Lecture 12: Tensors

Tianpei Xie

Oct. 20th., 2022

Contents

1 Multilinear Algebra

1.1 Multilinear Functions and Tensor Product

- **Definition** Suppose $V_1, \ldots, V_k$, and W are *vector spaces*. A map $F : V_1 \times \ldots \times V_k \to W$ is said to be ***multilinear*** if it is ***linear*** as a function of *each variable* ***separately*** when the others are held ***fixed***: for each i,

$$F(v_1, \ldots, av_i + a'v_i', \ldots, v_k) = a\,F(v_1, \ldots, v_i, \ldots, v_k) + a'F(v_1, \ldots, v_i', \ldots, v_k).$$

  A multilinear function of ***one variable*** is just ***a linear function***, and a multilinear function of ***two variables*** is generally called ***bilinear***.

- **Remark** Let us write $L(V_1, \ldots, V_k; W)$ for ***the set of all multilinear maps from*** $V_1 \times \ldots \times V_k$ ***to*** W. It is a ***vector space*** under the usual operations of *pointwise addition* and *scalar multiplication*:

$$\begin{aligned}(F' + F)(v_1, \ldots, v_i, \ldots, v_k) &= F(v_1, \ldots, v_i, \ldots, v_k) + F'(v_1, \ldots, v_i, \ldots, v_k),\\ (aF)(v_1, \ldots, v_i, \ldots, v_k) &= a\,F(v_1, \ldots, v_i, \ldots, v_k).\end{aligned}$$

- **Example** (***Some Familiar Multilinear Functions***).

  1. The ***dot product*** in $\mathbb{R}^n$ is a ***scalar-valued bilinear function*** of two vectors, used to compute ***lengths*** of vectors and ***angles*** between them.
  2. The ***cross product***, $(\cdot \times \cdot)$ in $\mathbb{R}^3$ is a ***vector-valued bilinear function*** of two vectors, used to compute ***areas*** of parallelograms and to find a third vector ***orthogonal*** to two given ones.
  3. The ***determinant***, $\det(\cdot)$ is a ***real-valued multilinear function*** of n *vectors* in $\mathbb{R}^n$, used to detect ***linear independence*** and to compute the ***volume*** of the parallelepiped spanned by the vectors.
  4. The ***bracket in a Lie algebra*** $\mathfrak{g}$ is a $\mathfrak{g}$***-valued bilinear function*** of two elements of $\mathfrak{g}$.

- **Example** (***Tensor Products of Covectors***).
  Suppose V is a vector space, and $\omega, \eta \in V^*$. Define a function $\omega \otimes \eta : V \times V \to \mathbb{R}$ by

$$(\omega \otimes \eta)(v_1, v_2) = \omega(v_1)\eta(v_2),$$

  where the product on the right is just ordinary multiplication of real numbers. The linearity of ω and η guarantees that $\omega \otimes \eta$ is a *bilinear function* of v_1 and v_2, so it is an element of $L(V, V; \mathbb{R})$. For example, if (e^1, e^2) denotes the standard dual basis for $\mathbb{R}^2$, then $e^1 \otimes e^2 : \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}$ is the bilinear function

$$e^1 \otimes e^2((w, x), (y, z)) = wz.$$

- **Definition** let $V_1, \ldots, V_k, W_1, \ldots, W_l$ be real vector spaces, and suppose $F \in L(V_1, \ldots, V_k; \mathbb{R})$ and $G \in L(W_1, \ldots, W_l; \mathbb{R})$. Define a function $F \otimes G : V_1 \times \ldots \times V_k \times W_1 \times \ldots \times W_l \to \mathbb{R}$ by

$$(F \otimes G)(v_1, \ldots, v_k, w_1 \ldots, w_l) = F(v_1, \ldots, v_k)\, G(w_1 \ldots, w_l) \tag{1}$$

It follows from the multilinearity of F and G that $(F \otimes G)(v_1, \ldots, v_k, w_1 \ldots, w_l)$ depends *linearly* on each argument v_i or w_j *separately*, so $F \otimes G$ is an element of $L(V_1, \ldots, V_k, W_1, \ldots, W_l; \mathbb{R})$ called ***the tensor product of*** F ***and*** G.

- **Remark** We can write tensor products of three or more multilinear functions unambiguously without parentheses. If $F_1, \ldots, F_l$ are *multilinear functions* depending on $k_1, \ldots, k_l$ variables, respectively, their tensor product $F_1 \times \ldots \times F_l$ is a *multilinear function* of $k = k_1 + \ldots + k_l$ variables, whose action on k vectors is given by inserting the first k_1 vectors into F_1, the next k_2 vectors into F_2, and so forth, and multiplying the results together.

- **Remark** The definition of multilinear function as well as tensor product show a ***recursion***. It *breaks* the complicated multi-variate calculation into product of smaller bivariate or univariate calculation.

- **Remark** If $\omega^j \in V_j^*$ for $j = 1, \ldots, k$, then $\omega^1 \otimes \ldots \otimes \omega^k \in L(V_1, \ldots, V_k; \mathbb{R})$ is the ***multilinear function*** given by

$$(\omega^1 \otimes \ldots \otimes \omega^k)(v_1, \ldots, v_k) = \omega^1(v_1) \ldots \omega^k(v_k). \tag{2}$$

We can see that $\omega^1 \otimes \ldots \otimes \omega^k$ is a multilinear extension of the linear functional ω.

- **Proposition 1.1** *(**A Basis for the Space of Multilinear Functions**).* *Let $V_1, \ldots, V_k$ be real vector spaces of dimensions $n_1, \ldots, n_k$, respectively. For each $j \in set1, \ldots, k$, let $(E_1^{(j)}, \ldots, E_{n_j}^{(j)})$ be a* **basis** *for V_j, and let $(\epsilon_{(j)}^1, \ldots, \epsilon_{(j)}^{n_j})$ be the corresponding **dual basis** for V_j^*. Then the set*

$$\mathfrak{B} = \left\{ \epsilon_{(1)}^{i_1} \otimes \ldots \otimes \epsilon_{(k)}^{i_k} :\ 1 \le i_j \le n_j, j = 1, \ldots, k \right\}$$

is a basis *for $L(V_1, \ldots, V_k; \mathbb{R})$, which therefore has **dimension equal to** $n_1 \ldots n_k$.*

1.2 Abstract Tensor Products of Vector Spaces

- We extend our result to abstract tensor product on multiple vector spaces. We need to first define the linear combinations.

