Tensors, Multilinear Algebra, and Differential Forms

Chenjia Lin

December 24, 2021

Introduction

1 December 15-17, 2021

Source. eigenchris, *Tensors for Beginners*, Videos 0-10

We begin with the following heuristic for tensors:

> "A *tensor* is an object that is invariant under a change of coordinates and has components that transform in a special, predictable way under a change of bases."

Objects or quantities that are "invariant under a change of coordinates" include vectors, covectors, linear maps, and inner products; each of them are abstract objects that can be realized concretely using coordinates relative to a basis. We will first consider some examples of tensors before discussing what they are generally.

1.1 Transformation of Bases in a Vector Space

Suppose V is a finite-dimensional vector space. If $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_n\}$ are bases of V, then we can express one in terms of the other, namely

$$u_j = \sum_{i=1}^{n} F_j^i e_i \quad \text{and} \quad e_j = \sum_{i=1}^{n} B_j^i u_i, \qquad\qquad \text{where } F_j^i, B_j^i \in \mathbb{R}.$$

To clarify, the superscript i here is not an exponent, but an index. This strange notation falls under *Einstein's summation convention*, which we examine closely later. We refer to the matrix $F = (F_j^i)$ as tbe *forward transform* from $\{e_1, \ldots, e_n\}$ to $\{u_1, \ldots, u_n\}$ and $B = (B_j^i)$ the *backward transform*. More concretely, we have

$$F = \begin{pmatrix} F_1^1 & F_2^1 & \cdots & F_n^1 \\ F_1^2 & F_2^2 & \cdots & F_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ F_1^n & F_2^n & \cdots & F_n^n \end{pmatrix}, \qquad\qquad B = \begin{pmatrix} B_1^1 & B_2^1 & \cdots & F_n^1 \\ B_1^2 & B_2^2 & \cdots & B_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ B_1^n & B_2^n & \cdots & B_n^n \end{pmatrix}.$$

Notice that i is the row index and j the column index. Now observe that

$$u_j = \sum_{i=1}^{n} F_j^i \left(\sum_{k=1}^{n} B_i^k u_k \right) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} B_i^k F_j^i \right) u_k. \tag{1}$$

Since $u_1, u_2, \ldots, u_n$ are linearly independent, we must have

$$\sum_{i=1}^{n} B_i^k F_j^i = \begin{cases} 1 & k = j, \\ 0 & k \neq j. \end{cases} \tag{2}$$

This result corresponds to

$$BF = \begin{pmatrix} B_1^1 & B_2^1 & \cdots & B_n^1 \\ B_1^2 & B_2^2 & \cdots & B_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ B_1^n & B_2^n & \cdots & B_n^n \end{pmatrix} \begin{pmatrix} F_1^1 & F_2^1 & \cdots & F_n^1 \\ F_1^2 & F_2^2 & \cdots & F_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ F_1^n & F_2^n & \cdots & F_n^n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = I.$$

Repeating the computations in (1) with e_j instead of u_j likewise tells us $FB = I$. Thus, F and B are inverses and $B = F^{-1}$. Before moving on, we define the *Kronecker delta*

$$\delta_{kj} = \begin{cases} 1 & k = j, \\ 0 & k \neq j. \end{cases}$$

This will continue to appear later in our studies.

1.2 Transformation of a Vector's Coordinates

The vector is our first example of a tensor. Given a vector space V, we can define a coordinate system on it by picking a basis, whereby every vector in V is a linear combination of the basis vectors. The coefficients in the linear combination expressing a vector—its "coordinates"—depend on the basis selected. Nonetheless, the expressions all refer to the same vector. We will now examine how the coordinates of a vector change when we consider different coordinate systems.

Let $v \in V$ and suppose that $v = \sum_{j=1}^n a^j e_j = \sum_{j=1}^n b^j u_j$, where $a^j, b^j \in \mathbb{R}$. Again, j here is a superscript index. We compute a^j and b^j using forward and backward transforms; observe that

$$\begin{aligned} v &= \sum_{j=1}^n a^j e_j \\ &= \sum_{j=1}^n a^j \left(\sum_{i=1}^n B_j^i u_i \right) \\ &= \sum_{i=1}^n \left(\sum_{j=1}^n B_j^i a^j \right) u_i. \end{aligned}$$

If $v = \sum_{j=1}^n b^j u_j$, then we must have $b^i = \sum_{j=1}^n B_j^i a^j$. Conversely, we have $a^i = \sum_{j=1}^n F_j^i b^j$. Notice that to transform the coordinates of v under $\{e_1, \ldots, e_n\}$ to $\{u_1, \ldots, u_n\}$, we use the *backwards transform* rather than the forward. In general, when an object's coordinates transform in this seemingly odd manner, we say that the coordinates transform *contravariantly*. The following example should motivate our observation here.

Example 1.1. Consider the bases $\{e_1, \ldots, e_n\}$ and $\{2e_1, \ldots, 2e_n\}$ of V. If $v = \sum_{i=1}^n v^i e_i$, then

$$v = \sum_{i=1}^n \frac{v^i}{2}(2e_i).$$

The doubling transformation applied to $\{e_1, \ldots, e_n\}$ to obtain $\{2e_1, \ldots, 2e_n\}$ requires us to halve the coefficient v^i. We see that a vector's coordinates transform inversely from how the bases transform.

1.3 Dual Spaces and Transformation of Covector Coordinates

Our next example of a tensor is the *covector*.

Definition 1.2. The *dual space* V^* of V is the collection of all linear maps $\alpha : V \to \mathbb{R}$. We call $\alpha \in V$ a *covector* of V.

We consider V^* as a vector space by defining addition and scalar multiplication as follows; if $\varphi, \psi \in V^*$ and $c \in \mathbb{R}$, then

$$(\varphi + \psi)(v) = \varphi(v) + \psi(v), \qquad (c \cdot \varphi)(v) = c \cdot \varphi(v)$$

for all $v \in V$. Since V^* is a vector space, we'd naturally want to consider some basis of V^*. We can produce one from a basis of V. If $\{e_1, \ldots, e_n\}$ is a basis of V, then consider the covectors $\epsilon^1, \ldots, \epsilon^n$ where $\epsilon^i(e_j) = \delta_{ij}$. By convention, covectors are denoted by Greek letters and use superscript indices. We call $\{\epsilon^1, \ldots, \epsilon^n\}$ the *dual basis* of $\{e_1, \ldots, e_n\}$, and we leave it as an exercise to check that they are indeed covectors and form a basis of V^*.

