hint: part of the work for the harder direction was already done when computing the dimension of the exterior power in this section.]

8. Mimic the computation of the dimension of an exterior power to find the dimension of a symmetric power.

5 Composition and Determinants

We'll finish these notes by applying these tools to give a satisfying presentation of determinants.

Suppose we have linear maps $f: V \to V'$ and $g: W \to W'$. We can use f and g to construct a bilinear map from V and W to $V' \otimes W'$: just send (v, w) to $f(v) \otimes g(w)$. (You should check that this is indeed bilinear.)

By the universal property of tensor products, this gives us a linear map $V \otimes W \to V' \otimes W'$. We'll write $f \otimes g$ for this map, and its construction will turn out to be very powerful.

The first thing to notice is that this tensor product construction respects composition of linear maps. That is, if $f: V \to V$, $f': V' \to V''$, $g: W \to W'$, $g': W' \to W''$ are linear maps, then

$$(f' \otimes g') \circ (f \otimes g) = (f'f) \otimes (g'g).$$

In fact, this is immediate: they both do the same thing to pure tensors, and the pure tensors span $V \otimes W$. (Alternatively, they're equal because they both correspond to the same bilinear map from V and W to $V'' \otimes W''$.)

Given a linear map $f: V \to W$, we can apply this repeatedly to get a map from $V^{\otimes k}$ to $W^{\otimes k}$. We'll call this map $f^{\otimes k}$. On pure tensors, it acts by

$$f^{\otimes n}(v_1 \otimes \cdots \otimes v_n) = f(v_1) \otimes \cdots \otimes f(v_n).$$

If $g: W \to X$ is another linear map, the fact that tensor products of linear maps respect composition also tells us that

$$g^{\otimes k} \circ f^{\otimes k} = (gf)^{\otimes k}.$$

(It's also straightforward to check this directly.)

Recall that we defined $\mathrm{Sym}^k V$ and $\bigwedge^k V$ as quotients of $V^{\otimes k}$ by subspaces we called S and D respectively. Given some linear map $f: V \to W$, write S_V and D_V for the S and D subspaces of $V^{\otimes k}$, and write S_W and D_W for the ones in $W^{\otimes k}$. You'll show in the exercises that $f^{\otimes k}$ takes S_V into S_W and D_V into D_W.

If we compose $f^{\otimes k}$ with the quotient map $W^{\otimes k} \to (W^{\otimes k})/S_W = \mathrm{Sym}^k W$, this tells us that the resulting map takes S_V to zero, meaning that it gives us a well-defined linear map $\mathrm{Sym}^k V \to \mathrm{Sym}^k W$. We'll call this map $\mathrm{Sym}^k f$. Similarly, we get a linear map $\bigwedge^k f: \bigwedge^k V \to \bigwedge^k W$. Yet again, it follows immediately from the preceding discussion that Sym^n and $\bigwedge^n$ respect composition of maps in the same way.

Why have we gone through all this? It turns out that looking at $\bigwedge^k f$ for various values of k tells you a lot about the map f itself:

Proposition. *Let $f: V \to W$ be a linear map. Then f has rank $< k$ if and only if $\bigwedge^k f = 0$.*

Proof. This follows almost immediately from Problem 7b in Section 4. Suppose $v_1, \ldots, v_k$ are vectors in V. Then that problem says that the element

$$f(v_1) \wedge \cdots \wedge f(v_k) \in \bigwedge^k W$$

is nonzero if and only if the $f(v_i)$'s are linearly independent. If $\mathrm{rk} f \geq k$, there will be some some set of vectors that makes this happen, and if $\mathrm{rk} f < k$ there won't. $\square$

We'll see later and in the exercises that there is a way to compute a matrix for $\bigwedge^k f$ from a matrix for f, which gives a way to determine the rank of a matrix just from its entries.

This construction is especially interesting in the case where $V = W$ and $k = \dim V$. Writing $n = \dim V$, in this case we have $\dim \bigwedge^n V = \binom{n}{n} = 1$. So $\bigwedge^n f$ is a map from a one-dimensional vector space to itself. Such a map is completely specified by single number: there is a unique d so that, for any nonzero $x \in \bigwedge^n V$, we have $(\bigwedge^n f)x = dx$. We call this number the *determinant* of f, and we write it $\det f$.