- **Definition** For any set S, ***a formal linear combination of elements of*** S is *a function* $f : S \to \mathbb{R}$ such that $f(s) = 0$ for all *but finitely many* $s \in S$. The ***free (real) vector space*** on S, denoted by $\mathcal{F}(S)$, is *the set of all formal linear combinations of elements of* S. Under pointwise addition and scalar multiplication, $\mathcal{F}(S)$ becomes ***a vector space over*** $\mathbb{R}$.

- **Remark** For each element $x \in S$, there is a function $\delta_x \in \mathcal{F}(S)$ that takes the value 1 on x and zero on all other elements of S; typically we identify this function with x itself, and thus think of S as a subset of $\mathcal{F}(S)$.

  Every element $f \in \mathcal{F}(S)$ can then be written uniquely in the form $f = \sum_i^m a_i\, x_i$, where $x_1, \ldots, x_m$ are the elements of S for which $f(x_i) \neq 0$, and $a_i = f(x_i)$. Thus, S ***is a basis*** for $\mathcal{F}(S)$, which is therefore ***finite-dimensional*** if and only if S is ***a finite set***.

- **Remark** Normally the linear combinations are introduced in *vector space* V in which *the scalar multiplication* and *addition* are defined. Here we **generalize** it to **any set** S through a special function f that only take nonzero values in finite number of elements in S. A typical example of such function is the indicator function $\delta_x(s) = \mathbb{1}\{s = x\}$. Since the function taking values in $\mathbb{R}$ which equipped with a proper definition of addition and scalar multiplication, we can represent any function f in terms of a linear combination of functions δ_{x_i} instead of x_i itself. This way helps us to circumvent the need to define algebraic structure on S.

- **Proposition 1.2** ***(Characteristic Property of the Free Vector Space).***
  For any set S and any vector space W, every map $A : S \to W$ has a ***unique extension*** *to a* ***linear map*** *$\bar{A} : \mathcal{F}(S) \to W$.*

- **Definition** Now let $V_1, \ldots, V_k$ be real vector spaces. We begin by forming ***the free vector space*** $\mathcal{F}(V_1 \times \ldots \times V_k)$, which is the set of all finite formal linear combinations of k-tuples $(v_1, \ldots, v_k)$ with $v_i \in V_i$ for $i = 1, \ldots, k$. Let $\mathcal{R}$ be the ***subspace*** of $\mathcal{F}(V_1 \times \ldots \times V_k)$ *spanned* by all elements of the following forms:

$$\begin{aligned} &(v_1, \ldots, a\, v_i, \ldots, v_k) - a\,(v_1, \ldots, v_i, \ldots, v_k) \\ &(v_1, \ldots, v_i + v_i', \ldots, v_k) - (v_1, \ldots, v_i, \ldots, v_k) - (v_1, \ldots, v_i', \ldots, v_k) \end{aligned} \tag{3}$$

  with $v_j, v_j' \in V_j$, $i \in \{1, \ldots, k\}$, and $a \in \mathbb{R}$.

  Define ***the tensor product of the spaces*** $V_1, \ldots, V_k$, denoted by $V_1 \otimes \ldots \otimes V_k$, to be the following ***quotient vector space***:

$$V_1 \otimes \ldots \otimes V_k = \mathcal{F}(V_1 \times \ldots \times V_k)/\mathcal{R}$$

  and let $\Pi : \mathcal{F}(V_1 \times \ldots \times V_k) \to V_1 \otimes \ldots \otimes V_k$ be ***the natural projection***. The ***equivalence class*** of an element $(v_1, \ldots, v_k)$ in $V_1 \otimes \ldots \otimes V_k$ is denoted by

$$v_1 \otimes \ldots \otimes v_k = \Pi(v_1, \ldots, v_k) \tag{4}$$

  and is called ***the (abstract) tensor product of*** $(v_1, \ldots, v_k)$.

  It follows from the definition that abstract tensor products satisfy

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

- **Remark** Note that the definition implies that every element of $V_1 \otimes \ldots \otimes V_k$ can be expressed as a linear combination of elements of the form $(v_1 \otimes \ldots \otimes v_k)$ for $v_i \in V_i$; but it is not true in general that every element of the tensor product space is of the form $(v_1 \otimes \ldots \otimes v_k)$.

- **Remark** Intuitively, we want to define the tensor product $v_1 \otimes \ldots \otimes v_k$ by concatenating all vectors into k-tuple $(v_1, \ldots, v_k)$. But this naive construction is not enough. We have the following challenges:

  1. The product space $V_1 \times \ldots \times V_k$ is *not necessarily a* ***vector space*** since we have not define the addition and scalar product for k-tuple $(v_1, \ldots, v_k)$
  2. We want the ***multilinearity holds*** for the operator on k-tuple $(v_1, \ldots, v_k)$, i.e. we want

$$(v_1, \ldots, a\, v_i' + b\, v_i'', v_k) = a\,(v_1, \ldots, v_i', v_k) + b\,(v_1, \ldots, v_i'', v_k) \tag{5}$$

     for any $i \in \{1, \ldots, k\}$ and any $a, b \in \mathbb{R}$.

The above constructions aim to solve these challenges:

1. Instead of defining the algebraic structure on product space $V_1 \times \ldots \times V_k$, we extend it to ***the free vector space*** $\mathcal{F}(V_1 \times \ldots \times V_k)$, the set of *all linear combinations* of k-tuples $(v_1, \ldots, v_k)$. By construction $\mathcal{F}(V_1 \times \ldots \times V_k) \supseteq V_1 \times \ldots \times V_k$ and *it is a vector space without defining the algebraic structure* since it use an indicator function to map to $\mathbb{R}$.