Proposition 1.3. If $\{e_1, \ldots, e_n\}$ is a basis of V and $\epsilon^1, \ldots, \epsilon^n \in V^*$ satisfy $\epsilon^i(e_j) = \delta_{ij}$, then $\{\epsilon^1, \ldots, \epsilon^n\}$ is a basis of V^*.

We now examine how the bases of a dual space transform; suppose $\{\epsilon^1, \ldots, \epsilon^n\}$ and $\{\eta^1, \ldots, \eta^n\}$ are dual bases of $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_n\}$, respectively. Suppose that

$$\eta^i = \sum_{j=1}^{n} Q^i_j \epsilon^j \qquad \text{where } Q^i_j \in \mathbb{R}.$$

The trick now is to plug u_k into η^i and e^j, and we see that

$$\begin{aligned}
\eta^i(u_k) &= \sum_{j=1}^{n} Q^i_j \epsilon^j(u_k) \\
\delta_{ik} &= \sum_{j=1}^{n} Q^i_j \epsilon^j \left(\sum_{\ell=1}^{n} F^\ell_k e_\ell \right) \\
&= \sum_{j=1}^{n} \sum_{\ell=1}^{n} Q^i_j F^\ell_k \; \epsilon^j(e_\ell) \\
&= \sum_{j=1}^{n} \sum_{\ell=1}^{n} Q^i_j F^\ell_k \cdot \delta_{j\ell}.
\end{aligned}$$

Note that $\delta_{j\ell} = 1$ if $j = \ell$ and $\delta_{j\ell} = 0$ if $j \neq \ell$, so

$$\delta_{ik} = \sum_{j=1}^{n} Q^i_j F^j_k.$$

We have $\delta_{ik} = \sum_{j=1}^{n} B^i_j F^j_k$ by (2). This expression holds uniquely because B is the unique inverse of F, so it follows that $Q^i_j = B^i_j$. Then

$$\eta^i = \sum_{j=1}^{n} B^i_j \epsilon^j,$$

and likewise

$$\epsilon^i = \sum_{j=1}^{n} F^i_j \eta^j.$$

Thus, we see that $\{\epsilon^1, \ldots, \epsilon^n\}$ transforms covariantly into $\{\eta^1, \ldots, \eta^n\}$. To obtain a rule for how a covector's coordinates transform, let $\alpha = \sum_{i=1}^n \alpha_i \epsilon^i$ where $\alpha_i \in \mathbb{R}$. We have

$$\begin{aligned}
\alpha &= \sum_{i=1}^n \alpha_i \epsilon^i \\
&= \sum_{i=1}^n \alpha_i \left(\sum_{j=1}^n F_j^i \eta^j \right) \\
&= \sum_{j=1}^n \left(\sum_{i=1}^n \alpha_i F_j^i \right) \eta^j .
\end{aligned}$$

It appears that covector coordinates transform *covariantly*—the same way $\{e_1, \ldots, e_n\}$ transforms into $\{u_1, \ldots, u_n\}$.

1.4 Transformation of Linear Maps

The third example of a tensor is a linear map from V to itself. Recall from linear algebra that a linear map can be represented by a matrix with respect to some choice of basis. We will not derive a relation between the matrix representations—the "coordinates"—of a linear map $L : V \to V$ under different bases. As usual, let $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_n\}$ be bases of V. Suppose that

$$L(e_j) = \sum_{i=1}^n L_j^i e_i \quad \text{and} \quad L(u_j) = \sum_{i=1}^n \widetilde{L_j^i} u_i \qquad\qquad \text{where } L_j^i, \widetilde{L_j^i} \in \mathbb{R}.$$

To compute $\widetilde{L_j^i}$, we will derive an expression for $L(u_j)$ in terms of $u_1, \ldots, u_n$ using forward and backward transforms. We have $u_j = \sum_{\ell=1}^n F_j^\ell e_\ell$, so

$$L(u_j) = L\left(\sum_{\ell=1}^n F_j^\ell e_\ell \right) = \sum_{\ell=1}^n F_j^\ell L(e_\ell)$$

by linearity of L. Note that $L(e_\ell) = \sum_{k=1}^n L_\ell^k e_k$, so

$$\begin{aligned}
L(u_j) &= \sum_{\ell=1}^n F_j^\ell \left(\sum_{k=1}^n L_\ell^k e_k \right) \\
&= \sum_{k=1}^n \sum_{\ell=1}^n F_j^\ell L_\ell^k e_k .
\end{aligned}$$

Notice that $e_k = \sum_{i=1}^n B_k^i u_i$, so

$$\begin{aligned}
L(u_j) &= \sum_{k=1}^n \sum_{\ell=1}^n F_j^\ell L_\ell^k \left(\sum_{i=1}^n B_k^i u_i \right) \\
&= \sum_{i=1}^n \sum_{k=1}^n \sum_{\ell=1}^n B_k^i L_\ell^k F_j^\ell u_i .
\end{aligned}$$

If $L(u_j) = \sum_{i=1}^n \widetilde{L_j^i} u_i$, then we must have

$$\widetilde{L_j^i} = \sum_{k=1}^n \sum_{\ell=1}^n B_k^i L_\ell^k F_j^\ell .$$

To make the sum above a bit more concrete, let (L^i_j) and $\left(\widetilde{L^i_j}\right)$ be the matrix representations of L with respect to $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_n\}$, respectively. The expression above describes

$$\underbrace{\begin{pmatrix} \widehat{L^1_1} & \widehat{L^1_2} & \cdots & \widehat{L^1_n} \\ \widehat{L^2_1} & \widehat{L^2_2} & \cdots & \widehat{L^2_n} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{L^n_1} & \widehat{L^n_2} & \cdots & \widehat{L^n_n} \end{pmatrix}}_{\left(\widetilde{L^i_j}\right)} = \underbrace{\begin{pmatrix} B^1_1 & B^1_2 & \cdots & B^1_n \\ B^2_1 & B^2_2 & \cdots & B^2_n \\ \vdots & \vdots & \ddots & \vdots \\ B^n_1 & B^n_2 & \cdots & B^n_n \end{pmatrix}}_{B=F^{-1}} \underbrace{\begin{pmatrix} L^1_1 & L^1_2 & \cdots & L^1_n \\ L^2_1 & L^2_2 & \cdots & L^2_n \\ \vdots & \vdots & \ddots & \vdots \\ L^n_1 & L^n_2 & \cdots & L^n_n \end{pmatrix}}_{(L^i_j)} \underbrace{\begin{pmatrix} F^1_1 & F^1_2 & \cdots & F^1_n \\ F^2_1 & F^2_2 & \cdots & F^2_n \\ \vdots & \vdots & \ddots & \vdots \\ F^n_1 & F^n_2 & \cdots & F^n_n \end{pmatrix}}_{F}.$$

It appears that matrix representations transform both covariantly and contravariantly.

