Multilinear Algebra
For the Undergraduate Algebra Student

Davis Shurbert
Department of Mathematics and Computer Science
University of Puget Sound

April 14, 2014

1 Introduction

When working in the field of linear algebra, it is natural to question whether or not we can expand our study to include linear transformations of multiple variables. Indeed, calculus students quickly learn that they can extend their tools for one variable functions to multivariate, vector valued functions. In this vein, we can extend the notion of a linear transformation to multilinear transformations.

A specific type of multilinear transformations are called *tensors*. In the study of tensors, we aim to examine the relationship between a multilinear transformation and representing that function as a multidimensional array. According to our intuition, we would like to have a similar method of representation as we have for linear transformations. Indeed, since linear transformations are by definition a type of tensor, whatever relationship we discover must be consistent with our understanding of this small dimensional case.

There are many approaches to the definition and study of tensors, as they are used in a wide array of mathematical fields. This paper takes an abstract mathematical approach to tensors, realizing them as multilinear transformations from vector spaces to the real numbers. As to not lose our intuition, we explore the connections between realizing tensors as multilinear functions and representing tensors as multidimensional arrays. Finally, for the sake of simplicity we will only consider vector spaces over the real numbers $\mathbb{R}$, leaving the extension to $\mathbb{C}$ as further research for the reader.

2 Dual Space

In order to define tensors, we must cover a few basic concepts in linear algebra. To begin, consider the set of all linear functions from a vector space V to $\mathbb{R}$, denoted $L(V : \mathbb{R})$. Recall that in order for $f : V \mapsto \mathbb{R}$ to be linear, it must be the case that

$$f(\alpha\vec{u} + \beta\vec{v}) = \alpha f(\vec{u}) + \beta f(\vec{v}), \text{ for all } \vec{u}, \vec{v} \in V.$$

If we treat $\mathbb{R}$ as a 1-dimensional vector space, then we can easily see that all elements of $L(V : \mathbb{R})$ are just linear transformations from one vector space to another. From previous work, we also know that such a set must form a vector space with $\dim L(V : \mathbb{R}) = \dim V \cdot \dim \mathbb{R} = \dim V$. We call $L(V : \mathbb{R})$ the *dual space* of V and we denote it as V^*

While the notation is similar to that of taking the conjugate transpose of a vector, this similarity has a purpose beyond confusing the reader. If we treat V as a space of column vectors, then $L(V : \mathbb{R}) = V^*$ can be represented as a space of row vectors with function input defined as matrix multiplication of a row vector by a column vector. That is, we can realize any $f(x)$ in $L(V : \mathbb{R})$ as fx for some row vector f, and all input column vectors x. In addition, since the dimensions of V and V^* are equal, we know they are isomorphic as vector spaces. Thus, we can realize the act of taking the conjugate transpose of a vector as an isomorphism between V and V^*.

While elements of V are called vectors, elements of V^* are called *covectors* in their relationship to V. We denote arbitrary covectors as v^*, while elements of V remain denoted v. This notation is distinct from our usual use of v^* in that v^* is no longer associated to a specific vector in V. We make this distinction because while the conjugate transpose forms an isomorphism between V and V^*, it may not be the only isomorphism and we need to account for this in order to maintain full generality.

We denote the value of v^* evaluated at v by $\langle v^*, v\rangle$. This notation purposefully mimics that of the inner product, and gives us our first multilinear function,

$$\langle\ ,\ \rangle : V^* \times V \mapsto \mathbb{R}.$$

Because we have equated $\langle v^*, v\rangle$ with matrix multiplication of a row vector by a column vector, it is no mistake that this bilinear function is almost identical to the standard inner product. However, we keep this distinction in order to preserve generality, and the similarity between the two operations brings us to our first theorem.