2. Instead of enforcing the *multilinearity* to hold, we ***partition the space*** $\mathcal{F}(V_1 \times \ldots \times V_k)$ ***according to the multilinearitiy rule***. That is, the set of tuples $(v_1, \ldots, a\, v_i' + b\, v_i'', v_k)$ and $(v_1, \ldots, v_i', v_k), (v_1, \ldots, v_i'', v_k)$ that satisfies the equation (5) *will be grouped together* via the equivalence relationship. The rule is actually a subspace of linear combinations of (difference of) tuples, denoted as $\mathcal{R} \subseteq \mathcal{F}(V_1 \times \ldots \times V_k)$.

   Now we instead focusing on the equivalent class itself. By construction, ***the equivalence class will satisfies the multilinear rule*** (5) (The representer of the equivalence class follow the rule). Thsu $V_1 \otimes \ldots \otimes V_k = \mathcal{F}(V_1 \times \ldots \times V_k)/\mathcal{R}$ is the tensor product space that we wants.

- **Remark** To understand a tensor product of vectors $v_1 \otimes \ldots \otimes v_k$, we need to know that it can be seen as *an equivalent class* of k-tuple $(v_1, \ldots, v_k)$. For tuples $(w_1, \ldots, w_k) \in (v_1, \ldots, v_k) + \mathcal{R}$, we have $(w_1, \ldots, w_k) - (v_1, \ldots, v_k) \in \mathcal{R}$.

$$\begin{aligned}
&\Leftrightarrow (w_1, \ldots, w_k) - (v_1, \ldots, v_k) \in \operatorname{span}\{(v_1, \ldots, v_k) \text{ follows rule (3)}\} \\
&\Leftrightarrow \text{for some } \; j \in \{1, \ldots, k\}\,, \text{ so that } w_j = a^j\, w_j', \\
&\quad \text{then } (v_1, \ldots, v_j, \ldots, v_k) = a^j (v_1, \ldots, w_j', \ldots v_k)
\end{aligned}$$

- **Proposition 1.3** ***(Characteristic Property of the Tensor Product Space).***
  Let $V_1, \ldots, V_k$ be finite-dimensional real vector spaces. If $A : V_1 \times \ldots \times V_k \to X$ is ***any multilinear map*** *into a vector space X, then there is a* ***unique linear map*** $\widetilde{A} : V_1 \otimes \ldots \otimes V_k \to X$ *such that the following diagram commutes:*

$$\begin{tikzcd}
V_1 \times \ldots \times V_k \arrow[r, "A"] \arrow[d, "\pi"'] & X \\
V_1 \otimes \ldots \otimes V_k \arrow[ur, dashed, "\widetilde{A}"'] &
\end{tikzcd} , \tag{6}$$

  where π is the map $\pi(v_1, \ldots, v_k) = v_1 \otimes \ldots \otimes v_k$.

- **Remark** The *characteristic property of the tensor product space* states that *any mulilinear function* $\tau : V_1 \times \ldots \times V_k \to \mathbb{R}$ descends into a *linear map* $\widetilde{\tau} : V_1 \otimes \ldots \otimes V_k \to \mathbb{R}$ so that any linear combinations of tensor products $v_{i_1} \otimes \ldots \otimes v_{i_k}$ is expressed as

$$\widetilde{\tau}\left(a^{i_1 \ldots i_k}\, v_{i_1} \otimes \ldots \otimes v_{i_k}\right) = a^{i_1 \ldots i_k}\, \tau(v_{i_1}, \ldots, v_{i_k})$$

- **Proposition 1.4** ***(A Basis for the Tensor Product Space).***
  Suppose $V_1, \ldots, V_k$ are real vector spaces of dimensions $n_1 \ldots n_k$, respectively. For each $j = 1, \ldots, k$, suppose $(E_1^{(j)}, \ldots, E_{n_j}^{(j)})$ is a ***basis*** *for V_j. Then the set*

$$\mathfrak{C} = \left\{ E_{i_1}^{(1)} \otimes \ldots \otimes E_{i_k}^{(k)} \,:\; 1 \le i_j \le n_j, j = 1, \ldots, k \right\}$$

  is a basis *for $V_1 \otimes \ldots \otimes V_k$, which therefore has* ***dimension equal to*** *$n_1 \ldots n_k$.*

- **Proposition 1.5** ***(Associativity of Tensor Product Spaces).***
  Let V_1, V_2, V_3 be finite-dimensional real vector spaces. There are ***unique isomorphisms***

$$V_1 \otimes (V_2 \otimes V_3) \simeq V_1 \otimes V_2 \otimes V_3 \simeq (V_1 \otimes V_2) \otimes V_3$$

  under which elements of the forms $v_1 \otimes (v_2 \otimes v_3)$, $v_1 \otimes v_2 \otimes v_3$ and $(v_1 \otimes v_2) \otimes v_3$ all correspond.

- The connection between tensor products in this abstract setting and the more concrete tensor products of ***multilinear functionals*** that we defined earlier is based on the following proposition.

  Proposition 1.6 ***(Abstract vs. Concrete Tensor Products)****.[Lee, 2003.]*
  If $V_1, \ldots, V_k$ are finite-dimensional vector spaces, there is a ***canonical isomorphism***

$$V_1^* \otimes \ldots \otimes V_k^* \simeq L(V_1, \ldots, V_k; \mathbb{R}) \tag{7}$$

  under which the ***abstract tensor product*** *defined by* (4) *corresponds to the* ***tensor product of covectors*** *defined by* (2)*.*

  Proof: First, define a map $\Phi : V_1^* \times \ldots \times V_k^* \to L(V_1, \ldots, V_k; \mathbb{R})$ by

$$\Phi(\omega^1, \ldots, \omega^k)(v_1, \ldots, v_k) = \omega^1(v_1) \ldots \omega^k(v_k).$$

  The expression on the right depends linearly on each v_i , so $\Phi(\omega^1, \ldots, \omega^k)$ is indeed an element of the space $L(V_1, \ldots, V_k; \mathbb{R})$. It is easy to check that Φ is multilinear as a function of $(\omega^1, \ldots, \omega^k)$, so by the characteristic property it descends uniquely to a linear map $\widetilde{\Phi}$ from $V_1^* \otimes \ldots \otimes V_k^*$ to $L(V_1, \ldots, V_k; \mathbb{R})$, which satisfies

$$\widetilde{\Phi}(\omega^1 \otimes \ldots \otimes \omega^k)(v_1, \ldots, v_k) = \omega^1(v_1) \ldots \omega^k(v_k).$$