1.5 Einstein Summation Convention

Since the beginning of this note, we have been using a mix of subscript and superscript indices on our basis vectors, vector components, and matrix entries. Our placement of indices has been intentional, and the following is a summary of the rules.

1. Basis vectors have lower indices (e.g. $e_1, \ldots, e_n$).
2. Vector components have upper indicies (e.g. $v = v^1 e_1 + v^2 e_2 + \cdots + v^n e_n$).
3. Covectors are denoted by Greek letters and use upper indices (e.g. $\epsilon^1, \ldots, \epsilon^n$).
4. Matrices use lower indices for columns and upper indices for rows (e.g. $u_i = \sum_{\ell=1}^n B^i_j e_i$).

If we stick to these indexing rules, then we can apply *Einstein's summation convention*:

> "If an index is repeated as both a lower and upper index, then disregard the summation sign."

Here are some examples:

$$\sum_{i=1}^n v^i e_i = v^i e_i,$$
$$\sum_{i=1}^n F^i_j e_i = F^i_j e_i,$$
$$\sum_{k=1}^n \sum_{\ell=1}^n B^i_k L^k_\ell F^\ell_j = B^i_k L^k_\ell F^\ell_j.$$

Notice that $B^i_k L^k_\ell F^\ell_j$ in the last line closely resembles the product BLF; the indexing simply tells us how to multiply matrices, so we can write B^i_k to represent B. Now observe that if $M \equiv M^i_j$ is an $n \times n$ matrix, then

$$MI = M^i_j \cdot \delta^k_i = M^k_j.$$

We see that δ^k_i "cancels" out the i's, which is as an upper index on M^i_j and lower index on δ^k_i. This will be a useful observation for computations.

Example 1.4. Recall that $\widetilde{L^i_j} = B^i_k L^k_\ell F^\ell_j$. Observe that

$$\begin{aligned} F^s_i \widetilde{L^i_j} B^j_t &= F^s_i B^i_k L^k_\ell F^\ell_j B^j_t \\ &= \delta^s_k \cdot L^k_\ell \cdot \delta^\ell_t \\ &= L^s_\ell \cdot \delta^\ell_t \\ &= L^s_t. \end{aligned}$$

1.6 Transformation of the Metric Tensor

Our fourth example of a tensor is the *metric tensor*. Assuming that V is an inner product space, the metric tensor is the inner product $\langle \cdot, \cdot \rangle$ endowed onto V. The *norm* of $v \in V$ is defined as $\|v\| := \sqrt{\langle v, v\rangle}$, and we take $\|v\|$ to mean the *length* of v. The *angle* θ between $v, w \in V$ is given by

$$\cos\theta = \frac{\langle v, w\rangle}{\|v\|\|w\|}.$$

The inner product between two vectors is invariant under any choice of coordinates, just as vectors, covectors, and linear objects are abstract, "invariant" objects that have coordinate representations. As such, the inner product is indeed a tensor (in terms of the definition we gave before), and it follows that lengths and angles are well-defined, invariant quantities.

To get a better grasp of the metric tensor, we first obtain a concrete "coordinate" representation of it. If $\{e_1, \ldots, e_n\}$ is a basis of V, $v = v^i e_i$, and $w = w^i e_i$, then

$$\langle v, w\rangle = \langle v^i e_i, w^j e_j\rangle = v^i w^j \langle e_i, e_j\rangle. \tag{3}$$

We see then that $\langle v, w\rangle$ depends only on the coordinates of v and w with respect to $\{e_1, \ldots, e_n\}$ and the values of the inner products $\langle e_i, e_j\rangle$. We can encapsulate the sums and products above using matrices:

$$\langle v, w\rangle = \begin{pmatrix} v^1 & v^2 & \cdots & v^n \end{pmatrix} \underbrace{\begin{pmatrix} \langle e_1, e_1\rangle & \langle e_1, e_2\rangle & \cdots & \langle e_1, e_n\rangle \\ \langle e_2, e_1\rangle & \langle e_2, e_2\rangle & \cdots & \langle e_2, e_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle e_n, e_1\rangle & \langle e_n, e_2\rangle & \cdots & \langle e_n, e_n\rangle \end{pmatrix}}_{\text{denote as } g} \begin{pmatrix} w^1 \\ w^2 \\ \vdots \\ w^n \end{pmatrix}.$$

Define $g_{ij} = \langle e_i, e_j\rangle$ and let g be the matrix $g = (g_{ij})$. Just as a linear map $L : V \to V$ has a matrix representation L^i_j, the metric tensor $\langle \cdot, \cdot \rangle$ has g as its concrete, coordinate-based description.

We are now ready to consider how the representation of a metric tensor transforms under different bases; let $\{e_1, \ldots, e_n\}$ and $\{u_1, \ldots, u_n\}$ be bases of V and define $g_{ij} = \langle e_i, e_j\rangle$ and $\widetilde{g_{ij}} = \langle u_i, u_j\rangle$. Observe that

$$\begin{aligned} \widetilde{g_{ij}} &= \langle u_i, u_j\rangle \\ &= \langle F^k_i e_k, F^\ell_j e_\ell\rangle \\ &= F^k_i F^\ell_j \langle e_k, e_\ell\rangle \\ &= F^k_i F^\ell_j g_{k\ell}. \end{aligned}$$

By inversion, it follows that

$$\begin{aligned}
B^t_k \widetilde{g_{ij}} &= B^t_k F^k_i F^\ell_j g_{k\ell} \\
&= \delta^t_i \cdot F^\ell_j g_{k\ell} \\
\delta^j_t \cdot B^t_k \widetilde{g_{ij}} &= \delta^j_t \cdot \delta^t_i \cdot F^\ell_j g_{k\ell} \\
B^j_k \widetilde{g_{ij}} &= \delta^j_i \cdot F^\ell_j g_{k\ell} \\
&= F^\ell_i g_{k\ell} \\
B^i_s B^j_k \widetilde{g_{ij}} &= B^i_s F^\ell_i g_{k\ell} \\
&= \delta^\ell_s \; g_{k\ell} \\
&= g_{ks}.
\end{aligned}$$

Thus, we have $g_{k\ell} = B^i_\ell B^j_k \widetilde{g_{ij}}$. It appears that the coordinates of a metric tensor transform covariantly under different bases.