Theorem 1.1. *For a vector space V, and any inner product*

$$\cdot : V \times V \mapsto \mathbb{R},$$

there exists a unique isomorphism $G : V \mapsto V^*$ defined as

$$G(v) = v^*(x) = v \cdot x$$

Proof. As a direct result of the definition of an inner product, G must be a linear transformation. This motivates our investigation of the kernel of G, $K(G) = \{v \in V | K(G) = 0_{V^*}\}$. Recall that the zero element in V^* is the linear transformation which takes all inputs to $0_{\mathbb{R}}$. 0_V must be in the kernel since $G(0_V) = 0^*_V(x) = 0_V \cdot x = 0_{\mathbb{R}}$ for all $x \in V$. Now, for any nonzero $v \in V$, we have

$$(G(v))(v) = v^*(v) = v \cdot v \neq 0,$$

due to the fact that the inner product of a nonzero vector with itself is nonzero. This shows that for any $v \neq 0$, there exists an input of $G(v)$ that is not sent to the zero vector. This directly implies that $G(v) \neq 0_{V^*}$ and thus $K(G) = \{0_V\}$. Because G is a linear transformation, this is enough to ensure that G is injective.

We now focus our attention on the surjectivity of G. Because G is a linear transformation from V to V^*, two isomorphic vector spaces, we can use the fundamental result of linear transformations that domain dimension is equal to the sum of rank and nullity. This allows us to conclude that our linear transformation must have full rank, and is thus invertible. $\square$

We can think of the isomorphism G from above as the canonical isomorphism between V and V^*, and we will use G to reference this isomorphism for the rest of our discussion. Indeed, if our inner product on V is simply the standard dot product, the difference between V and V^* becomes a nuisance rather than informative. However, our construction of the dual space is currently independent of the inner product assigned to V, allowing our results to hold regardless of this choice.

2.1 Dual Basis

Now that we have a solid understanding of the relationship between V and its dual space V^*, we aim to find a canonical basis for V^*. Given a basis $B = \{e_i\}_{i=1}^N$, we define the dual basis of V^* relative to B as the set $\{e^j\}_{j=1}^N$ such that

$$\langle e^j, e_i \rangle = \delta_i^j = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}.$$

Notationally, e^i denotes an index on covectors, rather than powers of vectors which we cannot make sense of at the moment. Furthermore, δ_i^j is called the *Kronecker delta*. While it is not obvious, given a basis B of V we can always find a unique dual basis of V^* relative to B.

3 Multilinear Functions

Now that we have built up some machinery for dealing with vector spaces and their duals, we can focus our attention to the study of multilinear functions. If we have a collection of s vector spaces $\{V_i\}_{i=1}^s$, then we define an s-linear function f as a function

$$f : V_1 \times V_2 \times \cdots \times V_s \mapsto \mathbb{R}$$

that is linear in all of its components. That is, if we fix all inputs of $f(v_1, v_2, \ldots, v_s)$ except for v_i, then f becomes a linear function of V_i. That is, for all $i \in \{1, \ldots, s\}$,

$$f(v_1, \ldots, \alpha u_i + \beta v_i, \ldots, v_s) = \alpha f(v_1, \ldots, u_i, \ldots, v_s) + \beta f(v_1, \ldots, v_i, \ldots, v_s).$$

Of course the concept of a multilinear function holds in the case where our codomain is not $\mathbb{R}$, but for the purpose of this paper and for the sake of intuition we ignore the generalized case. Before moving on to the specific example of a tensor, we will now explore some examples of multilinear functions.

Example 2.1. Consider the function $f : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ defined as

$$f(x, y, z) = xyz.$$

To show that f is 3-linear, we begin by fixing x and z and treat f as a function of one variable. Consider,

$$\begin{aligned} f(x, \alpha y_1 + \beta y_2, z) &= x(\alpha y_1 + \beta y_2)z \\ &= x(\alpha y_1)z + x(\beta y_2)z \\ &= \alpha x y_1 z + \beta x y_2 z \\ &= \alpha f(x, y_1, z) + \beta f(x, y_2, z). \end{aligned}$$

We have just shown that f is linear in y when we fix x and z, but we can use a similar argument to show that f is linear in all variables when the remaining inputs are fixed. Alternatively, since f is a symmetric function, linearity in one variable implies linearity in all variables.

Example 3.1. Consider any collection of s vector spaces $\{V_i\}_{i=1}^s$, and a collection of linear functions $\{A_i\}_{i=1}^s$ where $A_i : V_i \mapsto \mathbb{R}$. We can define a multilinear function $A^s : V_1 \times V_2 \times$

$\cdots \times V_s \mapsto \mathbb{R}$ as

$$A^s(v_1, v_2, \ldots, v_s) = \prod_{i=1}^{s}(A_i(v_i)).$$

It is relatively straightforward to show that A^s is s-linear, and we leave verification to the reader as an exercise.

