TENSOR PRODUCTS AND MULTILINEAR ALGEBRA

JOE HARRIS (WITH MINOR EDITING BY DENIS AUROUX)

1. Three definitions of the tensor product

Let's start with a pair of vector spaces V, W over a field K. (For the most part, the field K can be arbitrary, though later on when we talk about symmetric products we may want to restrict ourselves to characteristic 0.) There are typically three ways to define the *tensor product* $V \otimes W$:

1. The most concrete: we say that if V and W have bases $\{v_1, \dots, v_m\}$ and $\{w_1, \dots, w_n\}$, then $V \otimes W$ is the vector space with basis $\{v_i \otimes w_j\}_{1 \le i \le m;\, 1 \le j \le n}$. (An analogous definition works even when V and W may be infinite-dimensional, though it is dependent on the existence of bases.)

Note that if $v = a_1 v_1 + \cdots + a_m v_m \in V$ and $w = b_1 w_1 + \cdots + b_n w_n \in W$ are any two vectors, we can define an element $v \otimes w \in V \otimes W$ in a natural way: we set

$$v \otimes w = \sum_{i,j} a_i b_j (v_i \otimes w_j).$$

The elements $v \otimes w \in V \otimes W$ are called *pure* tensors, or *rank 1* tensors; by definition they span $V \otimes W$, though in general most tensors will not be pure. Indeed, we define the *rank* of an element $\eta \in V \otimes W$ to be the minimal number of pure tensors needed to express η as a sum of pure tensors.

This definition has two virtues: it's fairly concrete, and you see immediately that if V and W have dimensions m and n, then $V \otimes W$ is a vector space of dimension mn. (Warning: this will not be at all clear from either of the other two definitions below.) It has, however, two crucial defects: first, it involves making unnecessary choices (of bases); and (not unrelated) it isn't clear from the definition that tensor product defines a *functor*

$$\otimes : (Vect_K) \times (Vect_K) \to (Vect_K)$$

from the product of the category of vector spaces over K with itself, to itself. For a pure mathematician, this is a fairly ugly definition.

2. We can avoid the need to choose bases for our vector spaces V, W by a useful construction. To start with, in Step 1 we take U to be the vector space with basis $\{v \otimes w\}_{v \in V;\, w \in W}$; that is U is the space of all finite linear combinations of the symbols $v \otimes w$, where v and w range over all elements of V and W. This is a *huge* vector space—not even of countable dimension in general—but thankfully we won't have to deal with it for long, because:

Date: This version: October 2020.

In Step 2, we let $U_0 \subset U$ be the subspace spanned by elements of the form

$$(\lambda v) \otimes w - \lambda(v \otimes w); \quad v \otimes (\lambda w) - \lambda(v \otimes w)$$
$$(v + v') \otimes w - v \otimes w - v' \otimes w; \quad \text{and} \quad v \otimes (w + w') - v \otimes w - v \otimes w'.$$

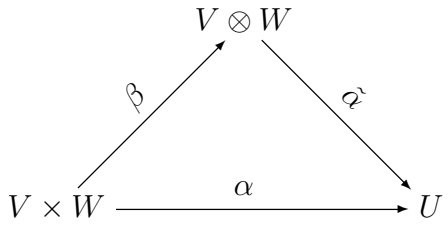
Finally, we can define the tensor product $V \otimes W$ to be just the quotient vector space U/U_0.

This is a convenient construction to have in your toolbag when you're crafting new vector spaces, and in the present circumstance it neatly avoids the need to choose bases. Correspondingly, it makes clear that tensor product is indeed a functor $\otimes : (Vect_K) \times (Vect_K) \to (Vect_K)$: if, for example, $\phi : V \to V'$ is any linear map, it clearly induces a linear map $V \otimes W \to V' \otimes W$ sending $v \otimes w$ to $\phi(v) \otimes w$.

It is not, however, without flaws; for one, it's far from clear from this definition that if V and W have dimensions m and n, then $V \otimes W$ is finite-dimensional, let alone that it's a vector space of dimension mn. And you do have to deal, however briefly, with vector spaces of uncountable dimension.

3. The third characterization of tensor products is the most abstract and difficult to grasp when you first encounter it (though hopefully less so than before we started talking about categories and functors). It is also by far the most common definition. This is not as perverse as it sounds; the fact is, this definition really does capture the essential aspect of the tensor product $V \otimes W$, which is that it's the universal object through which all bilinear maps from $V \times W$ factor.

To set this up, first some (not very new) language: if V and W are vector spaces, a *bilinear map* $\phi : V \times W \to U$ from $V \times W$ to another vector space U is simply a map linear in each variable separately (e.g., such that $\phi(\lambda v, w) = \lambda\phi(v, w)$, etc.). We then define the tensor product $V \otimes W$ to be a vector space (also denoted $V \otimes W$, unfortunately), together with a bilinear map $\beta : V \times W \to V \otimes W$, such that for any vector space U and bilinear map $\alpha : V \times W \to U$, there exists a unique map $\tilde{\alpha} : V \otimes W \to U$ such that $\alpha = \tilde{\alpha} \circ \beta$.