1.7 An Excursion into Bilinear Forms

The metric tensor (inner product) is a special case of what we call a *bilinear form.*

Definition 1.5. A *bilinear form* is a map $\mathcal{B} : V \times V \to \mathbb{R}$ such that

$$\begin{aligned}
\mathcal{B}(u, v + w) &= \mathcal{B}(u, v) + \mathcal{B}(u, w), \\
\mathcal{B}(u + w, v) &= \mathcal{B}(u, v) + \mathcal{B}(w, v), \\
\mathcal{B}(c \cdot u, v) &= c \cdot \mathcal{B}(u, v), \\
\mathcal{B}(u, c \cdot v) &= c \cdot \mathcal{B}(u, v).
\end{aligned}$$

for all $u, v, w \in V$ and $c \in \mathbb{R}$.

A bilinear form takes in a pair of vectors and returns a scalar such that $L(u) := \mathcal{B}(u, v_0)$ and $R(u) := \mathcal{B}(u_0, v)$ are linear maps for any fixed $u_0, v_0 \in V$. Notice that

$$\begin{aligned}
\langle u, v + w\rangle &= \langle u, v\rangle + \langle u, w\rangle, \\
\langle u + w, v\rangle &= \langle u, v\rangle + \langle w, v\rangle, \\
\langle c \cdot u, v\rangle &= c \cdot \langle u, v\rangle, \\
\langle u, c \cdot v\rangle &= c \cdot \langle u, v\rangle,
\end{aligned}$$

for all $u, v, w \in V$ and $c \in \mathbb{R}$, so the metric tensor is indeed a bilinear form. In fact, it is a *symmetric bilinear form* because $\langle u, v\rangle = \langle v, u\rangle$.

Definition 1.6. A bilinear form $\mathcal{B} : V \times V \to \mathbb{R}$ is *symmetric* if $\mathcal{B}(u, v) = \mathcal{B}(v, u)$ for all $u, v \in V$.

The point of our excursion into bilinear forms is that they are also tensors. The transformation rule $\widetilde{g_{ij}} = F^k_i F^\ell_j g_{k\ell}$ we derived for metric tensors generalize to that for bilinear forms, namely

$$\widetilde{\mathcal{B}_{ij}} = F^k_i F^\ell_j \mathcal{B}_{k\ell} \quad \text{and} \quad \mathcal{B}_{k\ell} = B^i_\ell B^j_k \widetilde{\mathcal{B}_{ij}}$$

where $\mathcal{B}_{ij} = \mathcal{B}(e_i, e_j)$ and $\widetilde{\mathcal{B}_{ij}} = \mathcal{B}(u_i, u_j)$. We have

$$\mathcal{B}(v^i e_i, w^j e_j) = v^i w^j \mathcal{B}(e_i, e_j) = v^i w^j \mathcal{B}_{ij},$$

so $\mathcal{B}(u, v)$ is completely determined by the coordinates of v and w relative to $\{e_1, \ldots, e_n\}$ and the $\mathcal{B}_{ij}$'s—just as we saw in (3).

1.8 Classification of Tensors

Recall our heuristic definition of a tensor:

> "A *tensor* is an object that is invariant under a change of coordinates and has components that transform in a special, predictable way under a change of bases."

We've seen that the coordinate representations of vectors, covectors, linear maps, and the metric tensor (inner product) transform according to specific rules. We classify tensors according to the number of covariant and contravariant components in their transformation rules. A (m, n)-*tensor* is one with m contravariant and n covariant components. For instance,

- vectors are $(1, 0)$-tensors because they transform contravariantly,
- covectors are $(0, 1)$-tensors because they transform covariantly,
- linear maps from V to V are $(1, 1)$-tensors because they transform both covariantly and contravariantly,
- bilinear forms are $(0, 2)$-tensors because they transform covariantly twice.

More generally, if T is an (m, n)-tensor with $T^{ijk\cdots}_{rst\cdots}$ and $\widetilde{T^{abc\cdots}_{xyz\cdots}}$ are coordinate representations of T, then

$$\widetilde{T^{abc\cdots}_{xyz\cdots}} = (B^a_i B^b_j B^c_k \cdots) T^{ijk\cdots}_{rst\cdots} (F^r_x F^s_y F^t_z \cdots).$$

Notice that the contravariant components $B^a_i, B^b_j, B^c_k, \ldots$ have $i, j, k, \ldots$ as lower indices, while $R^{ijk\cdots}_{rst\cdots}$ has them as upper indices. We observe an opposite relation with the covariant components $F^r_x, F^s_y, F^t_z, \ldots$. The expression above cooresponds to the "has components that transform in a special, predictable way under a change of bases" part of the heuristic definition of a tensor. Superscript indices are reserved for contravariant components, while subscript indices are reserved for covariant components.

2 December 18-19, 2021

Source. eigenchris, *Tensors for Beginners*, Videos 10-11

So far, we have considered tensors heuristically as abstract objects that are invariant under a change of bases/coordinates. We now present a more powerful and general heuristic for tensors:

> "A *tensor* is the tensor product of some collection of vectors and covectors."

Recall that a linear map from a vector space V onto itself is a tensor. As a first example for illuminating this new heuristic, we will show that a linear map is the tensor product of a vector and a covector (in that order). We will define tensor products more formally on a later date.

2.1 More on Dual Spaces: Interpreting Row Vectors

In linear algebra, we usually identify vectors with columns vectors—with respect to some basis—namely

$$v = v^1 e_1 + v^2 e_2 + \cdots + v^n e_n \longleftrightarrow \begin{pmatrix} v^1 \\ v^2 \\ \vdots \\ v^n \end{pmatrix}$$

where $\{e_1, \ldots, e_n\}$ is a basis of V. But what do row vectors represent? An insightful interpretation is that they represent covectors in V^* with respect to the dual basis $\{\epsilon^1, \ldots, \epsilon^n\}$ of $\{e_1, \ldots, e_n\}$ where $\epsilon^i(e_j) = \delta_{ij}$. More concretely, we identify

$$\alpha = \alpha_1 \epsilon^1 + \cdots + \alpha_n \epsilon^n \longleftrightarrow \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}.$$

Now observe that

$$\begin{aligned}\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}\begin{pmatrix}v^1\\v^2\\\vdots\\v^n\end{pmatrix} &= \alpha_1 v^1 + \cdots + \alpha_n v^n,\\ &= \left(\alpha_1\epsilon^1 + \cdots + \alpha_n\epsilon^n\right)(v^1 e_1 + \cdots + v^n e_n),\\ &= \alpha(v),\end{aligned}$$

so our identification of row vectors with covectors is accurate.

Proposition 2.1 (Riesz Representation Theorem)**.** Suppose V is a finite-dimensional inner product space. If $f \in V^*$, then there exists a unique $u_f \in V$ such that $f(v) = \langle u_f, v\rangle$.