4 Tensors

Restricting our focus to a specific type of multilinear function, we now give our definition of a tensor. For a vector space V, and positive integers p and q, we define a *tensor of order* (p, q) on V as a $(p+q)$-linear function

$$T : \underbrace{V^* \times \cdots \times V^*}_{p \text{ times}} \times \underbrace{V \times \cdots \times V}_{q \text{ times}} \mapsto \mathbb{R}.$$

In addition, we define a tensor of order $(0, 0)$ to be a scalar in $\mathbb{R}$. We denote the set of all tensors of order (p, q) on V as $\mathcal{T}^p_q(V)$. In fact, $\mathcal{T}^p_q(V)$ forms a vector space of dimension N^{p+q} (where $\dim(V) = N$) under the following operations. For any $\alpha \in \mathbb{R}$, and any $A, B \in \mathcal{T}^p_q(V)$,

$$(A+B)(v^1, \ldots, v^p, v_1, \ldots, v_q) = A(v^1, \ldots, v^p, v_1, \ldots, v_q) + B(v^1, \ldots, v^p, v_1, \ldots, v_q), \text{ and}$$
$$(\alpha A)(v^1, \ldots, v^p, v_1, \ldots, v_q) = \alpha(A(v^1, \ldots, v^p, v_1, \ldots, v_q))$$

While the fact that $\mathcal{T}^p_q(V)$ forms a vector space is apparent when we consider it as a set of all linear transformations from one vector space to another, we leave a more in-depth discussion of the dimension of this space for later in the paper.

Similarly to the fact that $(v^*)^* = v$, we make no distinction between the dual of the dual of V, $(V^*)^*$, and V itself. This fact, along with our definition of V^*, allows us to conclude that

$$\mathcal{T}^0_1(V) = L(V : \mathbb{R}) = V^*, \text{ and } \mathcal{T}^1_0(V) = L(V^* : \mathbb{R}) = V.$$

We can now begin to understand the relationship between $(p+q)$-dimensional arrays and tensors of order (p, q), as elements of V and V^* both have 1-dimensional array representations. Before we investigate this type of representation for higher order tensors, we now explore a few more examples of tensors.

Example (Bowen, Wang) 4.1. Given any linear transformation $A : V \mapsto V$, we can define a bilinear function $\hat{A} : V^* \times V \mapsto \mathbb{R}$ as

$$\hat{A}(u^*, v) = \langle u^*, Av \rangle.$$

We have already established that $\langle , \rangle : V^* \times V \mapsto \mathbb{R}$ is bilinear. This fact, together with the linearity of the matrix vector product allows us to conclude the following.

$$\hat{A}(\alpha u_1^* + \beta u_2^*, v) = \langle \alpha u_1^* + \beta u_2^*, v \rangle = \alpha \langle u_1^*, v \rangle + \beta \langle u_2^*, v \rangle = \alpha \hat{A}(u_1^*, v) + \beta \hat{A}(u_2^*, v), \text{ and}$$
$$\hat{A}(u^*, \alpha v_1 + \beta v_2) = \langle u^*, \alpha v_1 + \beta v_2 \rangle = \alpha \langle u^*, v_1 \rangle + \beta \langle u^*, v_2 \rangle = \alpha \hat{A}(u^*, v_1) + \beta \hat{A}(u^*, v_2)$$

This linearity qualifies $\hat{A}$ for membership in $\mathcal{T}_1^1(V)$. Note that in the case where our inner product on V is the dot product, and we consider V^* as a space of row vectors, we have $\hat{A}(u^*, v) = u^*Av$.

Proposition 5.1. *L(V:V) is isomorphic to $\mathcal{T}_1^1(V)$, and $Z : L(V : V) \mapsto \mathcal{T}_1^1(V)$ defined as*

$$Z(A) = \hat{A}(u^*, v) = \langle u^*, Av\rangle$$

is an isomorphism.