Note that any two such objects are the same up to isomorphism: if $(Z, \beta : V \times W \to Z)$ and $(Z', \beta' : V \times W \to Z')$ are any two such objects, by the basic condition satisfied by the two we see that β' factors through β and that β factors through β', and by uniqueness, that the induced maps $Z \to Z'$ and $Z' \to Z$ are inverse to each other. Thus the condition given in the third definition uniquely characterizes the tensor product; it's just not a priori clear that such a thing exists. So most mathematicians adopt the third definition as the definition of the tensor product, and view the first two as constructions of the tensor product.

2. A few basic identities

We'll start with a couple simple ones:

$$\begin{aligned}(V \otimes W)^* &= V^* \otimes W^*; \\ (U \oplus V) \otimes W &= (U \otimes W) \oplus (V \otimes W).\end{aligned}$$

(You should verify these yourself, to get some practice applying the "universal" definition.)

Next, we have a really crucial one, which we'll explain after we state it:

$$\mathrm{Hom}(V, W) = V^* \otimes W$$

One way to see this is to consider the ordinary product $V^* \times W$; that is, pairs of the form (ℓ, w) with $\ell \in V^*$ a linear functional on V and $w \in W$ any vector in W. We can associate to the pair (ℓ, w) the linear map $\phi : V \to W$ sending an arbitrary vector $v \in V$ to $\phi(v) = \ell(v) \cdot w$. This gives us a bilinear map $V^* \times W \to \mathrm{Hom}(V, W)$, which by our third characterization of tensor product factors through $V^* \otimes W$ to give an isomorphism $V^* \otimes W \to \mathrm{Hom}(V, W)$. (In other words, the rank 1 tensors in $V^* \otimes W$ correspond to maps $\phi : V \to W$ of rank 1, not coincidentally.)

Note one other aspect of this correspondence: we have

$$\mathrm{Hom}(V, W) = V^* \otimes W = (W^*)^* \otimes V^* = \mathrm{Hom}(W^*, V^*),$$

and the composition of these identifications gives us the *transpose map* $\mathrm{Hom}(V, W) \to \mathrm{Hom}(W^*, V^*)$.

Finally, one more identification that will lead us to the next section. As we've seen, if V is a vector space then the set B of bilinear forms $b : V \times V \to K$ also has the structure of a vector space; we can now identify it as

$$B = V^* \otimes V^*.$$

The tensor algebra. Let V be a single vector space. We will denote by $V^{\otimes n}$ the n-fold tensor product $V \otimes \cdots \otimes V$. (By convention, $V^{\otimes 0}$ is a copy of the ground field K.) We have natural bilinear maps $V^{\otimes m} \times V^{\otimes n} \to V^{\otimes (m+n)}$, and we can use these to give the direct sum

$$T(V) := \bigoplus_{n=0}^{\infty} V^{\otimes n}$$

the structure of a ring (not commutative, unless $\dim V = 1$). $T(V)$ is called the *tensor algebra* of V.

3. Symmetric powers

Throughout this section, the ground field K will be assumed to be of characteristic 0, as the theory of symmetric powers in positive characteristic requires extra care.

Let's revisit the identification $B = V^* \otimes V^*$. As we saw, the space B of bilinear forms on V naturally decomposes as a direct sum $B = B_{symm} \oplus B_{skew}$ of the subspaces of symmetric and skew-symmetric bilinear forms. Another way to express this is to say that there is an involution $\psi : (V^* \otimes V^*) \to (V^* \otimes V^*)$ exchanging the two factors; equivalently, we have a

map $\psi : B \to B$ sending the bilinear form $b(x, y)$ to the bilinear form $b(y, x)$. Since this is an automorphism of order 2 (and hence diagonalizable, with eigenvalues ± 1), we naturally get a direct sum decomposition of B into the $(+1)$-eigenspace and the (-1)-eigenspace for ψ. These are the subspaces we're identifying as B_{symm} and B_{skew}.

This also gives us a way of identifying subspaces of higher tensor powers of a given vector space V. For example, let $V^{\otimes d} = V \otimes \cdots \otimes V$ be the tensor product of V with itself d times. The symmetric group S_d on d letters acts on $V^{\otimes d}$ by permuting factors; we say that an element $\eta \in V^{\otimes d}$ is a *symmetric tensor* if it's invariant under this action, and we define the dth *symmetric power* $\mathrm{Sym}^d V$ to be the space of symmetric tensors.

There are, of course, other ways of defining the symmetric powers of a vector space V. For example, we could start with the tensor power $V^{\otimes d}$, and take the quotient by the subspace spanned by elements of the form

$$v_1 \otimes v_2 \otimes v_3 \otimes \cdots \otimes v_d \; - \; v_2 \otimes v_1 \otimes v_3 \otimes \cdots \otimes v_d$$

and similarly if we switch two other factors. The two definitions agree in characteristic 0, where we can apply the averaging operator

$$\alpha : v_1 \otimes v_2 \otimes \cdots \otimes v_d \; \mapsto \frac{1}{d!} \sum_{\sigma \in S_d} v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(d)}$$

but not in characteristics dividing $d!$. In general, these are two different constructions, which is why many people have opted to use instead the definition via universal properties: $\mathrm{Sym}^d(V)$ is the (unique up to isomorphism) vector space with d-linear map $\beta : V^d = V \times V \times \cdots \times V \to \mathrm{Sym}^d(V)$ such that for any symmetric d-linear map $\alpha : V^d \to U$ there exists a unique linear map $\tilde{\alpha} : \mathrm{Sym}^d V \to U$ with $\alpha = \tilde{\alpha} \circ \beta$.