Proof. Let $\{e_1, \dots, e_n\}$ be an orthonormal basis of V. To find $u_f \in V$ such that $f(v) = \langle u_f, v\rangle$ for all $v \in V$, notice that

$$\begin{aligned}f(v) &= f\left(\sum_{i=1}^{n}\langle v, e_i\rangle e_i\right)\\ &= \sum_{i=1}^{n}\langle v, e_i\rangle f(e_i)\\ &= \sum_{i=1}^{n} v^i f(e_i).\end{aligned}$$

Let $u_f = f(e_1)e_1 + \cdots + f(e_n)e_n$, and observe that $\langle u_f, v\rangle = f(v)$. $\square$

2.2 Linear Maps as Vector-Covector Tensor Products

Let $v = v^i e_i$ and $\alpha = \alpha_i \epsilon^i$. By the vector-column and covector-row identifications, observe that

$$\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}\begin{pmatrix}v^1\\v^2\\\vdots\\v^n\end{pmatrix}$$

returns the scalar $\alpha(v)$, while

$$\begin{pmatrix}v^1\\v^2\\\vdots\\v^n\end{pmatrix}\begin{pmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{pmatrix}$$

returns the matrix

$$\begin{pmatrix}v^1\alpha_1 & v^1\alpha_2 & \cdots & v^1\alpha_n\\ v^2\alpha_1 & v^2\alpha_2 & \cdots & v^2\alpha_n\\ \vdots & \vdots & \ddots & \vdots\\ v^n\alpha_1 & v^n\alpha_2 & \cdots & v^n\alpha_n\end{pmatrix}.$$

Coincidentally, the matrix represents a linear map from V to V with respect to $\{e_1, \dots, e_n\}$. For now, assume that the (ordered) column-row product above represents something more abstract, namely the *tensor product*

$v \otimes \alpha$ between v and α. We see that $v \otimes \alpha$ "produces" a linear map. With a bit of flexibility, the converse is valid as well; let $L : V \to V$ be a linear map with matrix representation

$$\begin{pmatrix} L^1_1 & L^1_2 & \cdots & L^1_n \\ L^2_1 & L^2_2 & \cdots & L^2_n \\ \vdots & \vdots & \ddots & \vdots \\ L^n_1 & L^n_2 & \cdots & L^n_n \end{pmatrix}$$

with respect to $\{e_1, \ldots, e_n\}$. Observe that

$$\begin{pmatrix} L^1_1 & L^1_2 & \cdots & L^1_n \\ L^2_1 & L^2_2 & \cdots & L^2_n \\ \vdots & \vdots & \ddots & \vdots \\ L^n_1 & L^n_2 & \cdots & L^n_n \end{pmatrix} = \underbrace{L^1_1 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} + L^1_2 \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} + \cdots + L^n_n \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}}_{\text{sum of } n^2 \text{ many matrices}},$$

so let I^i_j denote the matrix where all entries are 0 except for the ith row, jth column entry which we set equal to 1. Notice that I^i_j represents $e_i \otimes \epsilon^j$, as

$$I^i_j = \underbrace{\begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}}_{1 \text{ in } i\text{th row, } j\text{th column}} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}.$$

Thus, we have $\left(L^i_j\right) = L^i_j e_i \otimes \epsilon^j$, which shows that a linear map from V to V is "equivalent" to a linear combination of vector-covector tensor products. We emphasize that L is not equivalent to a single vector-covector tensor product like $v \otimes \alpha$, but actually a *linear combination* of tensor products. In particular, the $e_i \otimes \epsilon^j$ seem to behave like basis vectors of some vector space. This is, in fact, accurate as we will see in our formal definition of tensor products. In the meantime, let us extract a couple properties about tensor products from our present example.

Let α and β be covectors and $v, w \in V$. By our vector-column and covector-row identifications and some matrix properties, we have

$$\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \left(\begin{pmatrix} \alpha_1 & \cdots & \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 & \cdots & \beta_n \end{pmatrix} \right) = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \begin{pmatrix} \alpha_1 & \cdots & \alpha_n \end{pmatrix} + \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \begin{pmatrix} \beta_1 & \cdots & \beta_n \end{pmatrix},$$

$$\left(\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} + \begin{pmatrix} w^1 \\ \vdots \\ w^n \end{pmatrix} \right) \begin{pmatrix} \alpha_1 & \cdots & \alpha_n \end{pmatrix} = \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \begin{pmatrix} \alpha_1 & \cdots & \alpha_n \end{pmatrix} + \begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \begin{pmatrix} \beta_1 & \cdots & \beta_n \end{pmatrix},$$

which gives us

$$v \otimes (\alpha + \beta) = v \otimes \alpha + v \otimes \beta, \tag{4}$$

$$(v + w) \otimes \alpha = v \otimes \alpha + v \otimes \beta, \tag{5}$$

respectively. In addition, if $c \in \mathbb{R}$, then

$$\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \begin{pmatrix} c\alpha_1 & \cdots & c\alpha_n \end{pmatrix} = c \cdot \left(\begin{pmatrix} v^1 \\ \vdots \\ v^n \end{pmatrix} \begin{pmatrix} \alpha_1 & \cdots & \alpha_n \end{pmatrix} \right) = \begin{pmatrix} cv^1 \\ \vdots \\ cv^n \end{pmatrix} \begin{pmatrix} \alpha_1 & \cdots & \alpha_n \end{pmatrix},$$

which gives us

$$v \otimes (c \cdot \alpha) = c \cdot \big(v \otimes \alpha\big) = (c \cdot v) \otimes \alpha. \tag{6}$$

Equations (4), (5), and (6) are the key relations that will motivate our formal definition of tensors. As a final remark, recall that linear maps transform covariantly and contravariantly. Interestingly, the realization of linear maps as vector-covector tensor products seems to suggest that the contravariant and covariant properties of linear maps comes from its vector and covector components, respectively.

