

LINEAR ALGEBRA TENSOR PRODUCT NOTES

1. RECALLING BASES

Our notation: we let K be either \mathbb{R} or \mathbb{C} throughout the course.

If V and W are finite dimensional vector spaces, we typically write bases as

$$\mathcal{A} = \{v_1, v_2, \dots, v_p\} \subset V,$$

or

$$\mathcal{B} = \{w_1, w_2, \dots, w_q\} \subset W.$$

2. RECALLING DIRECT SUMS

We defined the direct sum of two vector spaces V and W to be the set

$$V \oplus W$$

whose elements are the pairs (v, w) for any $v \in V$ and $w \in W$. Addition and scaling are componentwise.

In this notation, given bases \mathcal{A} of V and \mathcal{B} of W , the set \mathcal{C} of vectors:

$$(v_1, 0), (v_2, 0), \dots, (v_p, 0), \\ (0, w_1), (0, w_2), \dots, (0, w_q)$$

is a basis of $V \oplus W$.

Similarly we define the direct sum $U \oplus V \oplus W$ to be the set of triples (u, v, w) with $u \in U, v \in V, w \in W$, etc.

3. THE IDEA OF TENSOR PRODUCT

Just as we can add vector spaces, we can also multiply them. The tensor product of two vector spaces V and W is a vector space, written $V \otimes W$, which we will define precisely later. The tensor product comes with a multiplication operation: for each vector $v \in V$ and vector $w \in W$, we have a choice of vector $v \otimes w \in V \otimes W$. This vector $v \otimes w$ depends linearly on v if you fix w , and linearly on w if you fix v . Moreover, if we have bases

$$v_1, v_2, \dots, v_p \in V,$$

and

$$w_1, w_2, \dots, w_q \in W,$$

then the products

$$\begin{array}{cccc} v_1 \otimes w_1, & v_1 \otimes w_2, & \dots, & v_1 \otimes w_q, \\ v_2 \otimes w_1, & v_2 \otimes w_2, & \dots, & v_2 \otimes w_q, \\ \vdots & \vdots & \vdots & \vdots \\ v_p \otimes w_1, & v_p \otimes w_2, & \dots, & v_p \otimes w_q \end{array}$$

form a basis for $V \otimes W$, called the *product basis*.

4. RECALLING DUAL SPACES

If V is a vector space, then V^* is the space of all linear functions

$$f: V \rightarrow K.$$

Under obvious adding and scaling, V^* is a vector space.

Each linear map

$$\phi: V \rightarrow W$$

has a *dual map*

$$\phi^*: W^* \rightarrow V^*$$

defined by

$$\phi^*(\alpha) = \alpha \circ \phi.$$

(This is a vector space version of transpose.)

Each basis

$$\mathcal{A} = \{v_1, v_2, \dots, v_p\} \subset V,$$

has a *dual basis* often written as

$$\mathcal{A}^* = \{\xi_1, \xi_2, \dots, \xi_p\} \subset V^*,$$

given by the linear functions which, on any vector

$$v = \sum a_j v_j$$

yield the number

$$\xi_j(v) = a_j.$$

Elements of V^* are called *covectors*, or *dual vectors*. But you can also just think of them as linear functions on V . We will write them using Greek letters like ξ instead of as functions like f , just to follow tradition.

Recall that to each vector $v \in V$, we associated a function $F_v: V^* \rightarrow K$ by the weird rule that, for any linear function $f: V \rightarrow K$, we let $F_v(f) = f(v)$.

This gives us an isomorphism

$$v \in V \mapsto F_v \in V^{**},$$

so that we can pretend that V^{**} is V .

5. MULTILINEAR FUNCTIONS

Definition 1. If U and V are vector spaces, then a map $f(u, v)$ defined for any $u \in U$ and $v \in V$ is *bilinear* if $f(u, v)$ is linear in u for any fixed constant choice of v and linear in v for any fixed constant choice of u .

Definition 2. If V_1, V_2, \dots, V_p are some vector spaces, then a function

$$f(v_1, v_2, \dots, v_p)$$

defined for $v_1 \in V_1, v_2 \in V_2, \dots, v_p \in V_p$ is *multilinear* if f is linear in each vector, when all of the other vectors are held constant.

Example 1. The function $f(x, y) = xy$ is multilinear, for x any y numbers. Note that f is *not* a linear function of the vector variable (x, y) . It scales quadratically if you scale both x and y . But it scales linearly if you scale x with fixed y , or scale y with fixed x .

