Introduction to the Tensor Product

James C Hateley

In mathematics, a tensor refers to objects that have multiple indices. Roughly speaking this can be thought of as a multidimensional array. A good starting point for discussion the tensor product is the notion of direct sums.

REMARK: The notation for each section carries on to the next.

1. Direct Sums

Let V and W be finite dimensional vector spaces, and let $\beta_v = \{e_i\}_{i=1}^n$ and $\beta_w = \{f_j\}_{j=1}^m$ be basis for V and W respectively. Now consider the direct sum of V and W, denoted by $V \oplus W$. Then

$$\beta_v \cup \beta_w = \{e_i\}_{i=1}^n \cup \{f_j\}_{j=1}^m$$

forms a basis for $V \oplus W$. Now it easy to see that if the direct sum of two vector spaces is formed, say $V \oplus W = Z$, then we have $V \cong V \oplus 0 \subset Z$ and $W \cong 0 \oplus W \subset Z$. So viewed as subspaces of Z, we have that V and W are orthogonal ($W \perp V$). So if $z \in Z$, then we have the decomposition for z as

$$z = v + w = \sum_{i=1}^n a_i e_i + \sum_{j=1}^m b_j f_j,$$

we can also write

$$z = \begin{pmatrix} v \\ w \end{pmatrix}$$

2. The Dual Space and Dual Transformation

For completeness sake, if $T \in \mathcal{L}(V, W)$ then $T : V \to W$ and T is linear. The dual space of a vector space V^* is defined to be the space of all linear functions $v^* : V \to \mathbb{R}$. Now if $T \in \mathcal{L}(V, W)$, we can define the dual transformation T^*, by $T^* : W^* \to V^*$. This operation T^* is also commonly known as the adjoint. With this definition we have the following action, if $v \in V$ and $v^* \in V^*$ then we have

$$(T^* v^*)\, v = v^* (Tv) \quad \text{or} \quad (T^* v^*, v) = (v^*, Tv)$$

The following propositions are properties of a linear maps and their duals.

Proposition 1. *If $I_v : V \to V$ is the identity on V, then $I_v^* : V^* \to V^*$ is the identity on V^*.*

Proof: By a simple calculation using our action we have,

$$(I_v^* v^*, v) = (v^*, I_v v) = (v^*, v) \ \square$$

Proposition 2. *If $T, S \in \mathcal{L}(V)$, then we have $(T \circ S)^* = S^* \circ T^*$.*

Proof: Again by a simple calculation using our action we have,

$$((T \circ S)^* v^*, v) = (v^*, T \circ S v) = (T^* v^*, S v) = (S^* \circ T^* v^*, v) \ \square$$

Proposition 3. *If $T \in \mathcal{L}(V, W)$ is bijective then T^* is bijective, and we have the following compositions;*

$$\left(T^{-1} \circ T\right)^* = I_v^* \quad \text{and} \quad \left(T \circ T^{-1}\right)^* = I_w^*$$

Proof: If T is invertible, then the compositions follow from the domain and range of definitions. To show that T^* is bijective consider the our action.

$$\begin{aligned} \left(\left(T^{-1} \circ T\right)^* v^*, v\right) &= \left(T^* \circ (T^{-1})^* v^*, v\right) \\ &= \left((T^{-1})^* v^*, Tv\right) = 0 \\ \Leftrightarrow \quad v = 0 \quad \Rightarrow & \quad \left(v^*, T^{-1} 0\right) \quad \Rightarrow \quad v^* = 0 \ \square \end{aligned}$$

Now if V is finite dimensional with basis β_v it's natural to wonder what the dimension of V^* is.

Definition 1. *If $\beta_v = \{e_i\}$ is a basis for V, then the set of covectors $\beta_v^* = \{e^i\}$ as a basis for V^* if*

$$(e^j, e_i) = \delta_i^j.$$

If $v \in V$, v can be expanded in terms of basis vectors. Consider our action on this expansion we observe that $\dim(V^*) = \dim(V)$.

3. Tensor Product

Now that we have an overview of a linear space and its dual we can start to define the tensor product.

Definition 2. *Let $\{V_i\}_{i=1}^k$ be a set of vector spaces. The map $\tau : \prod_{i=1}^k V_i \to \mathbb{R}$ is multilinear or k-linear if τ is linear in each coordinate. i.e, the map $v_i \to \tau(v_1, \ldots, v_i, \ldots v_k)$ is linear, or*

$$\begin{aligned} \tau(v_1, \ldots, av_i, \ldots v_k) &= a\tau(v_1, \ldots, v_i, \ldots v_k) \quad \textit{and} \\ \tau(v_1, \ldots, v_{1i} + v_{2i}, \ldots v_k) &= \tau(v_1, \ldots, v_{1i}, \ldots v_k) + \tau(v_1, \ldots, v_{2i}, \ldots v_k) \end{aligned}$$

Where $\prod_{i=1}^k$ denotes the k-fold Cartesian cross product.