3 December 20-24, 2021

Source. Lee, *Introduction to Smooth Manifolds*, Chapter 12

We now transition to a more abstract and rigorous framework for discussing tensor products. Previously, we saw that a linear map $L : V \to V$ is equivalent to a linear combination the $e_i \otimes \epsilon^j$ tensor products, where $\{e_1, \dots, e_n\}$ is a basis of V and $\{\epsilon^1, \dots, \epsilon^n\}$ the corresponding dual basis. It turns out that the collection of all $v \otimes \alpha$ for $v \in V$ and $\alpha \in V^*$ forms the vector space $V \otimes V^*$, which is the *tensor product of vector spaces* V and V^*. Note that $V \otimes V^*$ is not merely the Cartesian product $V \times V^*$ because it also possesses the relations seen in (4), (5), and (6). These properties extend to the tensor product of multiple vector spaces; if $V_1, \dots, V_k$ are finite-dimensional vector spaces, then $V_1 \otimes \cdots \otimes V_k$ satisfies

$$v_1 \otimes \cdots \otimes (v_i + v_i') \otimes \cdots \otimes v_k = (v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_k) + (v_1 \otimes \cdots \otimes v_i' \otimes \cdots \otimes v_k)\,, \tag{7}$$

$$v_1 \otimes \cdots \otimes (a \cdot v_i) \otimes \cdots \otimes v_k = a\,(v_1 \otimes \cdots \otimes v_i \otimes \cdots \otimes v_k)\,. \tag{8}$$

for $v_1, \dots, v_i, v_i', \dots, v_k \in V$ and $a \in \mathbb{R}$. In this note, we construct $V_1 \otimes \cdots \otimes V_k$ from $V_1 \times \cdots \times V_k$ and make our way back to results and definitions we have already seen.

3.1 A Refresher on Quotients of Vector Spaces

Source. Axler, *Linear Algebra Done Right*, Chapter 3

Before digging in, we review what it means to quotient a vector space.

Definition 3.1. Let V be a vector space and W a subspace of V. The *quotient V/W of V by W* is the collection of all equivalence classes on V by the equivalence relation $v \sim v'$ if and only if $v - v' \in W$ for $v, v' \in V$.

The following is an alternate, but equivalent formulation of V/W.

Proposition 3.2. If V is a vector space and W a subspace of V, then

$$V/W := \{v + W : v \in V\},$$

where $v + W := \{v + w : w \in W\}$ is the *affine translation of W by $v \in V$*.

In other words, V/W consists exactly of affine translations of W, so we can write $v + W$ to represent an equivalence class in V/W. We endow V/W with a vector space structure by defining addition and scalar multiplication on V/W as follows:

$$(v + W) + (v' + W) := (v + v') + W,$$
$$a \cdot (v + W) := a \cdot v + W,$$

where $v, v' \in V$ and $a \in \mathbb{R}$. An important function related to V/W is the *quotient map* $\Pi : V \to V/W$ defined by $\Pi(v) = v + W$.

Remark (Connections to Group Theory)**.** Note that vector spaces are additive abelian groups, so a subspace W of V is a normal subgroup of V. Hence, addition and scalar multiplication on V/W are well-defined.

The following is a powerful result we will rely on in a later proof.

Lemma 3.3 (Characteristic Property of Linear Maps)**.** Let V and U be vector spaces. If $T : V \to U$ is a linear map and W a subspace of $\ker T$, then there exists a unique linear map $\overline{T} : V/W \to U$ such that $\overline{T} \circ \Pi \equiv T$, where $\Pi : V \to V/W$ is the canonical quotient map by W.

Proof. If $v+W \in V/W$, then define $\overline{T}(v+W) := T(v)$. Note that $\overline{T}$ is well-defined: if $v+W = v'+W$ belong to V/W, then $v - v' = w$ for some $w \in W$. Since $W \subseteq \ker T$, we have

$$\begin{aligned} T(v-v') &= T(w), \\ T(v) - T(v') &= 0, \\ T(v) &= T(v'), \end{aligned}$$

so $\overline{T}$ is well-defined. We see that

$$\overline{T} \circ \Pi(v) = \overline{T}\,(v+W) = T(v)$$

for all $v \in V$, so $\overline{T} \circ \Pi = T$. To show uniqueness, suppose $\widehat{T} : V/W \to U$ is linear and $\widehat{T} \circ \Pi = T$. If $v + W \in V/W$, then notice that $\widehat{T}(v+W) = T \circ \Pi(v) = T(v) = \overline{T}(v)$. Hence, $\overline{T} \equiv \widehat{T}$. □

Remark. An important condition in the hypothesis is that $W \subseteq \ker T$; we require this to assert $\overline{T}$ is well-defined.

3.2 Constructing the Tensor Product of Vector Spaces

Suppose $V_1 \ldots, V_k$ are finite-dimensional vector spaces. If we squint hard enough, $V_1 \otimes \cdots \otimes V_k$ is roughly $V_1 \times \cdots \times V_k$ with the properties

$$(v_1, \ldots, v_i + v_i', \ldots, v_k) = (v_1, \ldots, v_i, \ldots, v_k) + (v_1, \ldots, v_i', \ldots, v_k), \tag{9}$$
$$(v_1, \ldots, a \cdot v_i, \ldots, v_k) = a \cdot (v_1, \ldots, v_i, \ldots, v_k). \tag{10}$$

However, these properties contradict addition and scalar multiplication if we were to treat $V_1 \times \cdots \times V_k$ as a vector space. To endow $V_1 \times \cdots \times V_k$ with identities (9) and (10), we define an equivalence relation on the *free vector space* $\mathcal{F}(V_1 \times \cdots \times V_k)$ of $V_1 \times \cdots \times V_k$.

The free vector space $\mathcal{F}(S)$ of an arbitrary set S is the collection of all *formal linear combinations of elements* in S, which are expressions of the form

$$a_1 \cdot s_1 + a_2 \cdot s_2 + \cdots + a_k \cdot s_k$$

where $a_i \in \mathbb{R}$ for all $1 \le i \le k$ and k is some positive integer. We emphasize that $+$ and $\cdot$ are just symbols; the most we can do is "combine like terms," such as

$$a \cdot s + b \cdot s = (a+b) \cdot s.$$

In short, S is a basis of $\mathcal{F}(S)$.

Now let $\mathcal{R}$ denote the subspace of $\mathcal{F}(V_1 \times \cdots \times V_k)$ generated by elements of the form

$$\begin{aligned} &(v_1, \ldots, v_i + v_i', \ldots, v_k) - (v_1, \ldots, v_i, \ldots, v_k) - (v_1, \ldots, v_i', \ldots, v_k), \\ &(v_1, \ldots, a \cdot v_i, \ldots, v_k) - a \cdot (v_1, \ldots, v_i, \ldots, v_k). \end{aligned}$$

Define the equivalence relation $\sim$ where $(v_1, \ldots, v_k) \sim (w_1, \ldots, w_k)$ if and only if

$$(v_1, \ldots, v_k) - (w_1, \ldots, w_k) \in \mathcal{R}.$$

We now have

$$\begin{aligned}(v_1, \ldots, v_i + v_i', \ldots, v_k) &\sim (v_1, \ldots, v_i, \ldots, v_k) + (v_1, \ldots, v_i', \ldots, v_k),\\ (v_1, \ldots, a \cdot v_i, \ldots, v_k) &\sim a \cdot (v_1, \ldots, v_i, \ldots, v_k),\end{aligned}$$

which are very similar to (9) and (10). We take $V_1 \otimes \cdots \otimes V_k$ to be the quotient $\mathcal{F}(V_1 \times \cdots \times V_k)/\mathcal{R}$ and $v_1 \otimes \cdots \otimes v_k$ the equivalence class containing $(v_1, \ldots, v_k)$. By our construction, $V_1 \otimes \cdots \otimes V_k$ satisfies (7) and (8) as desired.