  It follows immediately from the definition that $\widetilde{\Phi}$ takes abstract tensor products to tensor products of covectors. It also takes the basis of $V_1^* \otimes \ldots \otimes V_k^*$ to the basis for $L(V_1, \ldots, V_k; \mathbb{R})$, so it is an isomorphism. (Although we used bases to prove that $\widetilde{\Phi}$ is an isomorphism, $\widetilde{\Phi}$ itself is canonically defined without reference to any basis.) ∎

- **Remark** Using this canonical isomorphism, we henceforth use the notation $V_1^* \otimes \ldots \otimes V_k^*$ to denote either the abstract tensor product space or the space $L(V_1, \ldots, V_k; \mathbb{R})$, focusing on whichever interpretation is more convenient for the problem at hand.

- **Remark** Through this identification, an element $\omega^1 \otimes \ldots \otimes \omega^k \in V_1^* \otimes \ldots \otimes V_k^*$ is considered as a ***multi-linear functional***

$$(\omega^1 \otimes \ldots \otimes \omega^k)(v_1, \ldots, v_k) = \omega^1(v_1) \ldots \omega^k(v_k)$$

  which can also descend into a linear map on tensor product of v_i

$$\widetilde{\omega}^{1,2,\ldots,k}(v_1 \otimes \ldots \otimes v_k) = \omega^1(v_1) \ldots \omega^k(v_k)$$

- **Remark** Since we are assuming the vector spaces are all finite-dimensional, we can also identify each V_j with its second dual space V_j^{**}, and thereby obtain ***another canonical identification***

$$V_1 \otimes \ldots \otimes V_k \simeq L(V_1^*, \ldots, V_k^*; \mathbb{R}) \tag{8}$$

- **Remark** As we see, the space of tensor product defines a set of ***parallel linear systems***. All ***sub-systems*** are ***independent***. Each sub-system has its own ***basis***, its own ***linear operations*** and its own ***representation***. The tensor product operation *group* these independent linear systems together and *consider them as a whole*.

  For ***the whole system perspective***, its representations are collected locally and then concatenated together. The linear map on ***the concatenated representation*** is essentially the same as ***applying linear map*** in each sub-system and ***multiplying*** them together. This is the same as computing the joint distribution by product of marginal distributions. ***The multiplication principle*** is applied when counting the size of the whole system.

  The space of tensor product $V_1 \otimes \ldots \otimes V_k$ reflect a ***divide-and-conquer strategy*** and a ***local-global strategy*** to study the complex functions such as *multilinear functionals* $\alpha(v_1, \ldots, v_k)$. In the former, we study it by ***perturbing*** *the input of each sub-system*. In the latter, we think of it as ***a linear map*** on the k-tensors $v_1 \otimes \ldots \otimes v_k$.

1.3 Covariant and Contravariant Tensors on a Vector Space

- **Definition** Let V be a finite-dimensional real vector space. If k is a positive integer, ***a covariant k-tensor*** on V is an element of *the k-fold tensor product* $V^* \otimes \ldots \otimes V^*$, which we typically think of as ***a real-valued multilinear function of k elements of V***:

$$\alpha : \underbrace{V \times \ldots \times V}_{k} \to \mathbb{R}$$

  The number k is called ***the rank of*** α. A 0-tensor is, by convention, just a real number (a real-valued function depending multilinearly on *no vectors*!).

  We denote ***the vector space*** of ***all covariant k-tensors on V*** by the shorthand notation

$$T^k V^* = \underbrace{V^* \otimes \ldots \otimes V^*}_{k}$$

- **Example** (***Covariant Tensors***).
  Let V be a finite-dimensional vector space.

  1. *Every linear functional* $\omega : V \to \mathbb{R}$ is multilinear, so ***a covariant 1-tensor is just a covector***. Thus, $T^1(V^*)$ is equal to V^*.
  2. *A covariant 2-tensor* on V is *a real-valued* ***bilinear function*** of two vectors, also called ***a bilinear form***. One example is *the dot product* on $\mathbb{R}^n$. More generally, ***every inner product is a covariant 2-tensor***.
  3. *The determinant*, thought of as a function of n vectors, is ***a covariant n-tensor on*** $\mathbb{R}^n$.

- **Definition** For any finite-dimensional real vector space V, we define *the space of* ***contravariant tensors*** *on* V of ***rank*** k to be the vector space

$$T^k V = \underbrace{V \otimes \ldots \otimes V}_{k}$$

In particular, $T^1(V) = V$, and by convention $T^0(V) = \mathbb{R}$. Because we are assuming that V is finite-dimensional, it is possible to identify this space with *the set of* ***multilinear functionals of*** k ***covectors***:

$$T^k V = \left\{ \text{multilinear functionals } \alpha : \underbrace{V^* \times \ldots \times V^*}_{k} \to \mathbb{R} \right\}$$

- **Definition** Even more generally, for any nonnegative integers k, l, we define *the space of* ***mixed tensors on*** V ***of type*** (k, l) as

$$T^{(k,l)} V = \underbrace{V \otimes \ldots \otimes V}_{k} \otimes \underbrace{V^* \otimes \ldots \otimes V^*}_{l}$$

- **Remark** Some of these spaces are identical:

$$\begin{aligned} T^{(0,0)} V &= T^0 V = T^0 V^* = \mathbb{R} \\ T^{(0,1)} V &= T^1 V^* = V^* \\ T^{(1,0)} V &= T^1 V = V \\ T^{(k,0)} V &= T^k V \\ T^{(0,k)} V &= T^k V^* \end{aligned}$$