Now that we have this definition of multilinear, it's natural to wonder what sort of object the collection of all these maps $\{\tau\}$ form.

Proposition 4. *The set of all such k-linear maps $\tau : \prod_{i=1}^k V_i \to W$ form a vector space.*

Proof: Check the vector space axoims □

Now Suppose $V = V_i$ for $i = 1$ to k, define the a set of linear maps by;

$$\mathcal{T}^k(V), \quad \text{s.t. if} \quad \tau \in \mathcal{T}^k(V), \text{ then } \tau : \prod_{i=1}^k V \to \mathbb{R}.$$

Now by a simple observation we see that $\mathcal{T}^1(V) = V^*$. We have just generalized the notion of a dual space, these spaces lead to the definition of a tensor. Before we go through the definition of tensor space, we need to define the another dual map, and the tensor product

Proposition 5. *If $T \in \mathcal{L}(V, W)$, then there exists a map $T^* : \mathcal{T}^k(W) \to \mathcal{T}^k(V)$*

Proof: OMIT: see [1] chapter 16.

Now let $\tau \in \mathcal{T}^k(V), \sigma \in \mathcal{T}^k(V)$, we can define the tensor product $\otimes$, between τ and σ by

$$\tau \otimes \sigma \in \mathcal{T}^{k+l}(V) \text{ and } \tau \otimes \sigma(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+l}) = \tau(v_1, \ldots, v_k)\sigma(v_{k+1}, \ldots, v_{k+l})$$

Note with this definition we do not have commutativity, but we have associativity. i.e.,

$$\tau \otimes \sigma \neq \sigma \otimes \tau \qquad \text{and} \qquad (\sigma \otimes \tau) \otimes \nu = \sigma \otimes (\tau \otimes \nu)$$

Now let's digress from this formulation and give a more formal definition.

Definition 3. *Let V and W be two vector spaces. The tensor product of V and W denoted by $V \otimes W$ is a vector space with a bilinear map*

$$\otimes : V \times W \to V \otimes W$$

which has the universal property.

In otherwords, if $\tau : V \times W \to Z$, then there exists a unique linear map, up to isomorphism, $\tilde{\tau} : V \otimes W \Rightarrow Z$ such that $\otimes \circ \tilde{\tau} = \tau$. The diagram for universal property can be seen in figure 1 below. Another way to say this is that a map $\tau \in \mathcal{L}^2(V \times W, Z)$ induces a map $\tilde{\tau} \in \mathcal{L}(V \otimes W, Z)$

Proposition 6. *The tensor product between V and W always exists.*

Proof: OMIT: see [1] chapter 16.

Now that we have the a formal definition for the tensor product, using the notation from section 1, we can define a basis for $V \otimes W$.

Definition 4. *If β_v and β_w are basis for V and W respectively, then a basis for $V \otimes W$ is defined by*

$$\beta_v \otimes \beta_w = \{e_i \otimes f_j\}_{i,j=1}^{n,m}$$

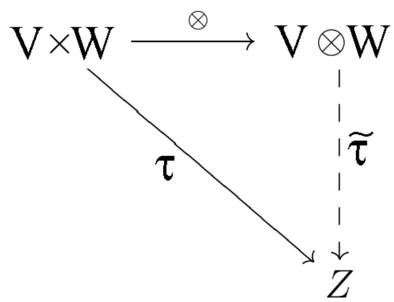

FIGURE 1. universal property for tensor product

With this definition we have that $\dim(V \otimes W) = mn$. Now if $\alpha \in \mathbb{R}$ the element $\alpha(e_i \otimes f_j)$ is called a simple tensor, and ff $v \in V$ and $w \in W$, the elements $v \otimes w$ are called tensors. Every tensor can be decomposed into simple tensors

$$v \otimes w = \sum_{i,j}^{n,m} a_{ij}(e_i \otimes f_j) \tag{3.1}$$

Now if $v_1, v_2, v \in V$, $w_1, w_2, w \in W$ and α is a scalar, then the follow properties of tensor are easily observed.

$$\begin{aligned} (v_1 + v_2) \otimes w &= v_1 \otimes w + v_2 \otimes w \\ v \otimes (w_1 + w_2) &= v \otimes w_1 + v \otimes w_2 \\ \alpha(v \otimes w) = (\alpha v) \otimes w &= v \otimes (\alpha w) \end{aligned}$$

4. General Tensors and Examples

Now that we have the a definition of the tensor product in general.