Example 2. If $\xi: U \rightarrow K$ and $\eta: V \rightarrow K$ are linear functions (in other words, $\xi \in U^*$ and $\eta \in V^*$), then $\zeta(u, v) = \xi(u)\eta(v)$ is bilinear.

6. MULTILINEAR MAPS

Definition 3. If U, V and Z are vector spaces, a map t associating a vector

$$t(u, v) \in Z$$

to any vectors $u \in U$ and $v \in V$ is *bilinear* if t is linear in u for any fixed constant choice of v and vice versa. Multilinear maps are defined similarly.

7. TENSORS

If V_1, V_2, \dots, V_p are any finite dimensional vector spaces, then we let

$$V_1 \otimes V_2 \otimes \dots \otimes V_p$$

be the set of all multilinear maps

$$t(\xi_1, \xi_2, \dots, \xi_p),$$

where

$$\xi_1 \in V_1^*, \xi_2 \in V_2^*, \dots, \xi_p \in V_p^*$$

are any covectors.

Example 3. The simplest example: take just a single vector space V_1 . Then each vector $v \in V_1$ gives a linear map $F_v: V^* \rightarrow K$. This F_v is a tensor. We will pretend $V^{**} = V$, by identifying F_v with v , and in this way pretend that every vector v is a tensor. In this way, a vector is the simplest type of tensor.

Definition 4. If V and W are finite dimensional vector spaces and $v \in V$ and $w \in W$, write $v \otimes w$ for the multilinear map

$$t(\xi_1, \xi_2) = \xi_1(v)\xi_2(w),$$

for any covectors $\xi_1 \in V^*$ and $\xi_2 \in W^*$. Call $v \otimes w$ the *tensor product* of the vectors v and w .

Definition 5. Similarly if

$$v_1 \in V_1, v_2 \in V_2, \dots, v_p \in V_p$$

are vectors in various finite dimensional vector spaces, then write

$$v_1 \otimes v_2 \otimes \dots \otimes v_p$$

for the multilinear map

$$t(\xi_1, \xi_2, \dots, \xi_p) = \xi_1(v_1)\xi_2(v_2)\dots\xi_p(v_p).$$

Definition 6. If s is a tensor in $V_1 \otimes V_2 \otimes \dots \otimes V_p$ and t is a tensor in $W_1 \otimes W_2 \otimes \dots \otimes W_q$, for various finite dimensional vector spaces, then define $s \otimes t$, the *tensor product* of s and t , to be the tensor

$$(s \otimes t)(\xi_1, \xi_2, \dots, \xi_p, \eta_1, \eta_2, \dots, \eta_q) = s(\xi_1, \xi_2, \dots, \xi_p)t(\eta_1, \eta_2, \dots, \eta_q).$$

Exercise: prove that

$$\begin{aligned} (av) \otimes w &= a(v \otimes w) = v \otimes (aw), \\ (v_1 + v_2) \otimes w &= v_1 \otimes w + v_2 \otimes w, \\ v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2, \end{aligned}$$

for any vectors $v, v_1, v_2 \in V$ and $w, w_1, w_2 \in W$ and any number $a \in K$.

Theorem 1. *Suppose that V and W are two finite dimensional vector spaces, with bases*

$$\mathcal{A} = \{v_1, v_2, \dots, v_p\} \subset V,$$

or

$$\mathcal{B} = \{w_1, w_2, \dots, w_q\} \subset W.$$

Then $V \otimes W$ is a finite dimensional vector space, with

$$\dim(V \otimes W) = (\dim V)(\dim W),$$

and with basis

$$\{v_1 \otimes w_1, v_1 \otimes w_2, \dots, v_p \otimes w_q\}.$$

Proof. Take the dual bases, say

$$\xi_1, \xi_2, \dots, \xi_p \in V^*,$$

and

$$\eta_1, \eta_2, \dots, \eta_q \in W^*.$$

Every tensor $t \in V \otimes W$ has the form

$$t(\xi, \eta) = \sum_{ij} a_{ij} b_j t(\xi_i, \eta_j),$$

if

$$\xi = \sum_i a_i \xi_i,$$

and

$$\eta = \sum_j b_j \eta_j.$$

So if we let

$$c_{ij} = t(\xi_i, \eta_j),$$

then t has an expression as

$$t = \sum_{ij} c_{ij} v_i \otimes w_j.$$

So t is a linear combination of the various $v_i \otimes w_j$.

