

On Tensors

Morgan Arnold

Compiled On: October 25, 2023

Contents

1	Vector Spaces & Dual Spaces	2
2	Tensors, Tensor Products, & Tensor Product Spaces	5

1 Vector Spaces & Dual Spaces

We begin our exploration of tensors with a review of the theory of vector spaces. Once we have reviewed relevant definitions and theorems regarding vector spaces, we will move to examining linear mappings from vectors to scalars. This will allow us to define the dual space corresponding to any vector space.

Definition 1 (Vector Spaces). *A vector space is composed of a field F and a set V , along with two binary operations called vector addition and scalar multiplication, typically denoted $\mathbf{u} + \mathbf{v}$ and $a\mathbf{u}$ respectively, where $a \in F$ and $\mathbf{u}, \mathbf{v} \in V$. The following axioms must also be satisfied:*

- $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$,
- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for all $\mathbf{u}, \mathbf{v} \in V$,
- there exists some element $\mathbf{0} \in V$ such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ for all $\mathbf{u} \in V$,
- for all $\mathbf{u} \in V$, there exists some element $-\mathbf{u} \in V$ such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$,
- $a(b\mathbf{u}) = (ab)\mathbf{u}$ for all $a, b \in F$ and $\mathbf{u} \in V$,
- $1\mathbf{u} = \mathbf{u}$ for all $\mathbf{u} \in V$, where 1 is the multiplicative identity of F ,
- $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ for all $a \in F$ and $\mathbf{u}, \mathbf{v} \in V$,
- $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$ for all $a, b \in F$ and $\mathbf{u} \in V$.

The elements of F are called scalars, and the elements of V are called vectors. V is called a vector space over F , or sometimes just a vector space if F is clear from context or not relevant.

In the event that the reader is not familiar with the definition of a field, and is interested in tensors for their applications to physics, little is lost by thinking of F as being one of \mathbb{R} or \mathbb{C} . In particular, in classical and relativistic physics it is typically the case that $F = \mathbb{R}$.

Notation 1 (Einstein Summation Convention). *Consider some indexed sets $\{a^1, \dots, a^n\}$ and $\{b_1, \dots, b_n\}$. The Einstein summation convention states that*

$$a^\mu b_\mu := \sum_{\mu=1}^n a^\mu b_\mu$$

In other words, if an unbound index variable occurs twice in a single term, at least once in superscript and at least once in subscript, summation is implied over all values of the index. In general relativity, the values of such indices are almost always $\{0, 1, 2, 3\}$, but in general it must be clear from context what the values taken on by the indices are. a^μ and b_μ are sometimes used to refer to the indexed sets themselves.

Definition 2 (Linear Dependence & Independence). Let V be a vector space over F . Let \mathbf{v}_μ be a subset of V . The vectors \mathbf{v}_μ are called linearly dependent if there exists a corresponding subset c^μ of F , not all zero, such that $c^\mu \mathbf{v}_\mu = 0$. \mathbf{v}_μ is called linearly independent if it is not linearly dependent.

Definition 3 (Bases). Let V be a vector space over F . A linearly independent subset \mathbf{e}_μ of V is called a basis if for all $\mathbf{v} \in V$ there exists a subset c^μ of F such that $\mathbf{v} = c^\mu \mathbf{e}_\mu$.

Theorem 1. Let V be a vector space. If \mathbf{u}_μ and \mathbf{v}_ν are finite bases of V , then $|\mathbf{u}_\mu| = |\mathbf{v}_\mu|$.

Proof. □

Definition 4 (Dimension). Let V be a vector space admitting a finite basis \mathbf{e}_μ . $|\mathbf{e}_\mu|$ is called the dimension of V , which we denote by $\dim V$.

While vector spaces which do not admit finite bases are of substantial interest in a variety of contexts in both mathematics and physics, the rest of this text will be concerned with vector spaces which do admit finite bases. From this point onwards in the text, all vector spaces should be taken to have finite dimension.

Notation 2 (Mappings). Given a mapping $f : A \rightarrow B$ and some $a \in A$, we sometimes denote the value of f at a , usually denoted $f(a)$, by $\langle f, a \rangle$.

Definition 5 (Linear Mappings). Let V be a vector space over F . A mapping $f : V \rightarrow F$ is called linear if for all $\mathbf{u}, \mathbf{v} \in V$ and $a \in F$, the following hold:

- $\langle f, \mathbf{u} + \mathbf{v} \rangle = \langle f, \mathbf{u} \rangle + \langle f, \mathbf{v} \rangle$
- $\langle f, a\mathbf{u} \rangle = a\langle f, \mathbf{u} \rangle$

Definition 6 (Isomorphism). Let U and V be vector spaces. A bijective linear mapping $f : U \rightarrow V$ is called an isomorphism. If any isomorphisms exist, U and V are called isomorphic.

Theorem 2. Let U and V be vector spaces. U and V are isomorphic if and only if $\dim U = \dim V$.

