MATH 725 NOTES
TOPICS IN MULTILINEAR ALGEBRA

MARIO L. GUTIERREZ ABED

The following definition is quite involved and we're going to use the following figure to guide us through it:

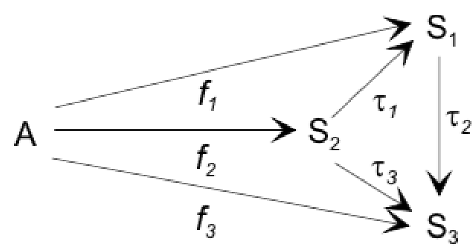

FIGURE 1

Definition. *Referring to* Figure 1, *let A be a set and let $\mathcal{S}$ be a family of sets. Let $\mathcal{F}$ be a family of functions from A to members of $\mathcal{S}$. Let $\mathcal{H}$ be a family of functions on members of $\mathcal{S}$. We assume that $\mathcal{H}$ has the following structure:*

- *$\mathcal{H}$ contains the identity function for each member of $\mathcal{S}$.*
- *$\mathcal{H}$ is closed under composition of functions.*
- *Composition of functions in $\mathcal{H}$ is associative.*

We also assume that for any $\tau \in \mathcal{H}$ and $f \in \mathcal{F}$, the composition $\tau \circ f$ is defined and is a member of $\mathcal{F}$.

*Let us refer to $\mathcal{H}$ as the **measuring set** and its members as **measuring functions**.*

*Then a pair $(S, f\colon A \to S)$ where $S \in \mathcal{S}$ and $f \in \mathcal{F}$ is said to have the **universal property for $\mathcal{F}$ as measured by $\mathcal{H}$**, or is a **universal pair** for $(\mathcal{F}, \mathcal{H})$, if for any $X \in S$ and any $g\colon A \to X$ in $\mathcal{F}$, there is a unique $\tau\colon S \to X$ in $\mathcal{H}$ for which the diagram in* Figure 1 *commutes, that is,*

$$g = \tau \circ f.$$

*Another way to express this is to say that any $g \in \mathcal{F}$ can be **factored through** f, or that any $g \in \mathcal{F}$ can be **lifted** to a function $\tau \in \mathcal{H}$ on S.* ★

<u>Example</u>: If V and W are vector spaces and $f \in V^*$ and $g \in W^*$, then the product map $\phi\colon V \times W \to F$ defined by

$$\phi(v, w) = f(v)g(w)$$

is bilinear. Dually, if $v \in V$ and $w \in W$, then the map $\lambda\colon V^* \times W^* \to F$ defined by

$$\lambda(f, g) = f(v)g(w)$$

is bilinear. ▲

Remark: It is precisely the *tensor product* that will allow us to generalize the previous example. In particular, if $\tau \in \mathcal{L}(U, W)$ and $\sigma \in \mathcal{L}(V, W)$, then we would like to consider a "product" map $\phi\colon U \times V \to W$ defined by

$$\phi(u, v) = f(u) \mathbin{?} g(v)$$

The tensor product $\otimes$ is just the thing to replace the question mark above, because it has the desired bilinearity property, as we will see. In fact, the tensor product is bilinear and nothing else, so it is exactly what we need!

Tensor Products

Referring to Figure 2 below, we seek to define a vector space T and a bilinear map $t\colon U \times V \to T$ so that any bilinear map f with domain $U \times V$ can be factored through t. Intuitively speaking, t is the most "general" or "universal" bilinear map with domain $U \times V$.

$$\begin{tikzcd}
U \times V \arrow[r, "t \text{ bilinear}"] \arrow[dr, "f \text{ bilinear}"'] & T \arrow[d, dashed, "\tau \text{ linear}"] \\
& W
\end{tikzcd}$$

Figure 2

Definition. *Let $U \times V$ be the cartesian product of two vector spaces over $\mathbb{F}$ and let $\mathcal{S} = \mathrm{Vect}(\mathbb{F})$.*[1] *Let*

$$\mathcal{F} = \{\mathrm{Hom}_{\mathbb{F}}(U, V; W) \mid W \in \mathcal{S}\}$$

be the family of all bilinear maps from $U \times V$ to a vector space W. Then the measuring set $\mathcal{H}$ is the family of all linear transformations $\{\tau\colon U \otimes V \to W \mid \tau \text{ is linear}\}$, and a pair $(T, t\colon U \times V \to T)$ is said to be ***universal for bilinearity*** *if it is universal for $(\mathcal{F}, \mathcal{H})$, that is, if for every bilinear map $g\colon U \times V \to W$, there is a unique linear transformation $\tau\colon U \otimes V \to W$ for which $g = \tau \circ t$.* ★

We now present the existence of tensor products in two different ways. First we present them in a way that is quite intuitive but that depends on the coordinates. The second approach is coordinate-free.

Existence I: Intuitive but not Coordinate-Free

The universal property for bilinearity captures the essence of bilinearity *and nothing more* (as is the intent for all universal properties). To understand better how this can be done, let $\mathcal{B} = \{e_i \mid i \in I\}$ be a basis for U and let $\mathcal{C} = \{f_j \mid j \in J\}$ be a basis for V. Then a bilinear map t on $U \times V$ is uniquely determined by assigning arbitrary values to the "basis" pairs (e_i, f_j). How can we do this and nothing more?