Now we need to prove that this linear combination is unique. But if we had any other expression as a linear combination, say

$$t = \sum_{ij} C_{ij} v_i \otimes w_j,$$

then we immediately check that

$$C_{ij} = t(v_i, w_j) = c_{ij},$$

so the coefficients of the two expressions are the same. □

8. UNIVERSAL MAPPING THEOREM

Theorem 2. *Suppose that V, W and Z are finite dimensional vector spaces. Every linear map $\phi: V \otimes W \rightarrow Z$ defines a bilinear map*

$$\Phi(v, w) = \phi(v \otimes w).$$

Every bilinear map Φ is uniquely determined this way. In particular, the vector space of all bilinear maps from $V \times W$ to Z is isomorphic to the vector space of linear maps from $V \otimes W$ to Z .

Proof. Let B be the set of all bilinear maps $\Phi: V \times W \rightarrow Z$. Make B into a vector space using the obvious addition and scaling of bilinear maps. Let A be the set of all linear maps $\phi: V \otimes W \rightarrow Z$. Make A into a vector space in the obvious way too. Clearly we can map $A \rightarrow B$ by taking ϕ to $\Phi(v, w) = \phi(v \otimes w)$. Call this map $\alpha: A \rightarrow B$. Clearly this map is linear.

The kernel of α is the maps ϕ so that

$$\phi(v \otimes w) = 0$$

for all $v \in V$ and $w \in W$. But the pairs $v \otimes w$ span $V \otimes W$ by theorem 1 on the facing page. Therefore $\phi = 0$. So α has kernel 0.

We leave the reader to calculate the dimensions of the vector spaces A and B , by using bases to write out the elements. □

9. RANK

Definition 7. Tensors of the form $v \otimes w$ are called *pure tensors*. (Also, tensors of the form $u \otimes v \otimes w$ are pure, etc.)

Example 4. Take $V = W = \mathbb{R}^3$. Then $e_1 \otimes e_2 + e_2 \otimes e_2$ can be written as $(e_1 + e_2) \otimes e_2$, so it is pure.

Example 5. But $e_1 \otimes e_2 + e_3 \otimes e_2$ cannot be written as $v \otimes w$ (let's see why later), so it is *not* pure.

Every tensor can be written as a sum of pure tensors, for example in a basis.

Example 6. Consider the tensor

$$t = e_1 \otimes e_1 + e_2 \otimes e_3 + e_3 \otimes e_3.$$

This tensor is *not* pure (which is certainly not obvious just looking at it). Let's see why. Any pure tensor $x \otimes y$ must be

$$\begin{aligned} x \otimes y &= (x^1 e_1 + x^2 e_2 + x^3 e_3) \otimes (y^1 e_1 + y^2 e_2 + y^3 e_3) \\ &= x^1 y^1 e_1 \otimes e_1 + x^2 y^1 e_2 \otimes e_1 + x^3 y^1 e_3 \otimes e_1 \\ &\quad + x^1 y^2 e_1 \otimes e_2 + x^2 y^2 e_2 \otimes e_2 + x^3 y^2 e_3 \otimes e_2 \\ &\quad + x^1 y^3 e_1 \otimes e_3 + x^2 y^3 e_2 \otimes e_3 + x^3 y^3 e_3 \otimes e_3. \end{aligned}$$

If we were going to have $x \otimes y = e_1 \otimes e_1 + e_2 \otimes e_2 + e_3 \otimes e_3$, we would need $x^1 y^1 = 1, x^2 y^2 = 1, x^3 y^3 = 1$, but also $x^1 y^2 = 0$, so $x^1 = 0$ or $y^2 = 0$, contradicting $x^1 y^1 = x^2 y^2 = 1$.

Definition 8. The *rank* of a tensor is the minimum number of pure tensors needed to write it.

Exercise: what is the rank of the tensor

$$t = e_1 \otimes e_1 + e_2 \otimes e_3 + e_3 \otimes e_3?$$

Exercise: given a tensor

$$t = \sum c_{ij} v_i \otimes w_j,$$

the rank of t is the rank of the matrix

$$C = (c_{ij}).$$

There is no easy way to find ranks of tensors in general, say in tensor products like $U \otimes V \otimes W$.