Proof. □

Definition 7 (Dual Space). Let V be a vector space over F . We call the set of all linear mappings $\mathbf{v}^* : V \rightarrow F$ the dual space of V , and denote it by V^* . V^* is a vector space over F once vector addition and scalar multiplication are defined as follows:

- $\langle \mathbf{u}^* + \mathbf{v}^*, \mathbf{v} \rangle = \langle \mathbf{u}^*, \mathbf{v} \rangle + \langle \mathbf{v}^*, \mathbf{v} \rangle$
- $\langle a\mathbf{v}^*, \mathbf{v} \rangle = a\langle \mathbf{v}^*, \mathbf{v} \rangle$

The elements of V^* are called covectors. Indexed sets of covectors are indexed with a superscript, rather than the subscript which is used for vectors.

Definition 8 (Dual Basis). *Let V be a vector space over F , and let V^* be its dual space. Given a basis \mathbf{e}_μ of V , we define a dual basis to be a set of $\mathbf{e}^\nu \in V^*$ which form a basis of V^* and satisfy $\langle \mathbf{e}^\nu, \mathbf{e}_\mu \rangle = \delta_\mu^\nu$.*

Theorem 3. *Let V be a vector space. Given a basis \mathbf{e}_μ of V , there exists a unique dual basis.*

Proof. □

Observe that the definition of the dual basis allows for the operation of a covector on a vector to be evaluated especially easily. Let V be a vector space over F , and let V^* be its dual space. Let \mathbf{e}_μ be a basis of V , then if $\mathbf{v} \in V$ we may write $\mathbf{v} = a^\mu \mathbf{e}_\mu$ for some a^μ . Similarly, let \mathbf{e}^ν be the dual basis of V^* corresponding to \mathbf{e}_μ , then if $\mathbf{v}^* \in V^*$ we may write $\mathbf{v}^* = b_\nu \mathbf{e}^\nu$ for some b_ν . Therefore, in all generality we have

$$\begin{aligned} \langle \mathbf{v}^*, \mathbf{v} \rangle &= \langle b_\nu \mathbf{e}^\nu, a^\mu \mathbf{e}_\mu \rangle \\ &= a^\mu b_\nu \langle \mathbf{e}^\nu, \mathbf{e}_\mu \rangle \\ &= a^\mu b_\nu \delta_\mu^\nu \\ &= a^\mu b_\mu, \end{aligned}$$

where the critical simplification $\langle \mathbf{e}^\nu, \mathbf{e}_\mu \rangle = \delta_\mu^\nu$ would not have occurred if \mathbf{e}^ν was not chosen to be the dual basis corresponding to \mathbf{e}_μ .

Having shown that every vector space has a corresponding dual space, it is natural to ask about the dual space of the dual space. At first glance, it seems as if our creation of a single vector space has lead to an infinite family of vector spaces simultaneously springing into existence. Fortunately (or unfortunately, depending on the reader's affinity for vector spaces), it turns out that the only vector spaces which are of any interest to us are the one which we explicitly constructed and its dual space. This is due to a particularly strong relationship between vector spaces and the dual spaces of their dual spaces.

Theorem 4. *Let V be a vector space over F . If $\mathbf{v}^{**} : V^* \rightarrow F$ is a linear mapping, then there exists a unique vector $\mathbf{v} \in V$ such that $\langle \mathbf{v}^{**}, \mathbf{v}^* \rangle = \langle \mathbf{v}^*, \mathbf{v} \rangle$ for all $\mathbf{v}^* \in V^*$. This defines an isomorphism between V and $(V^*)^*$.*

Proof. □

This correspondence is sometimes called the natural correspondence, and it is so strong that it is sensible to refer to elements of $(V^*)^*$ by their corresponding elements of V . In this sense, we can state that a vector $\mathbf{v} \in V$ is a linear mapping $\mathbf{v} : V^* \rightarrow F$, analogously to how a covector $\mathbf{v}^* \in V^*$ is a linear mapping $\mathbf{v}^* : V \rightarrow F$. One particularly nice way of thinking about this correspondence is to observe that just as $\langle \mathbf{v}^*, \circ \rangle$ can be thought of as taking a vector and returning a scalar, $\langle \circ, \mathbf{v} \rangle$ can be thought of as taking a covector and returning a scalar.

2 Tensors, Tensor Products, & Tensor Product Spaces

Having reviewed the theory of vector spaces and examined dual spaces, we have exhausted the set of all linear mappings from vectors to scalars and the set of all linear mappings from covectors to scalars. We now turn to the examination of multilinear mappings from multiple vectors and covectors to scalars, and indentify these mappings as tensors. We then carry on to define the tensor product as a method of combining tensors to create new tensors. This will allow us to define an infinite family of new vector spaces: the tensor product spaces.

