A short introduction to tensor products of vector spaces

ABSTRACT DEFINITION

1 BILINEAR MAPS. Let X, Y, and Z be vector spaces over some field $\mathbb{K}$. A map $\varphi\colon X \times Y \to Z$ is *bilinear* if

(i) for all $x \in X$, the map $Y \to Z$, $y \mapsto \varphi(x, y)$ is linear, and

(ii) for all $y \in Y$, the map $X \to Z$, $x \mapsto \varphi(x, y)$ is linear.

The set of bilinear maps from $X \times Y$ to Z is a vector space under pointwise operations; we denote it $B(X \times Y, Z)$.

2 EVALUATION OF LINEAR MAPS. Evaluation of linear maps is a bilinear map:

$$\begin{aligned} L(Y, Z) \times Y &\to Z \\ (A, y) &\mapsto Ay \end{aligned}$$

In a sense, this is the only kind of bilinear map: given a bilinear map $\psi\colon X \times Y \to Z$, we can define $T_\psi\colon X \to L(Y, Z)$ by $T_\psi(x) = \psi(x, \cdot)$; one can check that T_ψ is linear and that ψ is $T_\psi \times \mathrm{id}_Y$ followed by evaluation. (Here $T_\psi \times \mathrm{id}_Y$ means the map $T_\psi \times \mathrm{id}_Y(x, y) = (T_\psi(x), \mathrm{id}_Y(y))$.)

3 CURRYING. If $\psi\colon X \times Y \to Z$ is bilinear, then by definition $\psi(x, \cdot)\colon Y \to Z$ is linear; thus, as in §2, we obtain a map $T_\psi\colon X \to L(Y, Z)$, $x \mapsto \psi(x, \cdot)$. This map T_ψ is itself linear, and so we obtain a map $B(X \times Y, Z) \to L(X, L(Y, Z))$, $\psi \mapsto T_\psi$. This last map is a isomorphism; by symmetry, we have

$$B(X \times Y, Z) \cong L(X, L(Y, Z)) \cong L(Y, L(X, Z))$$

In §9 we will add $L(X \otimes Y, Z)$ to this list.

4 DEFINITION. Let X and Y be vector spaces over a field $\mathbb{K}$. A *tensor product* of X and Y is a vector space Z over $\mathbb{K}$, together with a bilinear map $\varphi\colon X \times Y \to Z$, satisfying the following *universal property*: for any vector space V and any bilinear map $\psi\colon X \times Y \to V$, there exists a unique linear map $\widetilde{\psi}\colon Z \to V$ such that the diagram

$$\begin{tikzcd} X \times Y \arrow[r, "\psi"] \arrow[d, "\varphi"'] & V \\ Z \arrow[ur, dashed, "\widetilde{\psi}"'] & \end{tikzcd}$$

commutes.

Note that if ψ is bilinear and T is linear then $T \circ \psi$ is bilinear; the definition of tensor product requires that *all* bilinear maps out of $X \times Y$ arise in this way from a single bilinear map.

5 EXISTENCE 1. Define

$$\Phi\colon X \times Y \to B(X \times Y, \mathbb{K})^\sharp$$

(where $^\sharp$ denotes the algebraic dual) by

$$\Phi(x, y)(\psi) = \psi(x, y) .$$

(So $\Phi(x, y)$ is the evaluation functional for (x, y).) Take $Z = \operatorname{span} \operatorname{Range} \Phi$ and $\varphi\colon X \times Y \to Z$, $\varphi(x, y) = \Phi(x, y)$. One can then show that Z has the desired universal property. (See Ryan (2002) for details.)

6 EXISTENCE 2. An alternative construction: Let W be the free vector space on $X \times Y$, that is, the space of formal linear combinations of elements of $X \times Y$. Impose the desired bilinearity relations as follows: let $\widetilde{W}$ be the subspace of W spanned by all elements of the form

$$\begin{aligned}
&\lambda(x, y) - (\lambda x, y) \\
&\lambda(x, y) - (x, \lambda y) \\
&(x, y + y') - (x, y) - (x, y') \\
&(x + x', y) - (x, y) - (x', y)
\end{aligned}$$

Then take $Z = W/\widetilde{W}$ and φ to be the natural embedding $X \times Y \to W$ followed by the quotient map $W \to Z$. One can then show that Z has the desired universal property.