The answer is that we should define t on the pairs (e_i, f_j) in such a way that the images $t(e_i, f_j)$ *do not interact* and then extend by bilinearity.

[1] $\mathrm{Vect}(\mathbb{F})$ denotes the family of all vector spaces over the field $\mathbb{F}$.

In particular, for each ordered pair (e_i, f_j), we invent a new formal symbol, say $e_i \otimes f_j$ and define T to be the vector space with basis

$$\mathcal{D} = \{e_i \otimes f_j \mid e_i \in \mathcal{B}, f_j \in \mathcal{C}\}.$$

Then define the map t by setting

$$t(e_i, f_j) = e_i \otimes f_j$$

and extending by bilinearity. This uniquely defines a bilinear map t that is as "universal" as possible among bilinear maps.

Indeed, if $g\colon U \times V \to W$ is bilinear, then the condition $g = \tau \circ t$ is equivalent to

$$\tau(e_i \otimes f_j) = g(e_i, f_j)$$

which uniquely defines a linear map $\tau\colon T \to W$. Hence, the pair (T, t) has the universal property for bilinearity.

A typical element of T is a finite linear combination

$$\sum_{i,j=1}^{n} \alpha_{i,j}(e_{k_i} \otimes f_{k_j}),$$

and if $u = \sum \alpha_i e_i$ and $v = \sum \beta_j f_j$, then

$$u \otimes v = t(u, v) = t\left(\sum \alpha_i e_i, \sum \beta_j f_j\right) = \sum \alpha_i \beta_j (e_i \otimes f_j).$$

Note that, as is customary, we have used the notation $u \otimes v$ for the image of *any* pair (u, v) under t. Strictly speaking, this is an abuse of the notation $\otimes$ as we have defined it.

Confusion may arise because while the elements $u_i \otimes v_j$ form a basis for T (by definition), the larger set of elements of the form $u \otimes v$ do span T, but are definitely not linearly independent. This raises various questions, such as when a sum of the form $\sum u_i \otimes v_j$ is equal to 0, or when we can define a map τ on T by specifying the values $\tau(u \otimes v)$ arbitrarily. The first question seems more obviously challenging when we phrase it by asking when a sum of the form $\sum t(u_i, v_j)$ is 0, since there is no algebraic structure on the cartesian product $U \times V$, and so there is nothing "obvious" that we can do with this sum. The second question is not difficult to answer when we keep in mind that the set $\{u \otimes v\}$ is not linearly independent.

The notation $\otimes$ is used in yet another way: T is generally denoted by $U \otimes V$ and called the ***tensor product*** of U and V. The elements of $U \otimes V$ are called ***tensors*** and a tensor of the form $u \otimes v$ is said to be ***decomposable***:

For example, in $\mathbb{R}^2 \otimes \mathbb{R}^2$, the tensor $(1, 1) \otimes (1, 2)$ is decomposable but the tensor $(1, 1) \otimes (1, 2) + (1, 2) \otimes (2, 3)$ is not.

It is also worth emphasizing that the tensor product $\otimes$ is not a product in the sense of a binary operation on a set, as is the case in rings and fields, for example. In fact, even when $V = U$, the tensor product $u \otimes u$ is not in U, but rather in $U \otimes U$. It is wise to remember that the decomposable tensor $u \otimes v$ is nothing more than the image of the ordered pair (u, v) under the bilinear map t, as are the basis elements $e_i \otimes f_j$.

Existence II: Coordinate-Free

The previous definition of tensor product is about as intuitive as possible, but has the disadvantage of not being coordinate free. The following customary approach to defining the tensor product does not require the choice of a basis.

Let $\mathbb{F}_{U\times V}$ be the vector space over $\mathbb{F}$ with basis $U \times V$. Let S be the subspace of $\mathbb{F}_{U\times V}$ generated by all vectors of the form

$$\alpha(u, w) + \beta(v, w) - (\alpha u + \beta v, w)$$

and

$$\alpha(u, v) + \beta(u, w) - (u, \alpha v + \beta w),$$

where $\alpha, \beta \in \mathbb{F}$ and u, v, and w are in the appropriate spaces. Note that these vectors are 0 if we replace the ordered pairs by tensors according to our previous definition.

The quotient space

$$U \otimes V = (\mathbb{F}_{U\times V})/S$$

is also called the ***tensor product*** of U and V. The elements of $U \otimes V$ have the form

$$\left(\sum \alpha_i(u_i, v_i)\right) + S = \sum \alpha_i[(u_i, v_i) + S].$$

However, since by definition $\alpha(u, v) - (\alpha u, v) \in S$ and $\alpha(u, v) - (u, \alpha v) \in S$, we can "absorb" the scalar in either coordinate, that is,

$$\alpha[(u, v) + S] = (\alpha u, v) + S = (u, \alpha v) + S,$$

and so the elements of $U \otimes V$ can be written simply as

$$\sum [(u_i, v_i) + S].$$

It is customary to denote the coset $(u, v) + S$ by $u \otimes v$ and so any element of $U \otimes V$ has the form

$$\boxed{\sum u_i \otimes v_i}$$

as in the previous "non-coordinate-free" definition.

Finally, the map $t\colon U \times V \to U \otimes V$ is defined by

$$t(u, v) = u \otimes v.$$

Theorem. *Let U and V be vector spaces. Then the pair $(U \otimes V, t\colon U \times V \to U \otimes V)$ has the universal property for bilinearity, as measured by linearity.*