- **Corollary 1.7** *Let V be an n-dimensional real vector space. Suppose (E_i) is any basis for V and (ϵ^j) is the dual basis for V^*. Then the following sets constitute bases for the tensor spaces over V:*

$$\begin{aligned} &\left\{ \epsilon^{i_1} \otimes \ldots \otimes \epsilon^{i_k} : \, 1 \le i_s \le n, s = 1, \ldots, k \right\} \text{ is basis for } T^k V^*; \\ &\left\{ E_{i_1} \otimes \ldots \otimes E_{i_k} : \, 1 \le i_s \le n, s = 1, \ldots, k \right\} \text{ is basis for } T^k V; \\ &\left\{ E_{i_1} \otimes \ldots \otimes E_{i_k} \otimes \epsilon^{j_1} \otimes \ldots \otimes \epsilon^{j_l} : \, 1 \le i_1, \ldots i_k, j_1, \ldots, j_l \le n \right\} \text{ is basis for } T^{(k,l)} V; \end{aligned} \tag{9}$$

Therefore, dim $T^k V^* =$ dim $T^k V = n^k$ and dim $T^{(k,l)} V = n^{k+l}$

- **Remark** (***Coordinate Representation of Covariant*** k***-Tensor***)
In particular, once a basis is chosen for V, every ***covariant*** k***-tensor*** $\alpha \in T^k(V^*)$ can be written uniquely in the form

$$\alpha = \alpha_{i_1, i_2, \ldots, i_k} \epsilon^{i_1} \otimes \ldots \otimes \epsilon^{i_k} \tag{10}$$

where the n^k coefficients $\alpha_{i_1, i_2, \ldots, i_k}$ are determined by

$$\alpha_{i_1, i_2, \ldots, i_k} = \alpha \left(E_{i_1}, \ldots, E_{i_k} \right) \tag{11}$$

For instance, covariant 2-tensor is bilinear form. Every *bilinear form* can be written as $\beta = \beta_{i,j} \epsilon^1 \otimes \epsilon^2$, for some uniquely determined $n \times n$ matrix $(\beta_{i,j})$.

- **Exercise 1.8** *Let* $v_1 = \sin(y)\frac{\partial}{\partial x}|_{(1,\pi/2)} - \frac{1}{2}x^2\frac{\partial}{\partial y}|_{(1,\pi/2)}$ *and* $v_2 = \cos(y)\frac{\partial}{\partial x}|_{(1,\pi/2)} + (x + y)\frac{\partial}{\partial y}|_{(1,\pi/2)}$. $\omega_1 = 2xdx|_{(1,\pi/2)} + \cos(y)dy|_{(1,\pi/2)}$, $\omega_2 = 2\cos(y)dx|_{(1,\pi/2)} - (x^2+y^2)dy|_{(1,\pi/2)}$.

$$\begin{aligned}
\omega_1 \otimes \omega_2 &= \left(2xdx_{(1,\pi/2)} + \cos(y)dy_{(1,\pi/2)}\right) \otimes \left(2\cos(y)dx_{(1,\pi/2)} - (x^2+y^2)dy|_{(1,\pi/2)}\right) \\
&= (2dx) \otimes \left(-(1+(\pi/2)^2)dy\right) = -(2+\pi^2/2)dx|_{(1,\pi/2)} \otimes dy|_{(1,\pi/2)} \\
\omega_1 \otimes \omega_2(v_1, v_2) &= \omega_1(v_1)\omega_2(v_2) \\
&= -(2+\pi^2/2)dx|_{(1,\pi/2)} \left(\sin(y)\frac{\partial}{\partial x}|_{(1,\pi/2)} - \frac{1}{2}x^2\frac{\partial}{\partial y}|_{(1,\pi/2)}\right) \\
&\quad dy|_{(1,\pi/2)} \left(\cos(y)\frac{\partial}{\partial x}|_{(1,\pi/2)} + (x+y)\frac{\partial}{\partial y}|_{(1,\pi/2)}\right) \\
&= -(2+\pi^2/2)dx|_{(1,\pi/2)} \left(\frac{\partial}{\partial x}|_{(1,\pi/2)} - \frac{1}{2}\frac{\partial}{\partial y}|_{(1,\pi/2)}\right) dy|_{(1,\pi/2)} \left((1+\pi/2)\frac{\partial}{\partial y}|_{(1,\pi/2)}\right) \\
&= -(2+\pi^2/2)(1+\pi/2) \quad \blacksquare
\end{aligned}$$

2 Symmetric and Alternating Tensors

2.1 Symmetric Tensors

- **Definition** Let V be a finite-dimensional vector space. A ***covariant k-tensor*** α on V is said to be ***symmetric*** if its value is ***unchanged*** by ***interchanging*** any pair of arguments:

$$\alpha(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k) = \alpha(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k)$$

whenever $i \leq i < j \leq k$.

- **Remark** The following statements are equivalent for a covariant k-tensor α:
  1. α is symmetric;
  2. For any vectors $v_1, \ldots, v_k \in V$, the value of $\alpha(v_1, \ldots, v_k)$ is unchanged when $v_1, \ldots, v_k$ are rearranged in any order.
  3. The components $\alpha_{i_1,\ldots,i_k}$ of α with respect to any basis are unchanged by any permutation of the indices.

- **Definition** The set of ***symmetric covariant k-tensors*** is a linear subspace of the space $T^k(V^*)$ of all covariant k-tensors on V; we denote this subspace by $\Sigma^k(V^*)$

- **Definition** There is a ***natural projection*** from $T^k(V^*)$ to $\Sigma^k(V^*)$ defined as follows. First, let S_k denote ***the symmetric group on k elements***, that is, the group of ***permutations*** of the set $\{1, \ldots, k\}$. Given a k-tensor α and a permutation $\sigma \in S_k$, we define a new k-tensor ${}^\sigma\alpha$ by

$${}^\sigma\alpha(v_1, \ldots, v_k) = \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)})$$

Note that ${}^\tau({}^\sigma\alpha) ={}^{\tau\sigma} \alpha$ where $\tau\sigma$ represents the composition of τ and σ, that is, $\tau\sigma(i) = \tau(\sigma(i))$.(*This is the reason for putting σ before α in the notation ${}^\sigma\alpha$ instead of after it.*)

We define a ***projection*** Sym : $T^k(V^*) \to \Sigma^k(V^*)$ called ***symmetrization*** by

$$\text{Sym } \alpha = \frac{1}{k!} \sum_{\sigma \in S_k} {}^{\sigma}\alpha$$

More explicitly, this means that

$$\text{Sym } \alpha(v_1, \ldots, v_k) = \frac{1}{k!} \sum_{\sigma \in S_k} \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)})$$