7 UNIQUENESS. Suppose Z and Z' are tensor products of X and Y, with associated bilinear maps φ and φ'. Factor φ' through φ using the universal property of Z:

$$\begin{tikzcd}
X \times Y \arrow[r, "\varphi'"] \arrow[d, "\varphi"'] & Z' \\
Z \arrow[ur, dashed, "\widetilde{\varphi}'"'] &
\end{tikzcd}$$

Factor φ through φ' using the universal property of Z':

$$\begin{tikzcd}
X \times Y \arrow[r, "\varphi'"] \arrow[d, "\varphi"'] & Z' \arrow[dl, dashed, "\widetilde{\varphi}"] \\
Z &
\end{tikzcd}$$

Now consider the identity map on Z and the composition $\widetilde{\varphi} \circ \widetilde{\varphi}'$:

$$\begin{tikzcd} X \times Y \arrow[r, "\varphi"] \arrow[d, "\varphi"'] & Z \\ Z \arrow[ur, "\mathrm{id}_Z"'] & \end{tikzcd} \qquad\qquad \begin{tikzcd} X \times Y \arrow[r, "\varphi"] \arrow[d, "\varphi"'] \arrow[dr, "\varphi'"] & Z \\ Z \arrow[r, "\widetilde{\varphi}'"'] & Z' \arrow[u, "\widetilde{\varphi}"'] \end{tikzcd}$$

Both id_Z and $\widetilde{\varphi} \circ \widetilde{\varphi}'$ factor φ through φ as in the universal property of Z; by the uniqueness part of the universal property,

$$\mathrm{id}_Z = \widetilde{\varphi} \circ \widetilde{\varphi}' \ .$$

Similarly, $\mathrm{id}_{Z'} = \widetilde{\varphi}' \circ \widetilde{\varphi}$, and so $Z \cong Z'$ (and in fact, the isomorphism is "canonical", meaning that it is given by the maps produced by the universal property). Thus there is essentially only one tensor product.

8 NOTATION. We write $X \otimes Y$ for "the" tensor product of vector spaces X and Y, and we write $x \otimes y$ for $\varphi(x, y)$.

9 LINEARIZATION OF BILINEAR MAPS. Given a bilinear map $X \times Y \to V$, the universal property of the tensor product yields a unique map $X \otimes Y \to V$; thus we have a map $B(X \times Y, V) \to L(X \otimes Y, V)$, which in fact gives an isomorphism

$$B(X \times Y, V) \cong L(X \otimes Y, V) \ .$$

One case of special interest is

$$(X \otimes Y)^\sharp = L(X \otimes Y, \mathbb{K}) \cong B(X \times Y, \mathbb{K}) \ ,$$

and so $(X \otimes Y)^{\sharp\sharp} \cong B(X \times Y, \mathbb{K})^\sharp$. (This observation makes Ryan's construction §5 quite natural; it's analogous to embedding a Banach space in its double dual.)

CONCRETE REPRESENTATION OF ELEMENTS

10 ELEMENTARY TENSORS. Elements of a tensor product $X \otimes Y$ having the form $x \otimes y$ are called elementary tensors. They span $X \otimes Y$. (Indeed, the factorization condition in the universal property only determines the value of $\widetilde{\psi}$ on the span of the elementary tensors; since $\widetilde{\psi}$ is uniquely determined, this span must be all there is.) So every element of $X \otimes Y$ is a linear combination of elementary tensors. In fact, since

$$\sum_{i=1}^{n} \lambda_i (x_i \otimes y_i) = \sum_{i=1}^{n} (\lambda_i x_i) \otimes y_i \ ,$$

every element of $X \otimes Y$ is a *sum* of elementary tensors.

Digression

Before continuing with the theory of writing tensors as sums of elementary tensors, we need to develop a few fundamental tools. The proofs here also serve as examples of the universal property in use.