Proof. Presuming that $\mathcal{T}_1^1(V)$ forms a vector space, we first show that Z is a linear transformation. For any $\alpha, \beta \in \mathbb{R}$ and any $A, B \in L(V : V)$, consider

$$\begin{aligned} Z(\alpha A + \beta B) &= (\alpha A \hat{+} \beta B)(u^*, v) = \langle u^*, (\alpha A + \beta B)v\rangle \\ &= \langle u^*, \alpha Av + \beta Bv\rangle = \alpha\langle u^*, Av\rangle + \beta\langle u^*, Bv\rangle \\ &= \alpha Z(A) + \beta Z(B). \end{aligned}$$

Establishing Z as a linear transformation gives us motivation for investigating the kernel of Z, $K(Z)$. Clearly, the zero linear transformation is in the kernel since $\langle u*, 0(v)\rangle = \langle u^*, 0_V\rangle = 0$ for all $u^* \in V^*$. Now, take any nonzero linear transformation $A \in L(V : V)$. Because A is nonzero, there must exist some $v \in V$ such that $Av = u$ is nonzero. Using our canonical isomorphism $G : V \mapsto V^*$, consider

$$Z(A)(G^{-1}(u), v) = \langle G^{-1}(u), Av\rangle = \langle G^{-1}(u), u\rangle = u \cdot u \neq 0.$$

Therefore, our kernel $K(Z)$ contains only the zero linear transformation and thus Z is an injective linear transformation. Since $\dim \mathcal{T}_1^1(V) = \dim V^{1+1} = \dim V \dim V = \dim L(V : V)$, this is enough to conclude that Z is an isomorphism. □

Example (Bowen, Wang) 6.1. Given any $v^* \in V^*$ and any $v \in V$, we can define the function $v \otimes v^* : V^* \times V \mapsto \mathbb{R}$ as

$$(v \otimes v^*)(u^*, u) = \langle u^*, v\rangle\langle v^*, u\rangle.$$

The membership of $v \otimes v^*$ in $\mathcal{T}_1^1(V)$ is left as an exercise to the reader. However, recall that this example is simply a case of Example 3.1 from above. Our isomorphism between $\mathcal{T}_1^1(V)$ and $L(V : V)$ allows us to conclude that $v \otimes v^*$ can be realized as a linear transformation from V to V such that

$$v \otimes v^*(u^*, u) = \langle u^*, (v \otimes v^*)u\rangle = \langle u^*, v\rangle\langle v^*, u\rangle.$$

We call $v \otimes v^*$ the *tensor product* of v and v^*, and we can represent $v \otimes v^*$ as the *Kronecker product* or *outer product* of v with $G^{-1}(v^*)$ (where G is our canonical isomorphism from V to V^*). If $v = (v_1, v_2, \ldots, v_N)$, and $G^{-1}(v^*) = u = (u_1, u_2, \ldots u_N)$, then the outer product of

v with u is defined as

$$v \otimes u = vu^T = \begin{pmatrix} v_1u_1 & v_1u_2 & \cdots & v_1u_N \\ v_2u_1 & v_2u_2 & \cdots & v_2u_N \\ \vdots & \vdots & \ddots & \vdots \\ v_Nu_1 & v_Nu_2 & \cdots & v_Nu_N \end{pmatrix}$$

Example 7.1. If we take $V = \mathbb{R}^2$, $v = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, and $v^* = \begin{bmatrix} 3 & 2 \end{bmatrix}$, we have

$$A \equiv \begin{bmatrix} 2 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 3 & 2 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ 3 & 2 \end{bmatrix}$$

Now, evaluate A at $(u^*, u) = (\begin{bmatrix} 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix})$ and investigate.

$$\begin{aligned} A(\begin{bmatrix} 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}) &= \langle \begin{bmatrix} 0 & 1 \end{bmatrix}, A\begin{bmatrix} 0 \\ 1 \end{bmatrix} \rangle = \langle \begin{bmatrix} 0 & 1 \end{bmatrix}, \begin{bmatrix} 6 & 4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rangle \\ &= \langle \begin{bmatrix} 0 & 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \end{bmatrix} \rangle = \langle \begin{bmatrix} 0 & 1 \end{bmatrix}, 2\begin{bmatrix} 2 \\ 1 \end{bmatrix} \rangle \\ &= \langle \begin{bmatrix} 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \rangle 2 = \langle \begin{bmatrix} 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \rangle \langle \begin{bmatrix} 3 & 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rangle \\ &= \langle u^*, v \rangle \langle v^*, u \rangle \end{aligned}$$

This property holds for all other combinations (u^*, u) of basis vectors of V^* and V respectively. However this can be shown in an entirely similar manner and is left as an exercise to the reader. The linearity of A ensures that if this property holds for all combinations of basis vectors, then it holds for all $u^* \in V^*$ and all $u \in V$.