- **Proposition 2.1** *(**Properties of Symmetrization**).*
  Let α be a covariant tensor on a finite-dimensional vector space.
  1. *Sym α is symmetric.*
  2. *Sym $\alpha = \alpha$ if and only if α is symmetric.*
- **Remark** If α and β are symmetric tensors on V, then $\alpha \otimes \beta$ is not symmetric in general. However, using the symmetrization operator, it is possible to define a new product that takes a pair of symmetric tensors and yields another symmetric tensor.
- **Definition** If $\alpha \in \Sigma^k(V^*)$ and $\beta \in \Sigma^k(V^*)$, we define their ***symmetric product*** to be the $(k+l)$-tensor $\alpha\,\beta$(denoted by juxtaposition with no intervening product symbol) given by

  $$\alpha\,\beta = \text{Sym } (\alpha \otimes \beta)$$

  More explicitly, the action of $\alpha\,\beta$ on vectors $v_1, \ldots, v_{k+l}$ is given by

  $$\alpha\,\beta(v_1, \ldots, v_{k+l}) = \frac{1}{(k+l)!} \sum_{\sigma \in S_{k+l}} \alpha(v_{\sigma(1)}, \ldots, v_{\sigma(k)})\beta(v_{\sigma(k+1)}, \ldots, v_{\sigma(k+l)})$$

- **Proposition 2.2** *(Properties of the Symmetric Product).*
  1. *The symmetric product is **symmetric** and **bilinear**: for all symmetric tensors α, β, γ and all $a, b \in \mathbb{R}$,*

     $$\begin{aligned} \alpha\,\beta &= \beta\,\alpha \\ (a\,\alpha + b\,\beta)\,\gamma &= a\,\alpha\,\gamma + b\,\beta\,\gamma = \gamma\,(a\,\alpha + b\,\beta) \end{aligned}$$

  2. *If α and β are **covectors**, then*

     $$\alpha\,\beta = \frac{1}{2}\left(\alpha \otimes \beta + \beta \otimes \alpha\right).$$

2.2 Alternating Tensors

- **Definition** Assume that V is a finite-dimensional real vector space. ***A covariant k-tensor*** α on V is said to be ***alternating*** (or ***antisymmetric*** or ***skew-symmetric***) if it ***changes***

sign whenever two of its arguments are *interchanged.* This means that for all vectors $v_1, \ldots, v_k \in V$ and every pair of distinct indices i, j it satisfies

$$\alpha\left(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_k\right) = -\alpha\left(v_1, \ldots, v_j, \ldots, v_i, \ldots, v_k\right)$$

Alternating covariant k-tensors are also variously called ***exterior forms***, ***multicovectors***, or ***k-covectors***.

The subspace of ***all alternating covariant*** k***-tensors*** on V is denoted by $\Lambda^k(V^*) \subseteq T^k(V^*)$.

- **Definition** Recall that for any permutation $\sigma \in S_k$, ***the sign of*** σ, denoted by sgn σ, is equal to $+1$ if σ *is* ***even*** (i.e., can be written as *a composition of an* ***even*** *number of transpositions*), and -1 if σ is ***odd***.

- **Remark** The following statements are equivalent for a covariant k-tensor α:

  1. α is alternating;

  2. For any vectors $v_1, \ldots, v_k \in V$, and any permutation $\sigma \in S_k$

  $$\alpha\left(v_{\sigma(1)}, \ldots, v_{\sigma(k)}\right) = (\text{sgn } \sigma)\alpha\left(v_1, \ldots, v_k\right)$$

  3. With respect to any basis, the components $\alpha_{i_1, \ldots, i_k}$ of α change sign whenever two indices are interchanged.

- **Remark** Regarding the symmetric and alternating tensors:

  1. Every 0-tensor (which is just a real number) is both *symmetric* and *alternating*.

  2. Every 1-tensor is both *symmetric* and *alternating*.

  3. An ***alternating*** 2-tensor on V is ***a skew-symmetric bilinear form***.

  4. Every ***covariant*** **2*-tensor*** β can be expressed as a ***sum*** *of an* ***alternating tensor*** *and a* ***symmetric one***, because

  $$\beta(v, w) = \frac{1}{2}\left(\beta(v, w) - \beta(w, v)\right) + \frac{1}{2}\left(\beta(v, w) + \beta(w, v)\right) \equiv \alpha(v, w) + \sigma(v, w)$$

  where $\alpha(v, w) = \frac{1}{2}\left(\beta(v, w) - \beta(w, v)\right)$ is an alternating tensor, and $\sigma(v, w) = \frac{1}{2}(\beta(v, w) + \beta(w, v))$ is symmetric.

  5. The above is not true for general higher tensor.

3 Tensors and Tensor Fields on Manifolds

3.1 Definitions

- **Definition** Now let M be a smooth manifold with or without boundary. We define the ***bundle of covariant*** k***-tensors*** on M by

$$T^k T^* M = \coprod_{p \in M} T^k\left(T_p^* M\right)$$

Analogously, we define ***the bundle of contravariant k-tensors*** by

$$T^k TM = \bigsqcup_{p \in M} T^k \left(T_p M\right)$$

and ***the bundle of mixed tensors of type*** (k, l) by

$$T^{(k,l)} TM = \bigsqcup_{p \in M} T^{(k,l)} \left(T_p M\right)$$

- **Remark** There are natural identifications

$$\begin{aligned}
&T^{(0,0)}TM = T^0 T^* M = T^0 TM = M \times \mathbb{R};\\
&T^{(0,1)}TM = T^1 T^* M = T^* M;\\
&T^{(1,0)}TM = T^1 TM = TM;\\
&T^{(0,k)}TM = T^k T^* M;\\
&T^{(k,0)}TM = T^k TM.
\end{aligned}$$

  Any one of these bundles is called ***a tensor bundle over*** M. (Thus, the tangent and cotangent bundles are special cases of tensor bundles.)

