LECTURE 16: TENSOR PRODUCTS

We address an aesthetic concern raised by bilinear forms: the source of a bilinear function is not a vector space. When doing linear algebra, we want to be able to bring all of that machinery to bear. That means we need to consider only vector spaces and linear maps. Our constructions today will answer that point. We begin with a generalization of a bilinear form.

Definition. A function $f\colon U \times V \to W$ is <u>bilinear</u> if

$$f(a\vec{u}_1 + b\vec{u}_2, \vec{v}) = af(\vec{u}_1, \vec{v}) + bf(\vec{u}_2, \vec{v}) \text{ and}$$
$$f(\vec{u}, a\vec{v}_1 + b\vec{v}_2) = af(\vec{u}, \vec{v}_1) + bf(\vec{u}, \vec{v}_2).$$

Let $Bilin(U, V; W)$ denote the set of all bilinear functions $U \times V \to W$. This is clearly a vector space with the usual operations.

Thus we have for each fixed $\vec{u}$ a linear map $f_{\vec{u}} = f(\vec{u}, -)\colon V \to W$, and for each $\vec{v} \in V$, a linear map $f_{\vec{v}} = f(-, \vec{v})\colon U \to W$. This already is a kind of generalization of our Reisz map we saw last time, and we will return to this later.

The tensor product is the unique (up to isomorphism) vector space such that for any W

$$\mathcal{L}(U \otimes V, W) = Bilin(U, V; W).$$

In some sense, this is our defining principal. We'll give two different constructions, and then show they are the same. We first review somewhat a "universal" property. We've actually already seen this notion twice before:

(1) The free vector space on a set X, $\mathbb{F}\{X\}$, had the defining property that to give a map of sets $X \to V$ was the same as to give a linear transformation $\mathbb{F}\{X\} \to V$, and that we recovered any map of sets by composing the linear map with a canonical map of sets $X \to \mathbb{F}\{X\}$ that was the inclusion of a basis.

(2) The quotient space V/W had the defining property that given a map $L\colon V \to U$ such that $W \subset \ker(L)$, we can find a unique map $V/W \to U$ such that the composite $V \to V/W \to U$ is L.

In both of these cases, we had two pieces of data: the vector space we wanted and a distinguished map to our vector space that factors any map of a certain kind (in the first example, it was just maps of sets, and in the second, it was linear maps with a kernel bigger than W). We can do the same for the tensor product by specifying its universal property.

Definition. Let U and V be vector spaces. Then the tensor product is vector space $U \otimes V$ such that

(1) There is a bilinear map $i\colon U \times V \to U \otimes V$ and
(2) Given any bilinear map $f\colon U \times V \to W$, there is a unique linear map $L_f\colon U \otimes V \to W$ such that $L_f \circ i = f$.

Of course, this definition might be defining nothing. On the other hand, if we have two vector spaces that satisfy the two conditions of the definition, then they are isomorphic (and if we work slightly harder and ensure that the map $U\times V \to U\otimes V$ is remembered, then the isomorphism is also unique).

Proposition. *If $(U\otimes V)_1$ and $(U\otimes V)_2$, together with their bilinear maps i_1 and i_2 respectively, both satisfy the conditions of the definition, then there is a unique isomorphism that takes i_1 to i_2 and vice-versa.*

Proof. Setting W in the definition equal to $(U\otimes V)_2$, we have a unique map

$$L_{i_2} : (U\otimes V)_1 \to (U\otimes V)_2$$

such that $L_{i_2}\circ i_1 = i_2$. On the other hand, if we swap the roles of $(U\otimes V)_1$ and $(U\otimes V)_2$, then we get a map

$$L_{i_1}\colon (U\otimes V)_2 \to (U\otimes V)_1.$$

Now the composite $L_2 = L_{i_2}\circ L_{i_1}\colon (U\otimes V)_2 \to (U\otimes V)_2$ is a linear map such that $L_2\circ i_2 = i_2$. Thus if we apply our defining property when $W = (U\otimes V)_2$ and factoring through $(U\otimes V)_2$, then we see that L_2 is the unique map such that $L_2\circ i_2 = i_2$ (this was the uniqueness part of the definition). Of course, the identity map is also such a map, so by uniqueness, L_2 is the identity map. The same argument shows that $L_1 = L_{i_1}\circ L_{i_2}$ is the identity, and we have constructed inverse isomorphisms. $\square$

We'll now focus on two constructions of very different flavors: an almost unuseable one that has very good formal properties and an unnaturally one (with a choice [boo!] of basis) that is easier to compute with.

Construction #1. In this construction, we will produce something that obviously satisfies the conditions of the definition. We begin by thinking again about what we want. We want a vector space $U\otimes V$, together with a bilinear map $i\colon U\times V\to U\otimes V$, such that for any W,

$$\mathcal{L}(U\otimes V, W) \cong Bilin(U,V;W)$$

and the isomorphism from $\mathcal{L}$ is given by $L\mapsto L\circ i$. We will build $U\otimes V$ and i by first considering all functions (not just bilinear ones). The free vector space $\mathbb{F}\{-\}$ has the desired universal property here:

$$\mathcal{L}(\mathbb{F}\{U\times V\}, W) = Map(U\times V, W),$$

and the isomorphism is again given by composing with the inclusion $i\colon U\times V\to \mathbb{F}\{U\times V\}$. Every bilinear function is in particular a map of sets $U\times V\to W$, so we need to simply impose restrictions that make our functions bilinear. We'll consider linearity in the first factor, as linearity in the second is an obvious extension.