Definition 5. *Let* $\mathcal{T}_s^r(V) = \overbrace{V \otimes \cdots \otimes V}^{r} \otimes \overbrace{V^* \otimes \cdots \otimes V^*}^{s} = \bigotimes_r V \otimes \bigotimes_s V^*$, *then* $\mathcal{T}_s^r(V)$ *is said to be a tensor of type* (r, s).

Earlier we saw how to multiply two tensors τ and σ of type $(k, 0)$ and $(l, 0)$ respectively. The new order is the sum of the orders of the original tensors. When described as multilinear maps the tensor product simply multiplies the two tensors; i.e,

$$\tau \otimes \sigma \in \mathcal{T}^{k+l}(V) \text{ and } \tau \otimes \sigma(v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+l}) = \tau(v_1, \ldots, v_k)\sigma(v_{k+1}, \ldots, v_{k+l}),$$

which again produces a map that is linear in all its arguments. On components the effect similarly is to multiply the components of the two input tensors. If τ is of type (k, l) and σ is of type (n, m), then the tensor product $\tau \otimes \sigma$ is of type $(k + n, l + m)$ and is given by

$$(\tau \otimes \sigma)^{i_1 \ldots i_{l+m}}_{j_1 \ldots j_{k+n}} = \tau^{i_1 \ldots i_l}_{j_1 \ldots j_k} \sigma^{i_{l+1} \ldots i_{l+m}}_{j_{k+1} \ldots j_{k+n}}$$

Examples: Below are examples of recognizable tensors.

- $\mathcal{T}_0^0(V)$ is a tensor of type $(0, 0)$, also known as scalars.
- $\mathcal{T}_0^1(V)$ is a tensor of type $(1, 0)$, also known as vectors.
- $\mathcal{T}_1^0(V)$ is a tensor of type $(0, 1)$, also known as covectors, linear functionals or 1-forms.
- $\mathcal{T}_1^1(V)$ is a tensor of type $(1, 1)$, also known as a linear operator.

More Examples:

- An an inner product, a 2-form or metric tensor is an example of a tensor of type $(0, 2)$

- A bivector(oriented plane segment) is a tensor of type $(2, 0)$.
- If $\dim(V) = 3$ then the cross product is an example of a tensor of type $(1, 2)$.
- If $\dim(V) = n$ then a tensor of type $(0, n)$ is an $N-$form i.e. determinant or volume form.

From looking at this we have a sort of natural extension of the cross product from $\mathbb{R}^3$. If $\dim(V) = n$, then a tensor of type $(1, n-1)$ is a sort of *crossproduct* for V.

A Simple Computational Example: Let $v \in \mathbb{R}^n$, and $w \in \mathbb{R}^m$, treating these like column vectors, we can form the tensor product of v and w by;

$$v \otimes w = vw^t \in \mathcal{M}_{n,m}(\mathbb{R}) \text{ or } w \otimes v = wv^t \in \mathcal{M}_{m,n}(\mathbb{R})$$

In each case we get a matrix of rank 1.

Another Computational Example: Consider $A, B \in \mathcal{M}_2(\mathbb{R})$, then we can compute the tensor product between these matricies as follows:

$$\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \otimes \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} = \begin{bmatrix} a_{1,1} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} & a_{1,2} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \\ & \\ a_{2,1} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} & a_{2,2} \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_{1,1}b_{1,1} & a_{1,1}b_{1,2} & a_{1,2}b_{1,1} & a_{1,2}b_{1,2} \\ a_{1,1}b_{2,1} & a_{1,1}b_{2,2} & a_{1,2}b_{2,1} & a_{1,2}b_{2,2} \\ a_{2,1}b_{1,1} & a_{2,1}b_{1,2} & a_{2,2}b_{1,1} & a_{2,2}b_{1,2} \\ a_{2,1}b_{2,1} & a_{2,1}b_{2,2} & a_{2,2}b_{2,1} & a_{2,2}b_{2,2} \end{bmatrix}.$$

5. The Wedge Product and Examples

A lot of time in when studying geometry we see the symbol $\wedge$, this symbol denotes the wedge product. Before we can define it we first need to define the alternating product. Consider $\mathcal{T}^r(V)$, this space is spanned by decomposable tensors

$$v_1 \otimes \cdots \otimes v_r, \quad v_i \in V.$$

The antisymmetrization of this tensor is defined by;

$$Alt(v_1 \otimes \cdots \otimes v_r) = \frac{1}{r!} \sum_{\gamma \in \mathcal{S}_r} \operatorname{sgn}(\gamma) v_{\gamma(1)} \otimes \cdots \otimes v_{\gamma(r)},$$

where $\mathcal{S}_r$ is the permutation group on r elements. Now the image $Alt(T^r(V)) := \mathcal{A}^r(V)$ is a subspace of $\mathcal{T}^r(V)$. The space $\mathcal{A}^r(V)$ inherits the structure from the vector space from that on $\mathcal{T}^r(V)$ and carries a graded product defined by $\tau \hat{\otimes} \sigma = Alt(\tau \otimes \sigma)$. Now suppose $\tau \in \mathcal{A}^r(V) \subset \mathcal{T}^r(V)$, writing out the components of τ, we have

$$\tau = \tau^{i_1 i_2 \ldots i_r} e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_r}$$