3.3 Some Technical Results on Tensor Products

We list some notable results in Chapter 12.

Lemma 3.4 (Characteristic Property of the Free Vector Space)**.** Suppose S is an arbitrary set and W a vector space. If $A : S \to W$ is some arbitrary function, then A extends uniquely to a linear map $\overline{A} : \mathcal{F}(S) \to W$.

Proof. If $\sum_{i=1}^k a_i s_i$ is a formal linear combination in $\mathcal{F}(S)$, then define

$$\overline{A}\left(\sum_{i=1}^k a_i s_i\right) = \sum_{i=1}^k a_i A(s_i).$$

We know $\overline{A}$ is well-defined because linear combinations in $\mathcal{F}(S)$ are unique, and by construction, $\overline{A}$ is linear. We have $\overline{A}(s) = A(s)$ for all $s \in S$, so $\overline{A}$ extends A. To see that $\overline{A}$ is unique, suppose $\overline{B}$ is also a linear extension of A. Observe that

$$\overline{A}\left(\sum_{i=1}^k a_i s_i\right) = \sum_{i=1}^k a_i A(s_i) = \overline{B}\left(\sum_{i=1}^k a_i s_i\right)$$

for all $\sum_{i=1}^k a_i s_i$ in $\mathcal{F}(S)$, so $\overline{A} \equiv \overline{B}$. $\square$

Recall that S is a basis of $\mathcal{F}(S)$. Lemma 3.4 is simply the general heuristic that a linear map is uniquely determined by where it sends the basis of a vector space.

Definition 3.5. Let $V_1, \ldots, V_k, X$ be vector spaces. A map $A : V_1 \times \cdots \times V_k \to X$ is multilinear if

$$\begin{aligned}A(v_1, \ldots, v_i + v_i', \ldots, v_k) &= A(v_1, \ldots, v_i, \ldots, v_k) + A(v_1, \ldots, v_i', \ldots, v_k),\\ A(v_1, \ldots, a \cdot v_i, \ldots, v_k) &= a \cdot A(v_1, \ldots, v_i, \ldots, v_k)\end{aligned}$$

for all $(v_1, \ldots, v_i, \ldots, v_k)$ and $(v_1, \ldots, v_i', \ldots, v_k)$ in $V_1 \otimes \cdots \otimes V_k$ and $a \in \mathbb{R}$.

The identities listed above generalize those for bilinear maps and are very similar to (9) and (10). The following result establishes the connection between multilinear maps and tensor products.

Lemma 3.6 (Characteristic Property of the Tensor Product Space)**.** Let $V_1, \ldots, V_k, X$ be vector spaces. If $V_1, \ldots, V_k$ have finite-dimension and $A : V_1 \times \cdots \times V_k \to \mathbb{R}$ is a multilinear map, then A extends uniquely to a linear map $\widetilde{A} : V_1 \otimes \cdots \otimes V_k \to X$ such that $\widetilde{A} \circ \pi = A$, where $\pi : V_1 \times \cdots \times V_k \to V_1 \otimes \cdots \otimes V_k$ is the map

$$\pi(v_1, \ldots, v_k) := v_1 \otimes \cdots \otimes v_k.$$

Proof. By Lemma 3.4, A extends uniquely to a linear map $\overline{A} : \mathcal{F}(V_1 \times \cdots \times V_k) \to \mathbb{R}$. Observe that $\mathcal{R} \subseteq \ker \overline{A}$; we have

$$\begin{aligned} &\overline{A}\big((v_1, \ldots, v_i + v_i', \ldots, v_k) - (v_1, \ldots, v_i, \ldots, v_k) - (v_1, \ldots, v_i', \ldots, v_k)\big) \\ &= \overline{A}(v_1, \ldots, v_i + v_i', \ldots, v_k) - \overline{A}(v_1, \ldots, v_i, \ldots, v_k) - \overline{A}(v_1, \ldots, v_i', \ldots, v_k) \\ &= A(v_1, \ldots, v_i + v_i', \ldots, v_k) - A(v_1, \ldots, v_i, \ldots, v_k) - A(v_1, \ldots, v_i', \ldots, v_k). \end{aligned}$$

If A is multilinear, then $A(v_1, \ldots, v_i + v_i', \ldots, v_k) - A(v_1, \ldots, v_i, \ldots, v_k) - A(v_1, \ldots, v_i', \ldots, v_k) = 0$. And so, we have $(v_1, \ldots, v_i + v_i', \ldots, v_k) - (v_1, \ldots, v_i, \ldots, v_k) - (v_1, \ldots, v_i', \ldots, v_k) \in \ker \overline{A}$. Similarly, we find that $(v_1, \ldots, a \cdot v_i, \ldots, v_k) - a \cdot (v_1, \ldots, v_i, \ldots, v_k) \in \ker \overline{A}$. Thus, it follows that $\mathcal{R} \subseteq \ker \overline{A}$.

Now let $\Pi : \mathcal{F}(V_1 \times \cdots \times V_k) \to V_1 \otimes \cdots \otimes V_k$ denote the quotient map by $\mathcal{R}$. Then by Lemma 3.3, there exists a unique linear map $\widetilde{A} : V_1 \otimes \cdots \otimes V_k \to \mathbb{R}$ such that $\widetilde{A} \circ \Pi = \overline{A}$. Let ι denote the inclusion $V_1 \times \cdots \times V_k \hookrightarrow \mathcal{F}(V_1 \times \cdots \times V_k)$ and observe that $\Pi \circ \iota = \pi$ and $\overline{A} \circ \iota = A$. Thus, $\widetilde{A} \circ \pi = A$. ☐

In rough terms, Lemma 3.6 tells us that multilinear maps on $V_1 \times \cdots \times V_k$ are really linear maps on $V_1 \otimes \cdots \otimes V_k$. The propositions that follow illustrate clever uses of this result. Now buckle up because the first one is a storm of indices!