A function $f\colon U\times V\to W$ is linear iff

$$f(a\vec{u}_1 + b\vec{u}_2, \vec{v}) = af(\vec{u}_1,\vec{v}) + bf(\vec{u}_2,\vec{v}), \text{or equivalently}$$

$$f(a\vec{u}_1 + b\vec{u}_2, \vec{v}) - af(\vec{u}_1,\vec{v}) - bf(\vec{u}_2,\vec{v}) = 0.$$

This is a linear combination of values of f on elements of $U\times V$, so we can record the same information with L_f, the linear transformation $\mathbb{F}\{U\times V\}\to W$ encoding f. Our identification now becomes: f is linear in the first factor iff

$$L_f\big((a\vec{u}_1 + b\vec{u}_2,\vec{v}) - a(\vec{u}_1,\vec{v}) - b(\vec{u}_2,\vec{v})\big) = 0,$$

since L_f is linear and $L_f((\vec{u},\vec{v})) = f(\vec{u},\vec{v})$.

This is a much nicer condition. We consider now the subspace I of $\mathbb{F}\{U \times V\}$ spanned by all possible vectors of this form

$$I = \big\langle (a\vec{u}_1 + b\vec{u}_2, \vec{v}) - a(\vec{u}_1, \vec{v}) - b(\vec{u}_2, \vec{v}), (\vec{u}, a\vec{v}_1 + b\vec{v}_2) - a(\vec{u}, \vec{v}_1) - b(\vec{u}, \vec{v}_2) \big\rangle.$$

Thus f is bilinear iff $I \subset \ker(L_f)$. The universal property of the quotient tells us then that f is bilinear iff L_f extends uniquely over $\mathbb{F}\{U \times V\}/I$. Here is a diagram that records much of this

$$\begin{tikzcd}
U \times V \arrow[r, "i"] \arrow[d, "f"'] & \mathbb{F}\{U \times V\} \arrow[r, "\pi_I"] \arrow[dl, "L_f"'] & \mathbb{F}\{U \times V\}/I \, . \arrow[dll, "\tilde{L}_f"] \\
W & &
\end{tikzcd}$$

The first diagonal map exists and is unique no matter what f is. The second diagonal map exists and is unique whenever f is bilinear. Thus we define the tensor product to be $\mathbb{F}\{U \times V\}/I$, and the structure map is the composite $\pi_I \circ i$. We must check that this is bilinear. This is easy, however. The element $(a\vec{u}_1 + b\vec{u}_2, \vec{v}) \in U \times V$ maps under i to the basis vector by the same name. However, since

$$(a\vec{u}_1 + b\vec{u}_2, \vec{v}) - a(\vec{u}_1, \vec{v}) - b(\vec{u}_2, \vec{v}) \in I,$$

we know that

$$\pi_I\big((a\vec{u}_1 + b\vec{u}_2, \vec{v})\big) = a\pi_I(\vec{u}_1, \vec{v}) + b\pi_I(\vec{u}_2, \vec{v}),$$

and thus $\pi_I \circ i$ is bilinear.

That this construction gives us the right universal property is immediate. There are a few draw-backs:

(1) If $|\mathbb{F}|$ is infinite, then $\mathbb{F}\{U \times V\}$ is HUGE (the dimension is $|\mathbb{F}|^2 \cdot \dim U \cdot \dim V$). Additionally, I is also HUGE (and it has essentially the same dimension).
(2) It's not immediately clear that $U \otimes V$ is not the zero space. In other words, that $I \subsetneq \mathbb{F}\{U \times V\}$.

Additionally, at the end of the day, we don't normally specific a linear transformation by specifying where EVERY element goes. We just work with a basis. That brings us to construction 2.

Construction 2. Now assume that U has a basis $\{\vec{u}_1, \dots, \vec{u}_n\}$ and V has a basis $\{\vec{v}_1, \dots, \vec{v}_m\}$ (and here infinite bases work no differently). We start with an easy proposition.