The proposition we just proved immediately tells us that $\det f \neq 0$ if and only if f has full rank. Moreover, the fact that $\bigwedge^n$ respects composition immediately gives us that, for two linear maps $f, g : V \to V$, we get

$$\det(gf) = \det g \cdot \det f,$$

since composing linear maps from a one-dimensional space to itself is the same as multiplying the corresponding numbers.

Suppose we're given a matrix for a linear map $f : V \to V$. That is, we've picked a basis $e_1, \ldots, e_n$ for V, and we've found the numbers a_{ij} so that

$$f(e_j) = \sum_{i=1}^{n} a_{ij} e_i.$$

How can we compute $\det f$ from this matrix? It's enough to see what $\bigwedge^n f$ does to one vector — it multiplies every vector in $\bigwedge^n V$ by the scalar $\det f$ — so let's plug in $e_1 \wedge \cdots \wedge e_n$. Writing e for this vector, we have that

$$\bigwedge^n f(e) = f(e_1) \wedge \cdots \wedge f(e_n) = \left(\sum_{i=1}^{n} a_{i1} e_i \right) \wedge \cdots \wedge \left(\sum_{i=1}^{n} a_{in} e_i \right).$$

By multilinearity, we can expand the right side of this equation by looking at all the ways of combining one term from each sum. If any two terms chosen in this way contain the same e_i, the result will be zero, so it's enough to just look at the ways of choosing distinct e_i's. This is the same as a permutation, so we get

$$\sum_{\pi} \left(\prod_{i=1}^{n} a_{i\pi(i)} \right) (e_{\pi(1)} \wedge \cdots \wedge e_{\pi(n)}).$$

But this last combination of e_i's is just $\operatorname{sgn} \pi \cdot e_1 \wedge \cdots \wedge e_n$, that is, $\operatorname{sgn} \pi \cdot e$. So, putting this all together, we get that

$$\det f = \sum_{\pi} \left((\operatorname{sgn} \pi) \prod_{i=1}^{n} a_{i\pi(i)} \right).$$

It's worth closing out this discussion by taking stock of how remarkable this is. If you've seen a different presentation of determinants before, it's probably quite surprising that once the machinery of tensor products and exterior powers has been developed, it's actually not nearly as difficult as one might expect to prove a bunch of facts about determinants: that they multiply when you multiply matrices, that they can tell you the rank of a matrix, and that they don't depend on the choice of basis.

In fact, you may not have even noticed that we proved independence of the choice of basis, but we did: nothing we did depended on any choice of basis for V. The expression we deduced for $\det f$ in terms of the entries of a matrix is often taken as the definition, but from this perspective all the facts I just listed are much more painful to prove.

This story is one of my favorite demonstrations of the power of mathematical abstraction. By being careful about when it's necessary to get into the details and when it's possibly to prove the thing you want formally without thinking too hard, you get a presentation in which the definitions feel well-motivated rather than arbitrary. We didn't exactly avoid doing the computations required by the ordinary approach — in a sense, all of it is packed up in the construction of a_I from the computation of the dimension of $\bigwedge^k V$ in Section 4 — but we were able to keep it where it was needed so that the rest of the story could be told more cleanly.

Exercises

1. Given linear maps $f : V \to V'$ and $g : W \to W'$, check that the map from $V \times W$ to $V' \otimes W'$ sending (v, w) to $f(v) \otimes g(w)$ is bilinear.

2. ($\star$) If $f : V \to W$ is a linear map and $S_V, D_V \subseteq V^{\otimes k}$ and $S_W, D_W \subseteq W^{\otimes k}$ are the subspaces defined in the section, prove that $f^{\otimes k}$ takes S_V into S_W and takes D_V into D_W.

3. Generalize the computation of $\det f$ from this section to figure out the entries of the matrix for $\bigwedge^k f$ when $k < n$. Use this to prove that a matrix T has rank $< k$ if and only if each $k \times k$ submatrix — that is, the matrix you get by selecting k different rows and k different columns from T — has determinant 0.