Definition 9 (Multilinear Mappings). *Let U_1, \dots, U_k be vector spaces over F . Let $\mathbf{u}_i \in U_i$ for all $i \in \{1, \dots, k\}$. A mapping $f : U_1 \times \dots \times U_k \rightarrow F$ is called multilinear if the mapping $g : U_i \rightarrow F$ defined by $\langle g, \mathbf{u}_i \rangle = \langle f, (\mathbf{u}_1, \dots, \mathbf{u}_i, \dots, \mathbf{u}_k) \rangle$ is linear for all $i \in \{1, \dots, k\}$.*

Definition 10 (Tensors). *Let V be a vector space over F , and let V^* be its dual space. A multilinear mapping $T : V^{*p} \times V^q \rightarrow F$ is called a type (p, q) tensor. The set of all type (p, q) tensors is denoted \mathcal{T}_p^q . \mathcal{T}_p^q is a vector space over F once vector addition and scalar multiplication have been defined as follows:*

- $\langle T + U, (\mathbf{v}^1, \dots, \mathbf{v}^p, \mathbf{v}_1, \dots, \mathbf{v}_q) \rangle = \langle T, (\mathbf{v}^1, \dots, \mathbf{v}^p, \mathbf{v}_1, \dots, \mathbf{v}_q) \rangle + \langle U, (\mathbf{v}^1, \dots, \mathbf{v}^p, \mathbf{v}_1, \dots, \mathbf{v}_q) \rangle$
- $\langle aT, (\mathbf{v}^1, \dots, \mathbf{v}^p, \mathbf{v}_1, \dots, \mathbf{v}_q) \rangle = a \langle T, (\mathbf{v}^1, \dots, \mathbf{v}^p, \mathbf{v}_1, \dots, \mathbf{v}_q) \rangle$.

Verification that the relevant axioms hold is mostly tedious, and so is omitted. The interested reader may wish to prove that the relevant axioms hold as an exercise.

Definition 11 (Rank). *Let $T \in \mathcal{T}_p^q$. $p + q$ is called the rank of T .*

Careful consideration of this definition will reveal that we have already defined two types of tensors: vectors and covectors. According to our new definition, vectors are type $(1, 0)$ tensors and covectors are type $(0, 1)$ tensors. We have actually seen a third type of tensor, although not a very interesting one. As a matter of convention, scalars are considered to be type $(0, 0)$ tensors.

Having defined tensors we have, in some sense, accomplished everything which we set out to. Unfortunately, it rapidly becomes apparent that there are substantial issues with the manner in which we defined tensors. One significant issue is that actually constructing a tensor while relying only on this definition represents a somewhat onerous task. The multilinearity of tensors helps somewhat, but it still appears to be necessary to define the operation of a type (p, q) tensor on all $(pq)^{|V|}$ combinations of basis vectors. We solve this issue by presenting a manner for combining tensors to produce new tensors of higher rank: the tensor product. This will allow us to represent all tensors in terms of vectors and covectors.

Definition 12 (Tensor Product). Let $T \in \mathcal{T}_p^q$ and $U \in \mathcal{T}_r^s$. Let $\mathbf{v}_i, \mathbf{u}_i \in V$ and $\mathbf{v}^i, \mathbf{u}^i \in V^*$ for all i . The tensor product of T and U defines a new tensor denoted by $T \otimes U \in \mathcal{T}_{p+r}^{q+s}$. Let $S = T \otimes U$. The operation of S is defined by

$$\langle S, (\mathbf{v}^1, \dots, \mathbf{v}^p, \mathbf{v}_1, \dots, \mathbf{v}_q, \mathbf{u}^1, \dots, \mathbf{u}^r, \mathbf{u}_1, \dots, \mathbf{u}_s) \rangle = \langle T, (\mathbf{v}^1, \dots, \mathbf{v}^p, \mathbf{v}_1, \dots, \mathbf{v}_q) \rangle \langle U, (\mathbf{u}^1, \dots, \mathbf{u}^r, \mathbf{u}_1, \dots, \mathbf{u}_s) \rangle$$

Definition 13 (Tensor Product Space). Let V be a vector space. Let V^* be its dual space. A tensor product space is a set defined by

$$\{\mathbf{v}_1 \otimes \dots \otimes \mathbf{v}_q \otimes \mathbf{v}^1 \otimes \dots \otimes \mathbf{v}^p : \mathbf{v}_i \in V, \mathbf{v}^i \in V^*\}$$

which is typically denoted by $\underbrace{V \otimes \dots \otimes V}_q \otimes \underbrace{V^* \otimes \dots \otimes V^*}_p$ or $V^q \otimes V^{*p}$.

Theorem 5. $V^q \otimes V^{*p} = \mathcal{T}_p^q$.

Proof.

□

This reframing of \mathcal{T}_p^q as a tensor product space is deeply advantageous. In particular, the fact that it is constructed directly from V and V^* induces a natural basis on \mathcal{T}_p^q , namely $\mathbf{e}_{\mu_1} \otimes \dots \otimes \mathbf{e}_{\mu_q} \otimes \mathbf{e}^{\nu_1} \otimes \dots \otimes \mathbf{e}^{\nu_p}$, where \mathbf{e}_{μ_i} and \mathbf{e}^{ν_i} are all simply copies of a basis of V and the corresponding dual basis. Similarly to our definition of the dual basis, this choice of basis for \mathcal{T}_p^q allows for the operation of an element of \mathcal{T}_p^q on the appropriate number of vectors and covectors to be evaluated with ease. We will, however, defer our discussion of this until we have more fully developed the theory of tensors.