- **Definition** *A **section** of a tensor bundle* is called *a **(covariant, contravariant, or mixed) tensor field** on M*. ***A smooth tensor field*** is a section that is smooth in the usual sense of smooth sections of vector bundles.

  So ***contravariant*** 1***-tensor fields*** are the same as ***vector fields***, and ***covariant*** 1***-tensor fields*** are ***covector fields***.

- **Remark** The ***spaces of smooth sections*** of these tensor bundles, $\Gamma\left(T^k T^* M\right)$, $\Gamma\left(T^k TM\right)$, and $\Gamma(T^{(k,l)}TM)$, are ***infinite-dimensional vector spaces** over* $\mathbb{R}$, and ***modules*** over $\mathcal{C}^\infty(M)$.

  We also denote the *space of smooth covariant tensor fields* as

$$\mathcal{T}^k(M) = \Gamma\left(T^k T^* M\right).$$

- **Remark** (***Coordinate Representation of Tensor Fields***)
  In any smooth local coordinates (x^i), sections of these bundles can be written (using the summation convention) as

$$A = \begin{cases} A_{i_1,\dots,i_k}\, dx^{i_1} \otimes \dots \otimes dx^{i_k}, & A \in \Gamma\left(T^k T^* M\right);\\[2ex] A^{i_1,\dots,i_k}\, \dfrac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \dfrac{\partial}{\partial x^{i_k}}, & A \in \Gamma\left(T^k TM\right);\\[2ex] A^{i_1,\dots,i_k}_{j_1,\dots,j_l}\, \dfrac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \dfrac{\partial}{\partial x^{i_k}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_k} & A \in \Gamma\left(T^{(k,l)}TM\right); \end{cases} \tag{12}$$

  The functions $A_{i_1,\dots,i_k}$, $A^{i_1,\dots,i_k}$, or $A^{i_1,\dots,i_k}_{j_1,\dots,j_l}$ are called the ***component functions*** of A in the chosen coordinates.

- **Proposition 3.1** ***(Smoothness Criteria for Tensor Fields).***
  Let M be a smooth manifold with or without boundary, and let $A : M \to T^k T^* M$ be a rough section. The following are equivalent.

  1. *A is smooth.*
  2. *In **every** smooth coordinate chart, the **component functions** of A are smooth.*
  3. *Each point of M is contained in **some** coordinate chart in which A has **smooth component functions**.*
  4. *If $X_1, \ldots, X_k \in \mathfrak{X}(M)$, then the function $A(X_1, \ldots, X_k) : M \to \mathbb{R}$, defined by*
  $$A(X_1, \ldots, X_k)(p) = A_p \left(X_1\big|_p, \ldots, X_k\big|_p \right) \tag{13}$$
  is smooth
  5. *Whenever $X_1, \ldots, X_k$ are smooth vector fields defined on **some open subset** $U \subseteq M$, the function $A(X_1, \ldots, X_k)$ is smooth on U.*

- **Proposition 3.2** *Suppose M is a smooth manifold with or without boundary, $A \in \mathcal{T}^k(M)$, $B \in \mathcal{T}^l(M)$, and $f \in \mathcal{C}^\infty(M)$. Then fA and $A \otimes B$ are also **smooth tensor fields**, whose **components** in any smooth local coordinate chart are*
  $$\begin{aligned} (fA)_{i_1,\ldots,i_k} &= f A_{i_1,\ldots,i_k}, \\ (A \otimes B)_{i_1,\ldots,i_{k+l}} &= A_{i_1,\ldots,i_k} B_{i_{k+1},\ldots,i_{k+l}}. \end{aligned}$$

- **Lemma 3.3** ***(Tensor Characterization Lemma)****.[Lee, 2003.]*
  A map
  $$\mathcal{A} : \underbrace{\mathfrak{X}(M) \times \ldots \times \mathfrak{X}(M)}_{k} \to \mathcal{C}^\infty(M). \tag{14}$$
  *is **induced** by a **smooth covariant k-tensor field** A as in (13) if and only if it is **multilinear** over $\mathcal{C}^\infty(M)$.*

  Remark Note that for any $f, f' \in \mathcal{C}^\infty(M)$, any $X_i, X_i' \in \mathfrak{X}(M)$
  $$A(X_1, \ldots, fX_i + f'X_i', \ldots, X_k) = fA(X_1, \ldots, X_i, \ldots, X_k) + f'A(X_1, \ldots, X_i', \ldots, X_k)$$

- **Definition** For symmetric and alternating tensor field, we have the following definition:

  1. ***A symmetric tensor field*** on a manifold (with or without boundary) is simply *a covariant tensor field* whose value at each point is a *symmetric tensor*.

     The ***symmetric product*** of two or more tensor fields is defined pointwise, just like the tensor product. Thus, for example, if A and B are *smooth covector fields*, their symmetric product is ***the smooth 2-tensor field*** AB, which is given by
     $$AB = \frac{1}{2}(A \otimes B) + \frac{1}{2}(B \otimes A).$$

  2. ***Alternating tensor fields*** are called ***differential forms***;

3.2 Pullbacks of Tensor Fields

- **Definition** Suppose $F: M \to N$ is a smooth map. For any point $p \in M$ and any *k-tensor* $\alpha \in T^k(T^*_{F(p)}N)$, we define a tensor $dF^*_p(\alpha) \in T^k\left(T^*_pM\right)$, called ***the pointwise pullback*** of α by F at p, by
$$dF^*_p(\alpha)(v_1, \ldots, v_k) = \alpha\left(dF_p(v_1), \ldots, dF_p(v_k)\right)$$
for any $v_1, \ldots, v_k \in T_pM$.

- **Definition** If A is *a covariant k-tensor field on* N, we define a rough***k-tensor field*** F^*A ***on*** M; called ***the pullback of*** A ***by*** F, by
$$(F^*A)_p = dF^*_p(A_{F(p)}).$$
This tensor field acts on vectors $v_1, \ldots, v_k \in T_pM$ by
$$(F^*A)_p(v_1, \ldots, v_k) = A_{F(p)}\left(dF_p(v_1), \ldots, dF_p(v_k)\right).$$

- **Proposition 3.4** ***(Properties of Tensor Pullbacks)***.
Suppose $F: M \to N$ and $G: N \to P$ are smooth maps, A and B are covariant tensor fields on N, and f is a real-valued function on N.
    1. $F^*(fB) = (f \circ F)\, F^*(B)$
    2. $F^*(A \otimes B) = F^*A \otimes F^*(B)$
    3. $F^*(A + B) = F^*A + F^*(B)$
    4. *$F^*(B)$ is a **(continuous) tensor field**, and is **smooth** if B is **smooth**.*
    5. $(G \circ F)^*B = F^*(G^*B)$.
    6. $(Id_N)^*B = B$.