Proposition. *Let f be a bilinear function $U \times V \to W$ then f is uniquely determined by its values on $(\vec{u}_i, \vec{v}_j)$, and any choice of values on these points determines a bilinear function.*

Proof. Let $\vec{u} = \sum a_i \vec{u}_i$ and let $\vec{v} = \sum b_j \vec{v}_j$. Then by linearity in the first factor

$$f(\vec{u}, \vec{v}) = \sum_i a_i f(\vec{u}_i, \vec{v}).$$

By linearity in the second factor, this is

$$= \sum_i a_i \sum_j b_j f(\vec{u}_i, \vec{v}_j) = \sum_{i,j} a_i b_j f(\vec{u}_i, \vec{v}_j).$$

Thus specifying the values on $(\vec{u}_i, \vec{v}_j)$ defines f. Since these form a basis, there are no additional restrictions and hence any collection of values gives a bilinear function. $\square$

This gives us our second definition:

$$U \otimes V = \mathbb{F}\{(\vec{u}_i, \vec{v}_j)\}.$$

From the above proposition, we know that bilinear functions $U \times V$ are in 1-1 correspondence with linear maps $\mathbb{F}\{(\vec{u}_i, \vec{v}_j)\}$, and therefore we have almost verified the universal property. The only remaining point is to produce a bilinear map $U \times V \to U \otimes V$. However, using our 1-1 correspondence, we just have to produce a linear transformation $U \otimes V \to U \otimes V$. The identity works very nicely, and the corresponding bilinear map is the last piece of structure for the universal property.

This gives us a very nice, easy consequence.

Proposition. *The dimension of* $U \otimes V$ *is* $\dim(U) \cdot \dim(V)$.

Moving On. We already showed that any two things that satisfy the universal property are uniquely isomorphic. Thus our two definitions of $\otimes$ are the same, and we will feel free to use both depending on context. In all cases, we will denote the image of $(\vec{u}, \vec{v})$ in $U \otimes V$ by $\vec{u} \otimes \vec{v}$.

We turn now to maps. Let's consider $S\colon U \to U'$ and $T\colon V \to V'$. Then we want to produce a new linear map $S \otimes T\colon U \otimes V \to U' \otimes V'$. To produce a linear map like this is the same as producing a bilinear map $U \times V \to U' \otimes V'$. This, however, is essentially formal (one of the advantages of a universal property).

Proposition. *If* $i\colon U' \times V' \to U' \otimes V'$ *is the bilinear structure map, then*

$$i \circ (S \times T)\colon U \times V \to U' \otimes V'$$

is bilinear. Here $S \times T\colon U \times V \to U' \times V'$ *is given by* $(S \times T)(\vec{u}, \vec{v}) = (S(\vec{u}), T(\vec{v}))$.

Proof. We first check linearity in the first factor. Our map is a composite, so we break it into two pieces and analyze them separately.

$$(S \times T)(a\vec{u}_1 + b\vec{u}_2, \vec{v}) = \big(S(a\vec{u}_1 + b\vec{u}_2), T(\vec{v})\big) = \big(aS(\vec{u}_1) + bS(\vec{u}_2), T(\vec{v})\big).$$

The map i is bilinear, so we know that

$$i\Big(\big(aS(\vec{u}_1) + bS(\vec{u}_2), T(\vec{v})\big)\Big) = ai\big(S(\vec{u}_1), T(\vec{v})\big) + bi\big(S(\vec{u}_2), T(\vec{v})\big)$$

This is exactly what expresses linearity in the first factor. The second factor is handled identically. $\square$

Thus we get a linear map $S \otimes T\colon U \otimes V \to U' \otimes V'$ by

$$(S \otimes T)(\vec{u} \otimes \vec{v}) = S(\vec{u}) \otimes T(\vec{v})$$

and extending by linearity.

Proposition. *The assignment* $(S, T) \mapsto S \otimes T$ *is a bilinear map*

$$\mathcal{L}(U, U') \times \mathcal{L}(V, V') \to \mathcal{L}(U \otimes V, U' \otimes V'),$$

and thus descends to a linear map

$$\mathcal{L}(U, U') \otimes \mathcal{L}(V, V') \to \mathcal{L}(U \otimes V, U' \otimes V').$$

When all spaces are finite dimensional, this is an isomorphism.

We can drop the dimensionality restrictions, but we won't explore that here.

Proof. This kind of argument is the same as the ones we have been giving. We check directly that $S \otimes T$ is linear in each factor, and then we conclude we have a map.

For this, consider $(aS_1 + bS_2) \otimes T$. Equality of functions means they have the same values at each point, and so it suffices to show that on a generic test point $(\vec{u} \otimes \vec{v})$,

$$(aS_1 + bS_2) \otimes T(\vec{u} \otimes \vec{v}) = aS_1 \otimes T(\vec{u} \otimes \vec{v}) + bS_2 \otimes T(\vec{u} \otimes \vec{v}).$$

We write out the left most term:

$$(aS_1 + bS_2 \otimes T)(\vec{u} \otimes \vec{v}) = \big((aS_1 + bS_2)(\vec{u})\big) \otimes T(\vec{v}) = \big(aS_1(\vec{u}) + bS_2(\vec{u})\big) \otimes T(\vec{v}).$$

Since the tensor product in $U' \otimes V'$ (where the last term in the equality lives) is bilinear, we conclude that this is the right-hand side of the desired equality.

Showing that this is an isomorphism is slightly tricker. Careful details will be left as an exercise. For now, note that if we choose a basis for all of our vector spaces, then we get a basis for the linear transformations and for the tensor product by mirroring the "dual basis" construction. With this basis, it is immediate that basis vectors map to basis vectors in a bijective way, and hence that the map is an isomorphism. □