Then we can define the wedge product of two alternating tensors τ and σ of ranks r and p by

$$\tau \wedge \sigma = \frac{(r+p)!}{r!p!} Alt(\tau \otimes \sigma) = \frac{1}{r!p!} \sum_{\gamma \in \mathcal{S}_{r+p}} \operatorname{sgn}(\gamma) \tau^{i_{\gamma(1)} \ldots i_{\gamma(r)}} \sigma^{i_{\gamma(r+1)} \ldots i_{\gamma(r+p)}} e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_{r+p}}.$$

Basic properties of $\wedge$:
bilinearity:
anti-commutativity: $\tau \wedge \sigma = (-1)^{rp} \sigma \wedge \tau$
Note that if r is odd we have $\tau \wedge \tau = 0$.

Before we proceed to a couple examples, first a little terminology from geometry. Let M, N be two manifolds, and let $f : M \to N$ and let f_{*_p} be the map defined at a point p on the tangent space of M defined by

$$f_{*_p} : TM_p \to TN_{f(p)}$$

sections of TM are called covariant, and sections of T^*M are called contravariant.

With respect to raising or lowering an index of a a curvature tensor, when a vector space is equipped with a metric tensor, there are operations that convert a contravariant (upper) index into a covariant (lower) index and vice versa. A metric itself is a (symmetric) (0,2)-tensor, it is thus possible to contract an upper index of a tensor with one of lower indices of the metric. This produces a new tensor with the same index structure as the previous, but with lower index in the position of the contracted upper index. This operation is as lowering an index. Conversely, a metric has an inverse which is a (2,0)-tensor. This inverse metric can be contracted with a lower index to produce an upper index. This operation is called raising an index.

Examples:

Let $\dim(M) = 2$, and let dx_1, dx_2 be the dual basis to the tangent bundle TM. Now consider the wedge product between dx_1, dx_2 applied to two elements in TM.

$$\begin{aligned}(dx_1 \wedge dx_2)(v_1, v_2) &= \frac{2!}{1!1!} Alt(dx_1 \otimes dx_2)(v_1, v_2) \\ &= dx_1(v_1)dx_2(v_2) - dx_1(v_2)dx_2(v_1)\end{aligned}$$

Example: grad, curl operators

Consider $f = f(x, y, z)$, and consider df as a linear functional, this can be computed as

$$df = f_x dx + f_y dy + f_z dz.$$

This is commonly knows as the gradient of f, or grad f or ∇f.

Let f, g, h be functions from $\mathbb{R}^3$ to $\mathbb{R}$ and let $\omega = f dx + g dy + h dz$, what about $d(\omega)$?. Using the fact that $dx_i \wedge dx_i = 0$ and the wedge product is skewsymmetric in $\mathbb{R}^3$ we have,

$$\begin{aligned} d\omega &= d\omega = f dx + g dy + h dz) \\ &= (f_x dx + g_x dy + h_x dz) \wedge dx + (f_y dx + g_y dy + h_y dz) \wedge dy + (f_z dx + g_z dy + h_z dz) \wedge dz \\ &= g_x dy \wedge dx + h_x dz \wedge dx + f_y dx \wedge dy + h_y dz \wedge dy + f_z dx \wedge dz + g_z dy \wedge dz \\ &= (h_y - g_z)\, dy \wedge dz + (h_x - f_z)\, dx \wedge dz + (g_x - f_y)\, dx \wedge dy \end{aligned}$$

This is commonly knows as the curl of ω or curlω or $\nabla \times \omega$.

Example: volume form

If ω is the volume-form of an n-dimensional manifold, then ω can be written in terms dx_i

$$\omega = \sqrt{|\det g_{ij}|} dx_1 \wedge \cdots \wedge dx_n$$

where g_{ij} is the metric on the manifold.

References

[1] Lang, S., *Algebra, Graduate Texts in Mathematics, (Revised third ed.)* New York: Springer-Verlag, 2002, MR1878556, ISBN 978-0-387-95385-4

[2] Spivak, M. *A comprehensive introduction to differential geometry. Vol 1* Publish or Perish Inc., 1999, ISBM 0-914098-83-7

[3] Wikipedia; *Tensor,* http://en.wikipedia.org/wiki/Tensor