- **Remark** If f is a continuous real-valued function (i.e., a 0-tensor field) and B is a k-tensor field, then it is consistent with our definitions to interpret $f \circ B$ as fB, and F^*f as $f \circ F$.

- **Corollary 3.5** ***(Coordinate Representation of Pullback Tensor Fields)***
Let $F: M \to N$ be smooth, and let B be a covariant k-tensor field on N. If $p \in M$ and (y^i) are smooth coordinates for N on a neighborhood of $F(p)$, then F^*B has the following expression in a neighborhood of p:
$$F^*\left(B_{i_1,\ldots,i_k}\, dy^{i_1} \otimes \ldots \otimes dy^{i_k}\right) = \left(B_{i_1,\ldots,i_k} \circ F\right) d\left(y^{i_1} \circ F\right) \otimes \ldots \otimes \left(y^{i_k} \circ F\right).$$

- **Remark** F^*B is computed as follows: whereaver you see y^i in the expression for B, just substitute the ith component function of F and expand.

- **Exercise 3.6** ***(Pullback of a Tensor Field)***.
Let $M = set(r, \theta) : r > 0, |\theta| < \pi/2$ and $N = \{(x, y) : x > 0\}$, and let $F: M \to \mathbb{R}^2$ be the smooth map $F(r, \theta) = (r\cos(\theta), r\sin(\theta))$. The pullback of the tensor field $A = x^{-2}dy \otimes dy$ by F can be computed easily by substituting $x = r\cos(\theta)$, $y = r\sin(\theta)$ and simplifying:
$$\begin{aligned} F^*(A) = F^*\left(x^{-2}dy \otimes dy\right) &= (r\cos(\theta))^{-2}\ d\left(r\sin(\theta)\right) \otimes d\left(r\sin(\theta)\right) \\ &= (r\cos\theta)^{-2}\ (\sin\theta dr + r\cos\theta d\theta) \otimes (\sin\theta dr + r\cos\theta d\theta) \\ &= r^{-2}\tan^2\theta dr \otimes dr + r^{-1}\tan\theta\,(d\theta \otimes dr + dr \otimes d\theta) + d\theta \otimes d\theta \quad \blacksquare \end{aligned}$$

3.3 Contraction

- **Proposition 3.7** *Let V be a finite-dimensional vector space. There is a natural (basis-independent)* ***isomorphism*** *between $T^{(k+1,l)}V$ and the space of* ***multilinear*** *maps*

$$\underbrace{V^* \times \ldots \times V^*}_{k} \times \underbrace{V \times \ldots \times V}_{l} \to V$$

- **Remark** For instance, there is an *isomorphism* : $T^{(1,1)}V \to L(V,V)$ that $(V \otimes \omega) \mapsto (F^i_j)$ where under basis (E_i) and co-basis (ϵ^j)

$$\begin{aligned} &V \otimes \omega = F^i_j \, E_i \otimes \epsilon^j \\ &\Rightarrow F^i_j = (V \otimes \omega)(\epsilon^i, E_j) = \epsilon^i(V)\, \omega(E_j) = V^i \, \omega_j = (\omega(V))^i_j \end{aligned}$$

- **Definition** We can use the result of Proposition 3.7 to define a natural operation called ***trace*** or ***contraction***, which *lowers the rank of a tensor by* 2.

  For $F = v \otimes \omega \in T^{(1,1)}V$. Define the operator tr : $T^{(1,1)}V \to \mathbb{R}$ is just ***the trace of*** F for i.e. the sum of the diagonal entries of any matrix representation of F. More generally, we define tr $: T^{(k+1,l+1)}V \to T^{(k,l)}V$ by letting tr $F(\omega^1, \ldots, \omega^k, v_1, \ldots, v_l)$ be the ***trace*** of the $(1,1)$***-tensor***

$$F(\omega^1, \ldots, \omega^k, \cdot, v_1, \ldots, v_l, \cdot) \in T^{(1,1)}V$$

  In terms of a basis, the ***components*** of tr F are

$$(\text{tr } F)^{i_1,\ldots,i_k}_{j_1,\ldots,j_l} = F^{i_1,\ldots,i_k,m}_{j_1,\ldots,j_l,m}.$$

  In other words, just ***set the last upper and lower indices equal*** and ***sum***.

- **Remark** We consider a $(1,1)$-tensor $F = v \otimes \omega$. Under standard basis, $v = v^i E_i$ and $\omega = \omega_j \, \epsilon^j$, F has representation

$$\begin{aligned} F &= v \otimes \omega \\ &= (v^i E_i) \otimes (\omega_j \, \epsilon^j) \\ &= (\omega_j \, v^i) E_i \otimes \epsilon^j := F^i_j \, E_i \otimes \epsilon^j. \end{aligned}$$

  There is an isomorphism $T^{(1,1)}V \to L(V;V)$ as $F \mapsto [F^i_j]_{j,i}$. Then the ***trace*** of F is

$$\begin{aligned} \text{tr } (v \otimes \omega) &= \omega(v) \\ &= \omega_i \, v^i \\ &= \text{tr } \left(\begin{bmatrix} \omega_1 \, v^1 & \ldots & \omega_1 \, v^n \\ \vdots & \ddots & \vdots \\ \omega_n \, v^1 & \ldots & \omega_n \, v^n \end{bmatrix} \right) = \text{tr } [F^i_j]_{j,i}. \end{aligned}$$

3.4 Lie Derivatives of Tensor Fields

References

John Marshall Lee. *Introduction to smooth manifolds.* Graduate texts in mathematics. Springer, New York, Berlin, Heidelberg, 2003. ISBN 0-387-95448-1.