11 Tensoring with the field. $\mathbb{K} \otimes X \cong X$.

Proof Factor the scalar multiplication map through the tensor product, obtaining a linear map $M \colon \mathbb{K} \otimes X \to X$ such that $M(\lambda \otimes x) = \lambda x$.

$$\begin{tikzcd} \mathbb{K} \times X \arrow[r, "\text{scal.}", "\text{mul.}"'] \arrow[d, "\varphi"'] & X \\ \mathbb{K} \otimes X \arrow[ur, dashed, "M"'] & \end{tikzcd}$$

Define $N \colon X \to \mathbb{K} \otimes X$ by $N(x) = 1 \otimes x$. Note that N is linear. Furthermore,

$$MN(x) = M(1 \otimes x) = 1x = x\,,$$

so $MN = \mathrm{id}_X$, and

$$NM(\lambda \otimes x) = N(\lambda x) = 1 \otimes (\lambda x) = \lambda \otimes x\,,$$

and so $NM = \mathrm{id}_{\mathbb{K} \otimes X}$ (on elementary tensors, which is enough because both sides are linear and the elementary tensors span; see §10). □

12 Tensoring linear maps. Let $S \colon U \to V$ and $T \colon W \to X$ be linear maps. Then there is a unique linear map $S \otimes T \colon U \otimes W \to V \otimes X$ such that

$$S \otimes T(u \otimes w) = Su \otimes Tw$$

for all $u \in U$ and $w \in W$.

Proof Such a map is unique if it exists because we have specified its values on elementary tensors, and they span (§10). For existence, factor $\varphi_{V,X} \circ (S \times T)$ through $\varphi_{U,W}$:

$$\begin{tikzcd} U \times W \arrow[r, "S \times T"] \arrow[d, "\varphi_{U,W}"'] \arrow[dr] & V \times X \arrow[d, "\varphi_{V,X}"] \\ U \otimes W \arrow[r, dashed, "S \otimes T"'] & V \otimes X \end{tikzcd}$$

(Here $S \times T$ is the map $S \times T(u, w) = (Su, Tw)$. Note that it's not linear; but $\varphi_{V,X} \circ (S \times T)$ is bilinear.) □

13 AMBIGUITY 1. Note that "$S \otimes T$" here means an element of $L(U \otimes W, V \otimes X)$. It is also the name of an elementary tensor in $L(U, V) \otimes L(W, X)$. We will partially fix this ambiguity in §39.

14 COMPOSITION OF LINEAR MAPS. $S_1 T_1 \otimes S_2 T_2 = (S_1 \otimes S_2)(T_1 \otimes T_2)$. (Indeed, both are linear and have the same effect on elementary tensors.)

CONCRETE REPRESENTATION OF ELEMENTS, CONTINUED

15 LINEAR INDEPENDENCE 1. If $(y_i)_1^n$ are linearly independent and $\sum_{i=1}^n x_i \otimes y_i = 0$, then all $x_i = 0$. (Compare to the analogous statement with scalar multiplication in place of "$\otimes$".)

Proof Let $\psi \in X^\sharp$. Define $\Psi \colon X \otimes Y \to Y$ by $\Psi(x \otimes y) = \psi(x)y$. (This is $\psi \otimes \mathrm{id}_Y$ as in §12 followed by the identification $\mathbb{K} \otimes Y \cong Y$ as in §11.) We have

$$0 = \Psi(0) = \Psi\Big(\sum_{i=1}^n x_i \otimes y_i\Big) = \sum_{i=1}^n \Psi(x_i \otimes y_i) = \sum_{i=1}^n \psi(x_i) y_i \,.$$

Since the y_i are linearly independent, this implies $\psi(x_i) = 0$ for all i. Since ψ was arbitrary, all $x_i = 0$. □

16 LINEAR INDEPENDENCE 2. If $(y_i)_1^n$ are linearly independent and $\sum_{i=1}^n x_i \otimes y_i = \sum_{i=1}^n x_i' \otimes y_i$, then $x_i = x_i'$ for all i. (Again, compare to the analogous statement with scalar multiplication in place of "$\otimes$".)

17 LINEAR INDEPENDENCE 3. If $(x_\alpha)_\alpha$ are linearly independent and $(y_\beta)_\beta$ are linearly independent then $(x_\alpha \otimes y_\beta)_{\alpha,\beta}$ are linearly independent. (The proof is direct, once the observation that

$$\sum_\alpha \sum_\beta \lambda_{\alpha,\beta} x_\alpha \otimes y_\beta = \sum_\alpha \Big(\sum_\beta \lambda_{\alpha,\beta} x_\alpha\Big) \otimes y_\beta$$

is made.)