5 Tensor Product

We can extend our idea of the tensor product to handle an arbitrary number of vectors and covectors. Given a vector space V, a collection of vectors $\{v_i\}_{i=1}^p$, and a collection of covectors $\{v^j\}_{j=1}^q$, we can build a function

$$v_1 \otimes \cdots \otimes v_p \otimes v^1 \otimes \cdots \otimes v^q : \underbrace{V^* \times \cdots \times V^*}_{p \text{ times}} \times \underbrace{V \times \cdots \times V}_{q \text{ times}} \mapsto \mathbb{R}$$

defined as

$$(v_1 \otimes \cdots \otimes v_p \otimes v^1 \otimes \cdots \otimes v^q)(u^1, \ldots u^p, u_1, \ldots u_q) = \langle u^1, v_1 \rangle \cdots \langle u^p, v_p \rangle \langle v^1, u_1 \rangle \cdots \langle v^q, u_q \rangle$$

for all $\{u_i\}_{i=1}^p \in V$ and all $\{u^j\}_{j=1}^q \in V^*$. Again, the linearity of this function follows from Example 3.1. We call such a construction the *tensor product* of the vectors $\{v_i\}_{i=1}^p \cup \{v^j\}_{j=1}^q$, and we use this definition to form a basis for $\mathcal{T}^p_q(V)$. Fix a basis of V $\{e_i\}_{i=1}^N$, and let $\{e^j\}_{j=1}^N$ be the corresponding dual basis of V^*. It is the case that the set

$$B = \{e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes e^{j_1} \otimes \cdots \otimes e^{j_q}\}$$

for all combinations of $\{i_k\}_{k=1}^p \in \{1, \ldots, N\}$ and $\{j_z\}_{z=1}^q \in \{1, \ldots, N\}$, forms a basis of $\mathcal{T}^p_q(V)$. Note that B has cardinality N^{p+q} as there are $p+q$ indices, each ranging from 1 to N.

Example 8.1. Before we prove that the above B is a basis for $\mathcal{T}^p_q(V)$, we will first demonstrate that we can form a basis for $\mathcal{T}^1_1(V)$. Let $\dim V = N$, pick a basis of V, $\{e_i\}_{i=1}^N$, and find the corresponding dual basis for V^*, $\{e^j\}_{j=1}^N$. Consider the following relation of linear dependence

$$\sum_{i=1}^N \sum_{j=1}^N \alpha_i^j e_i \otimes e^j = \mathcal{O},$$

where $\{\alpha_i^j\}_{i,j=1}^{N,N}$ are scalars and $\mathcal{O}$ denotes the zero tensor which takes all inputs to $0_{\mathbb{R}}$. Now fix $z, k \in \{1, \ldots, N\}$ and investigate

$$\begin{aligned} 0_{\mathbb{R}} = (\sum_{i=1}^N \sum_{j=1}^N \alpha_i^j e_i \otimes e^j)(e^k, e_z) &= \sum_{i=1}^N \sum_{j=1}^N [(\alpha_i^j e_i \otimes e^j)(e^k, e_z)] \\ = \sum_{i=1}^N \sum_{j=1}^N \alpha_i^j \langle e^k, e_i\rangle \langle e^j, e_z\rangle &= \sum_{i=1}^N \sum_{j=1}^N \alpha_i^j \delta_i^k \delta_z^j. \end{aligned}$$

Recall that δ_j^i is the Kronecker delta and is defined on page 4.

So, we have a double sum ranging over the indices i and j, the terms of which all contain the product $\delta_i^k \delta_z^j$. By definition, $\delta_i^k \delta_z^j = 0$ unless $i = k$ and $j = z$, in which case $\delta_i^k \delta_z^j = 1$. This allows us to reduce our sum to

$$\sum_{i=1}^N \sum_{j=1}^N \alpha_i^j \delta_i^k \delta_z^j = \alpha_z^k$$

which we already know is equal to zero. If we repeat this process for all $z, k \in \{1, \ldots, N\}$, we can conclude that all scalars α_i^j are equal to 0. This implies that the set $\{e_i \otimes e^j\}_{i,j=1}^N$ is linearly independent.