It's not hard to write down a basis for $\mathrm{Sym}^d V$: if $\{e_1, \ldots, e_n\}$ is a basis for V, then to any n-tuple $I = (i_1, \ldots, i_n)$ of nonnegative integers adding up to d we can associate the element $e^I \in \mathrm{Sym}^d V$ that is the image, under the averaging operator α above, of the tensor $e_1 \otimes \cdots \otimes e_1 \otimes e_2 \otimes \cdots \otimes e_2 \cdots \otimes e_n \otimes \cdots \otimes e_n$, where the factor e_1 appears i_1 times, etc. In fact, in this way the direct sum

$$\mathrm{Sym}^\bullet V = \bigoplus_{d=0}^{\infty} \mathrm{Sym}^d V,$$

called the *symmetric algebra* of V, can be given the structure of a commutative ring (note that the multiplication here is not the product in the tensor algebra, but rather the product in the tensor algebra followed by the symmetrizing operator α), and this ring is in fact the polynomial ring in n variables, $K[e_1, \ldots, e_n]$.

4. Exterior powers

Everything we've just done has an analogue with an extra sign factor thrown in, though the results are quite different. Briefly, we say an element $\eta \in V^{\otimes d}$ is *alternating* if when we permute the factors of $V^{\otimes d}$ by a permutation $\sigma \in S_d$, the tensor η is carried into $(-1)^\sigma \eta$, where we denote by $(-1)^\sigma$ the sign of the permutation σ. The space of all alternating tensors of degree d on V is denoted $\bigwedge^d V$, and called the dth *exterior power* of V.

As before, we have two alternative definitions: we can define $\bigwedge^d V$ to be the quotient of $V^{\otimes d}$ by the subspace spanned by tensors of the form

$$v_1 \otimes v_2 \otimes v_3 \otimes \cdots \otimes v_d \ + \ v_2 \otimes v_1 \otimes v_3 \otimes \cdots \otimes v_d$$

(and similarly for other transpositions swapping two factors); and we can similarly define it as the universal object through which all skew-symmetric multilinear maps $V^d \to U$ factor.

Just as in the case of the symmetric powers, it's not hard to write down a basis for $\bigwedge^d V$, this time using the "skew-symmetrizing" operator β, defined by

$$\beta : v_1 \otimes v_2 \otimes \cdots \otimes v_d \ \mapsto \frac{1}{d!} \sum_{\sigma \in S_d} (-1)^\sigma v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(d)}.$$

In this case, though, the image of a given tensor $v_1 \otimes v_2 \otimes \cdots \otimes v_d$, denoted $v_1 \wedge v_2 \wedge \cdots \wedge v_d$, is 0 whenever a factor v_i is repeated. Thus, if $\{e_1, \ldots, e_n\}$ is a basis for V, then to any subset $I = \{i_1, \ldots, i_d\}$ of cardinality d we can associate the element $e^I \in \bigwedge^d V$ that is the image, under the averaging operator β above, of the tensor $e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_d}$; and these form a basis for $\bigwedge^d V$.

In fact, in this way the direct sum

$$\bigwedge^\bullet V := \bigoplus_{d=0}^{\infty} \bigwedge^d V,$$

called the *exterior algebra* of V, can be given the structure of a ring (note that the multiplication here is not the product in the tensor algebra, but rather the product in the tensor algebra followed by the skew-symmetrizing operator β). There are two big differences: first, multiplication in the exterior algebra is not commutative but skew-commutative; and second, the exterior algebra is actually a vector space of finite dimension 2^n, with all the direct summands of degree $d > n$ equal to 0.

5. Trace and determinant

Two of the key invariants of an operator $\phi : V \to V$ on a finite-dimensional vector space are its *trace* and its *determinant*. They are commonly defined in terms of a matrix representation of ϕ, which is convenient for explicit calculations but still unfortunate: you shouldn't have to choose a basis for V in order to define these two quantities. Happily, we don't have to; here forthwith are the basis-free definitions of trace and determinant.

To start with the relatively straightforward: in terms of the identification $\mathrm{Hom}(V, V) = V^* \otimes V$, the trace map $\mathrm{Hom}(V, V) \to K$ is simply the linear operator $V^* \otimes V \to K$, called *contraction*, which sends a rank one tensor $\ell \otimes v$ to $\ell(v)$.

As for the determinant, this emerges from the simple observation that if V is an n-dimensional vector space, then $\bigwedge^n V$ is one-dimensional. If $\phi : V \to V$ is any linear map it induces a map $\wedge^n \phi : \bigwedge^n V \to \bigwedge^n V$, which must then be multiplication by a scalar; that scalar is the determinant of ϕ.