18 SPANNING SETS. If $(x_\alpha)_\alpha$ spans X and $(y_\beta)_\beta$ spans Y then $(x_\alpha \otimes y_\beta)_{\alpha,\beta}$ spans $X \otimes Y$.

19 BASES. If $(x_\alpha)_\alpha$ is a basis for X and $(y_\beta)_\beta$ is a basis for Y then $(x_\alpha \otimes y_\beta)_{\alpha,\beta}$ is a basis for $X \otimes Y$.

20 DIMENSION. $\dim(X \otimes Y) = \dim(X)\dim(Y)$.

21 NO ZERO DIVISORS. $x \otimes y = 0$ if and only if either $x = 0$ or $y = 0$. (Consider the dimensions of the spaces spanned by x, by y, and by $x \otimes y$.)

22 SHORTEST REPRESENTATIONS 1. If $\sum_{i=1}^{n} x_i \otimes y_i$ is a shortest representation of its value as a sum of elementary tensors, then $(x_i)_1^n$ are linearly independent and $(y_i)_1^n$ are linearly independent.

Proof By contraposition: if, say, the x_i have a dependency relation, then one of them can be expressed in terms of the others, which allows us to construct a shorter representation. □

(Compare the proof that all bases of a vector space have the same number of elements.)

23 SHORTEST REPRESENTATIONS 2. All representations of a tensor as a sum of elementary tensors with linearly independent factors have the same number of terms (which is called the *rank* of the tensor). More explicitly: if $(x_i)_{i=1}^m$ are linearly independent, and $(y_i)_{i=1}^m$ are linearly independent, and $(u_j)_{j=1}^n$ are linearly independent, and $(v_j)_{j=1}^n$ are linearly independent, and $\sum_{i=1}^m x_i \otimes y_i = \sum_{j=1}^n u_j \otimes v_j$, then $m = n$.

Proof Let $(x_i^\sharp)_1^m$ be coordinate functionals for the x_i. Applying $x_i^\sharp \otimes \mathrm{id}$ to both sides of the hypothesized equality yields

$$y_i = \sum_{j=1}^{n} x_i^\sharp(u_j) v_j \ .$$

(We've identified $\mathbb{K} \otimes Y$ with Y again.) Therefore $y_i \in \operatorname{span}(v_j)_1^n$. Since this is true for all i, and the y_i are linearly independent, this yields $m \leq n$; by symmetry, $m = n$. □

ALGEBRAIC FACTS

We have seen (§11) that the underlying field is the identity for tensor product of vector spaces, that is, $\mathbb{K} \otimes X \cong X$ (with correspondence given by $\lambda \otimes x \leftrightarrow \lambda x$). Here are some more algebraic facts.

24 COMMUTATIVITY. $X \otimes Y \cong Y \otimes X$, with correspondence $x \otimes y \leftrightarrow y \otimes x$.

25 ASSOCIATIVITY. $(X \otimes Y) \otimes Z \cong X \otimes (Y \otimes Z)$

26 DISTRIBUTIVITY. $X \otimes (Y \oplus Z) \cong (X \otimes Y) \oplus (X \otimes Z)$

27 INCLUSION. If $X \hookrightarrow X'$ then $X \otimes Y \hookrightarrow X' \otimes Y$, and the embedding sends $x \otimes y$ to $x \otimes y$, that is, $\varphi(x, y)$ to $\varphi'(x, y)$, where $\varphi\colon X \times Y \to X \otimes Y$ and $\varphi'\colon X' \times Y \to X' \otimes Y$ are the canonical bilinear maps.