Now that linear independence has been established for $C = \{e_i \otimes e^j\}_{i,j=1}^{N,N}$, we must show that C spans $\mathcal{T}^1_1(V)$. First, define the scalars $\alpha_z^k = A(e^k, e_z)$. We already know that $(\sum_{i=1}^N \sum_{j=1}^N \alpha_i^j e_i \otimes e^j)(e^k, e_z) = \alpha_z^k$, which by definition is equal to $A(e^k, e_z)$. This result tells us that for any combination of basis vectors e_i and e^j, we can realize $A(e_i, e^j)$ as a linear combination of elements from C evaluated at (e_i, e^j). The multilinearity of A allows us to conclude that for any $v \in V$ and $v^* \in V^*$, we can realize $A(v^*, v)$ as a a linear combination of elements from C evaluated at (v^*, v). That is for some scalars β_i^j,

$$A(v^*, v) = (\sum_{i=1}^{N} \sum_{j=1}^{N} \beta_i^j e_i \otimes e^j)(v^*, v).$$

This implies that C spans $\mathcal{T}_1^1(V)$, and we have found our desired basis.

The case of $\mathcal{T}_q^p(V)$ with q or p greater than one can be shown using entirely similar proof techniques. However, completing this proof in full generality becomes very cumbersome notationally. Furthermore, the abridged proof can be found in [1], which is freely available.

While we have shown that tensors of the form $v_1 \otimes \cdots \otimes v_p \otimes v^1 \otimes \cdots \otimes v^q$ form a generating set for $\mathcal{T}_q^p(V)$, not every tensor in $\mathcal{T}_q^p(V)$ can be realized as the tensor product of vectors and covectors. For example, if we let $V = \mathbb{R}^2$ and define the identity tensor in $\mathcal{T}_1^1(V)$ as

$$I(v^*, v) = \langle v^*, I_V v \rangle = \langle v^*, v \rangle$$

where I_V denotes the identity transformation from $V \mapsto V$, we can see that I_V cannot be constructed as the Kronecker product of a vector and a covector. This is due to the fact that the Kronecker product must yield singular matrices, as each row in $v \otimes v^*$ is a scalar multiple of v^* by entries in v. However, we can easily compute a basis of $L(V : V) \cong L(V^* \times V : \mathbb{R}) \cong \mathcal{T}_1^1(V)$, using the procedure outlined in Example 8.1 and our realization of the tensor product of vectors as the Kronecker product.

$$\begin{aligned}
e_1 \otimes e^1 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
e_1 \otimes e^2 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\
e_2 \otimes e^1 &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \\
e_2 \otimes e^2 &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}$$

6 Tensor Components

While we have discussed how to realize matrices, vectors, and covectors as tensors, we now focus our attention to representing tensors as indexed components. In particular, we would like to expand our definition of the Kronecker product to encompass any number of vectors, matrices, or higher order tensors. For any A in $\mathcal{T}_q^p(V)$, we define the $N^{(p+q)}$ components of A to be the scalars

$$A_{j_1,\ldots,j_q}^{i_1,\ldots,i_p} = A(e^{i_1}, \ldots, e^{i_p}, e_{j_1}, \ldots, e_{j_q}).$$

This form is similar to the matrix representation M of a linear transformation T relative to a given basis B, as we build the rows of M by evaluating T at each element of B. Furthermore, just as any matrix yields a linear transformation, any indexed collection of scalars $A_{j_1,\ldots,j_q}^{i_1,\ldots,i_p}$ yields a tensor in $\mathcal{T}_q^p(V)$.

Example 9.1. Let $V = \mathbb{R}^2$ and consider the identity tensor $\mathcal{I}$ in $\mathcal{T}_1^1(V)$ defined as

$$\mathcal{I}(v^*, v) = \langle v^*, v\rangle.$$

We can construct the scalars $\mathcal{I}_j^i$ in row i column j using our definition of tensor components.

$$\mathcal{I}_j^i = \mathcal{I}(e^i, e_j) = \langle e^i, e_j\rangle = \delta_j^i$$

So, our component representation of $\mathcal{I}$ is the matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

as expected.