Proposition 3.7. Let $V_1, \ldots, V_k$ be finite-dimensional vector spaces with dimension $d_1, \ldots, d_k$. If $\left\{e_1^{(k)}, \ldots, e_{d_k}^{(k)}\right\}$ is the basis of V_k, then

$$\mathcal{C} = \left\{e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_k}^{(k)} : 1 \leq i_1 \leq d_1, \ldots, 1 \leq i_k \leq d_k\right\}$$

is a basis of $V_1 \otimes \cdots \otimes V_k$. Hence, $\dim(V_1 \otimes \cdots \otimes V_k) = d_1 \cdot \cdots \cdot d_k$.

Proof. For a pure tensor $v_1 \otimes \cdots \otimes v_k$, expressing v_i in terms of $\left\{e_1^{(i)}, \ldots, e_{d_i}^{(i)}\right\} \subseteq V_i$ for all $1 \leq i \leq n$ and applying (9) and (10) expresses $v_1 \otimes \cdots \otimes v_k$ in terms of the elements in $\mathcal{C}$, so $\operatorname{span}(\mathcal{C}) = V_1 \otimes \cdots \otimes V_k$. To see that the elements are linearly independent, suppose that

$$a^{i_1, \ldots, i_k} e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_k}^{(k)} = 0.$$

(thanks Einstein!) Let $\left\{\epsilon_{(i)}^1, \ldots, \epsilon_{(i)}^k\right\}$ be the dual basis of $\left\{e_1^{(i)}, \ldots, e_{d_i}^{(i)}\right\}$. For any k-tuple $(m_1, \ldots, m_k)$ of indices, define the multilinear map $A^{m_1, \ldots, m_k} : V_1 \otimes \cdots \otimes V_k \to \mathbb{R}$ by

$$A^{m_1, \ldots, m_k}(v_1, \ldots, v_k) = \epsilon_{(1)}^{m_1}(v_1) \cdot \cdots \cdot \epsilon_{(k)}^{m_k}(v_k).$$

By Lemma 3.6, $A^{m_1, \ldots, m_k}$ extends uniquely to a linear map $\widetilde{A}^{m_1, \ldots, m_k}$ on $V_1 \otimes \cdots \otimes V_k$. We have

$$\widetilde{A}^{m_1, \ldots, m_k}\left(a^{i_1, \ldots, i_k} e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_k}^{(k)}\right) = a^{i_1, \ldots, i_k} \widetilde{A}^{m_1, \ldots, m_k}\left(e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_k}^{(k)}\right) = 0.$$

Note that

$$\begin{aligned} \widetilde{A}^{m_1, \ldots, m_k}\left(e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_k}^{(k)}\right) &= \widetilde{A}^{m_1, \ldots, m_k} \circ \pi\left(e_{i_1}^{(1)}, \ldots, e_{i_k}^{(k)}\right), \\ &= A\left(e_{i_1}^{(1)}, \ldots, e_{i_k}^{(k)}\right), \\ &= \epsilon_{(1)}^{m_1}\left(e_{i_1}^{(1)}\right) \cdot \cdots \cdot \epsilon_{(k)}^{m_1}\left(e_{i_k}^{(k)}\right), \\ &= \delta_{i_1}^{m_1} \cdot \cdots \cdot \delta_{i_k}^{m_k}. \end{aligned}$$

Thus, we have

$$\begin{aligned} a^{i_1, \ldots, i_k} \widetilde{A}^{m_1, \ldots, m_k}\left(e_{i_1}^{(1)} \otimes \cdots \otimes e_{i_k}^{(k)}\right) &= a^{i_1, \ldots, i_k}\ \delta_{i_1}^{m_1} \cdot \cdots \cdot \delta_{i_k}^{m_k}, \\ 0 &= a^{m_1, \ldots, m_k}, \end{aligned}$$

which means the elements in $\mathcal{C}$ are linearly independent. $\square$

Proposition 3.8 (Associativity of Tensor Products)**.** If V_1, V_2, V_3 are finite-dimensional vector spaces with dimensions d_1, d_2, and d_3, respectively, then there are unique isomorphisms

$$V_1 \otimes (V_2 \otimes V_3) \cong V_1 \otimes V_2 \otimes V_3 \cong (V_1 \cong V_2) \cong V_3.$$

under which $v_1 \otimes (v_2 \otimes v_3)$, $v_1 \otimes v_2 \otimes v_3$, and $(v_1 \otimes v_2) \otimes v_3$ all correspond.

Proof. First, we show $V_1 \otimes (V_2 \otimes V_3) \cong V_1 \otimes V_2 \otimes V_3$. Consider the multilinear map $A : V_1 \otimes V_2 \otimes V_3 \to V_1 \otimes (V_2 \otimes V_3)$ defined by

$$A(v_1, v_2, v_3) := v_1 \otimes (v_2 \otimes v_3).$$

Lemma 3.6 tells us A extends uniquely to a linear map $\widetilde{A} : V_1 \otimes V_2 \otimes V_3 \to V_1 \otimes (V_2 \otimes V_3)$ such that $\widetilde{A} \circ \pi = A$. If $\left\{e_1^{(i)}, \dots, e_{d_i}^{(i)}\right\}$ is a basis of V_i for each $1 \leq i \leq 3$, then notice that

$$\left\{e_{i_1}^{(1)} \otimes \left(e_{i_2}^{(2)} \otimes e_{i_3}^{(3)}\right) : 1 \leq i_1 \leq d_1, 1 \leq i_2 \leq d_2, 1 \leq i_3 \leq d_3\right\}$$

is a basis of $V_1 \otimes (V_2 \otimes V_3)$, and we have $\dim(V_1 \otimes (V_2 \otimes V_3)) = d_1 d_2 d_3 = \dim(V_1 \otimes V_2 \otimes V_3)$. Observe that

$$\begin{aligned} \widetilde{A}\left(e_{i_1}^{(1)} \otimes e_{i_2}^{(2)} \otimes e_{i_3}^{(3)}\right) &= A\left(e_{i_1}^{(1)}, e_{i_2}^{(2)}, e_{i_3}^{(3)}\right), \\ &= e_{i_1}^{(1)} \otimes \left(e_{i_2}^{(2)} \otimes e_{i_3}^{(3)}\right), \end{aligned}$$

so $\widetilde{A}$ is a bijection between the bases of $V_1 \otimes V_2 \otimes V_3$ and $V_1 \otimes (V_2 \otimes V_3)$ and $\widetilde{A}$ is an isomorphism. By a similar argument, we have $V_1 \otimes V_2 \otimes V_3 \cong (V_1 \otimes V_2) \otimes V_3$. $\square$