Proof Since $i\colon X \hookrightarrow X'$ is injective, it has a left inverse, say $j\colon X' \to X$. Then $(j \otimes \mathrm{id}_Y)(i \otimes \mathrm{id}_Y) = \mathrm{id}_X \otimes \mathrm{id}_Y$ (see §14), so $i \otimes \mathrm{id}_Y$ has a left inverse, so it too is injective. □

28 QUOTIENTS. If X is a subspace of X' then $(X'/X) \otimes Y \cong (X' \otimes Y)/(X \otimes Y)$.

Proof Let $Q\colon X' \to X'/X$ be the quotient map. Then $Q \otimes \mathrm{id}_Y\colon X' \otimes Y \to (X'/X) \otimes Y$ is surjective (because it has a right inverse; analogous to §27). Let $z \in X' \otimes Y$, say, with shortest representation $z = \sum_{i=1}^n x_i \otimes y_i$. Then

$$\begin{aligned} z \in \ker(Q \otimes \mathrm{id}_Y) &\iff Q \otimes \mathrm{id}_Y(z) = 0 \\ &\iff \sum_{i=1}^n Qx_i \otimes y_i = 0 \\ &\iff (\forall i\colon Qx_i = 0) \\ &\iff (\forall i\colon x_i \in \ker Q) \\ &\iff (\forall i\colon x_i \in X) \end{aligned}$$

Therefore $\ker(Q \otimes \mathrm{id}_Y) = X \otimes Y$. The desired result follows by the first isomorphism theorem. □

(Unpacking this proof yields that the correspondence is $(a + X) \otimes y \leftrightarrow a \otimes y + X \otimes Y$. In other words, $[a] \otimes y \leftrightarrow [a \otimes y]$, where $[a]$ represents the equivalence class of a in X'/X and $[a \otimes y]$ represents the equivalence class of $a \otimes y$ in $(X' \otimes Y)/(X \otimes Y)$. In short, tensor product (of elements) commutes with taking the equivalence class.)

FUNCTION SPACES AND LINEAR OPERATORS

29 NOTATION. Let Ω be a set and X a vector space. We write $\mathcal{F}(\Omega, X)$ for the vector space of functions $\Omega \to X$ under pointwise operations.

30 Tensoring a function space. $\mathcal{F}(\Omega, X) \otimes Y \hookrightarrow \mathcal{F}(\Omega, X \otimes Y)$.

Proof Define $\Phi\colon \mathcal{F}(\Omega, X) \times Y \to \mathcal{F}(\Omega, X \otimes Y)$ by $\Phi(f, y)(\omega) = f(\omega) \otimes y$. Then Φ is bilinear, so we get a linear map $\widetilde{\Phi}\colon \mathcal{F}(\Omega, X) \otimes Y \to \mathcal{F}(\Omega, X \otimes Y)$ such that $\widetilde{\Phi}(f \otimes y) = \Phi(f, y)$. We wish to show $\widetilde{\Phi}$ is injective; we will show it has trivial kernel. Let $F \in \mathcal{F}(\Omega, X) \otimes Y$ be such that $\widetilde{\Phi}(F) = 0$. Let $F = \sum_{i=1}^{n} f_i \otimes y_i$ be a shortest representation of F as a sum of elementary tensors. Then $\sum_{i=1}^{n} \Phi(f_i, y_i) = 0$, that is, $\sum_{i=1}^{n} f_i(\omega) \otimes y_i = 0$ for all $\omega \in \Omega$. Since the y_i are linearly independent, this yields that $f_i(\omega) = 0$ for all i and all ω, which yields $F = 0$. □

So we can think of $f \otimes y \in \mathcal{F}(\Omega, X) \otimes Y$ as a function $\Omega \to X \otimes Y$, via $(f \otimes y)(\omega) = f(\omega) \otimes y$.

31 Nonsurjectivity. $\widetilde{\Phi}$ is usually not surjective: its image consists of those $F\colon \Omega \to X \otimes Y$ such that there exists a finite-dimensional subspace Y' of Y and $\operatorname{Range} F \subseteq X \otimes Y'$. But $\widetilde{\Phi}$ is surjective if Ω is a finite set or Y is finite-dimensional.

32 Functions of finite support. Let $\mathcal{F}_{\text{fin}}(\Omega, X)$ denote the space of functions $\Omega \to X$ having finite support. Then $\mathcal{F}_{\text{fin}}(\Omega, X) \otimes Y \cong \mathcal{F}_{\text{fin}}(\Omega, X \otimes Y)$.