Before exploring a tensor representation which requires more than 2 indices, we must generalize the Kronecker product. For two matrices $A_{m\times n}$ and $B_{p\times q}$, the Kronecker product of A and B, $A \otimes B$ is defined as

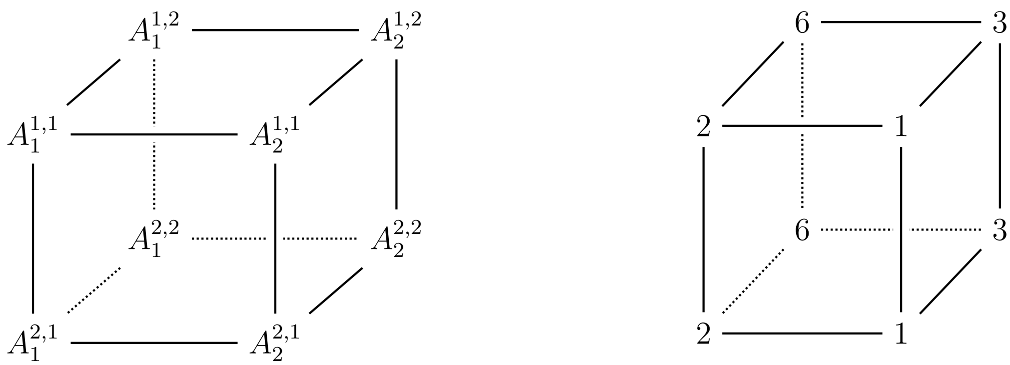
$$A \otimes B = \begin{pmatrix} [A]_{1,1}B & [A]_{1,2}B & \cdots & [A]_{1,n}B \\ [A]_{2,1}B & [A]_{2,2}B & \cdots & [A]_{2,n}B \\ \vdots & \vdots & \ddots & \vdots \\ [A]_{m,1}B & [A]_{m,2}B & \cdots & [A]_{m,n}B \end{pmatrix}$$

While this product can be treated as an array of dimension 2, we will primarily focus on the Kronecker product of a matrix or tensor by a vector, and treat the resulting object as a higher dimensional array.

Example 10.1. For $V = \mathbb{R}^2$, consider the vectors $u = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $v^* = \begin{bmatrix} 2 & 1 \end{bmatrix}$, and $w^* = \begin{bmatrix} 1 & 3 \end{bmatrix}$. We can use our standard basis together with its dual basis to construct the component representation of the tensor $A = (u \otimes v^* \otimes w^*) \in \mathcal{T}_1^2$.

$$A_1^{1,1} = A(e^1, e_1, e_1) = \langle\begin{bmatrix} 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\rangle\langle\begin{bmatrix} 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\rangle\langle\begin{bmatrix} 1 & 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\rangle = 2$$
$$A_1^{2,1} = A(e^1, e_2, e_1) = \langle\begin{bmatrix} 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\rangle\langle\begin{bmatrix} 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\rangle\langle\begin{bmatrix} 1 & 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\rangle = 1$$
$$A_1^{1,2} = A(e^1, e_1, e_2) = \langle\begin{bmatrix} 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\rangle\langle\begin{bmatrix} 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\rangle\langle\begin{bmatrix} 1 & 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\rangle = 6$$

Continuing in this manner, we can calculate the rest of our 8 components to arrive at the component representation of A.



Now, we compare this component representation of the tensor $A = (u \otimes v^* \otimes w^*)$ with the Kronecker product $u \otimes v^* \otimes w^*$. Since the Kronecker product is associative, we have

$$\begin{aligned} u \otimes v^* \otimes w^* &= \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 2 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 2\begin{bmatrix} 1 & 3 \end{bmatrix} & 1\begin{bmatrix} 1 & 3 \end{bmatrix} \\ 2\begin{bmatrix} 1 & 3 \end{bmatrix} & 1\begin{bmatrix} 1 & 3 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 2 & 6 \\ 2 & 6 \end{bmatrix} & \begin{bmatrix} 1 & 3 \\ 1 & 3 \end{bmatrix} \end{bmatrix}, \end{aligned}$$

which is identical to our construction of the components of A from above.