33 Cartesian product of domains. $\mathcal{F}_{\text{fin}}(\Omega, \mathbb{K}) \otimes \mathcal{F}_{\text{fin}}(\Omega', \mathbb{K}) \cong \mathcal{F}_{\text{fin}}(\Omega, \mathbb{K} \otimes \mathcal{F}_{\text{fin}}(\Omega', \mathbb{K})) \cong \mathcal{F}_{\text{fin}}(\Omega, \mathcal{F}_{\text{fin}}(\Omega', \mathbb{K})) \cong \mathcal{F}_{\text{fin}}(\Omega \times \Omega', \mathbb{K})$

34 Finite-dimensional vector spaces. $\mathbb{K}^m \otimes \mathbb{K}^n \cong \mathbb{K}^{mn} \cong M_{m \times n}(\mathbb{K})$.

35 Power of vector spaces. $\mathbb{K}^n \otimes X \cong X^n$, because $\mathbb{K}^n = \mathcal{F}(\{1, \ldots, n\}, \mathbb{K})$ and $X^n = \mathcal{F}(\{1, \ldots, n\}, X)$. (Or use $\mathbb{K}^n \otimes X = (\mathbb{K} \oplus \cdots \oplus \mathbb{K}) \otimes X$ and §26.)

36 Field-valued vs vector-valued. $\mathcal{F}(\Omega, \mathbb{K}) \otimes X \hookrightarrow \mathcal{F}(\Omega, X)$.

37 Linear maps 1. $L(X, Y) \otimes Z \hookrightarrow L(X, Y \otimes Z)$, because $L(X, Y) = \mathcal{F}(\mathcal{B}, Y)$, where $\mathcal{B}$ is a basis for X. (We have isomorphism here if X is finite-dimensional or Z is.)

38 Linear maps 2. $X^\sharp \otimes Y \hookrightarrow L(X, Y)$, since $X^\sharp = L(X, \mathbb{K})$. The correspondence is given by $x^\sharp \otimes y(x) = \langle x, x^\sharp \rangle y$; the matrix representation of such maps is

$$\begin{bmatrix} \\ y \\ \\ \end{bmatrix} \begin{bmatrix} \quad x^\sharp \quad \end{bmatrix},$$

that is, $(x_j^\sharp y_i)_{i,j}$. The image of this embedding is the finite-rank operators.

39 Ambiguity 2. Using the inclusions above we obtain

$$\begin{aligned} L(U, V) \otimes L(W, X) &\hookrightarrow L(U, V \otimes L(W, X)) \\ &\hookrightarrow L(U, L(W, V \otimes X)) \\ &\cong B(U \times W, V \otimes X) \\ &\cong L(U \otimes W, V \otimes X) \end{aligned}$$

The resulting inclusion sends the elementary tensor $S \otimes T$ to the (unique) linear map $U \otimes W \to V \otimes X$ that sends $u \otimes w$ to $Su \otimes Tw$; we called this map $S \otimes T$ before (see §12, §13), so the ambiguity is somewhat acceptable. (It is very acceptable in cases when we have isomorphism, that is, when U and W are finite-dimensional, or U and V are finite-dimensional, or W and X are finite-dimensional.)

40 Duality. $X^\sharp \otimes Y^\sharp \hookrightarrow (X \otimes Y)^\sharp$, with isomorphism when X is finite-dimensional or Y is. Identifying tensors with their images here yields the rule

$$\langle x \otimes y, f \otimes g \rangle = \langle x, f \rangle \langle y, g \rangle \ .$$

(On the right we should strictly speaking have $\langle x, f \rangle \otimes \langle y, g \rangle$, but the space is $\mathbb{K} \otimes \mathbb{K}$, which we identify with $\mathbb{K}$.)

References

A more formal and complete treatment of tensor products of vector spaces can be found in the first chapter of Greub (1967). For tensor products of Banach spaces, see Ryan (2002). For tensor products of modules, see Lang (2002).

W. H. Greub. *Multilinear Algebra*. Number 136 in Die Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1967.

Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, London, 2002.

Raymond A. Ryan. *Introduction to Tensor Products of Banach Spaces*. Springer Monographs in Mathematics. Springer-Verlag, London, 2002.