We can now represent all simple tensors as Kronecker products of vectors and covectors. While this is a powerful tool, we must realize the distinction between tensors and higher dimensional arrays. Every tensor has a component representation, and every collection of indexed scalars corresponds to a tensor. However, decomposition of tensors into the tensor product of vectors and covectors is not always possible, nor is it necessarily unique. For example, because of the multilinearity of $\langle , \rangle : V^* \times V \mapsto \mathbb{R}$ we have

$$(2v \otimes \frac{1}{2}v^*)(u^*, u) = \langle u^*, 2v \rangle \langle \frac{1}{2}v, u \rangle = \langle u^*, v \rangle \langle v, u \rangle = (v \otimes v^*)(u^*, u).$$

7 Conclusions and Further Research

While our discussion so far has been mathematically rich, we are simply studying the fundamentals of tensors and tensor algebra. In addition, we have left many important questions unanswered. Now that we have realized tensors as higher dimensional arrays, can we come up with an effective definition of tensor multiplication? We can extend the tensor product to accept general tensors as inputs, however this results in a higher order tensor. If we have a suitable notion of tensor multiplication, can we decompose tensors into pieces which have "nice" properties? Are there useful definitions of the determinant, trace, or eigenvalues? These question have complicated answers, as for example there are multiple distinct notions of eigenvalues for higher dimensional tensors. While we do not cover these questions in the context of this paper, relevant papers can be found in the references.

The logical next step for those who are interested is to study *contractions* of tensors. Contractions act as linear functions taking tensors of order (p, q) to tensors of order $(p-1, q-1)$. However, in the study of contractions we expand our definition of tensors as linear functions

$$T : V^* \times \cdots \times V^* \times V \times \cdots \times V \mapsto U$$

for some vector space U. Naturally, tensors become significantly more complicated when working with them in this context, however we can use this definition to develop a richer structure and understanding of tensors.

After establishing a strong mathematical understanding of tensors, one is free to explore the countless applications in both pure and applied fields in mathematics. While tensors are consistently used in the mathematical sciences, I am particularly interested in using them to understand hypergraphs. Just as we have expanded matrices into higher dimensions, we can realize hypergraphs (graphs which have multiple types of edges) as tensors. This field of study is currently used to analyze monumental data sets, such as social network graphs. For example, hypergraphs are used to study social network data because there can simultaneously exist many different relationships between two people. I can be Bob's roommate, friend, coworker, and fellow student, or we can live in the same geographical location, have the same hair color, or share economic status. Bob and I can share any subset of the above relations, and if we represent our relationship as a graph with Bob and I both occupying vertices, each of these relationships require their own type of edge. Such a hypergraph has an adjacency tensor, just as regular graphs have adjacency matrices, which is used to calculate and understand certain properties of the hypergraph.

Sources

Sources listed in the descending order based on the frequency of use in the paper.

1. INTRODUCTION TO VECTORS AND TENSORS, Linear and Multilinear Algebra, Volume 1. Ray M. Bowen, C.-C. Wang. Updated 2010.
   (Freely available at: http://repository.tamu.edu/bitstream/handle/1969.1/2502/IntroductionToVectorsAndTensorsVol1.pdf?s..)

2. Matrix Analysis for Scientists and Engineers, Alan J. Laub.
   (Relevant chapter at http://www.siam.org/books/textbooks/OT91sample.pdf)

3. Notes on Tensor Products and the Exterior Algebra, K. Purbhoo. Updated 2008
   (Lecture notes, available at http://www.math.uwaterloo.ca/ kpurbhoo/spring2012-math245/tensor.pdf)

4. Linear Algebra via Exterior Products. Sergei Winitzki. Copyright 2002.

5. Tensor Spaces and Exterior Algebra. Takeo Yokonuma. English version translated 1991.

Further Research

7. Multilinear Algebra. Werner Greub. Springer, updated 2013.

8. Eigenvalues of the Adjacency Tensor on Products of Hypergraphs. Kelly J. Pearson, Tan Zhang. October, 2012
   http://www.m-hikari.com/ijcms/ijcms-2013/1-4-2013/pearsonIJCMS1-4-2013.pdf