Lecture 17

Today we begin studying the material that is also found in the Multi-linear Algebra Notes. We begin with the theory of *tensors*.

4.3 Tensors

Let V be a n-dimensional vector space. We use the following notation.

Notation.

$$V^k = \underbrace{V \times \cdots \times V}_{k \text{ times}}. \tag{4.14}$$

For example,

$$V^2 = V \times V, \tag{4.15}$$

$$V^3 = V \times V \times V. \tag{4.16}$$

Let $T : V^k \to \mathbb{R}$ be a map.

Definition 4.1. The map T is *linear in its ith factor* if for every sequence $v_j \in V, 1 \leq j \leq n, j \neq i$, the function mapping $v \in V$ to $T(v_1, \ldots, v_{i-1}, v, v_{i+1}, \ldots, v_k)$ is linear in v.

Definition 4.2. The map T is *k-linear* (or is a *k-tensor*) if it is linear in all k factors.

Let T_1, T_2 be k-tensors, and let $\lambda_1, \lambda_2 \in \mathbb{R}$. Then $\lambda_1 T_1 + \lambda_2 T_2$ is a k-tensor (it is linear in all of its factors).

So, the set of all k-tensors is a vector space, denoted by $\mathcal{L}^k(V)$, which we sometimes simply denote by $\mathcal{L}^k$.

Consider the special case $k = 1$. The the set $\mathcal{L}^1(V)$ is the set of all linear maps $\ell : V \to \mathbb{R}$. In other words,

$$\mathcal{L}^1(V) = V^*. \tag{4.17}$$

We use the convention that

$$\mathcal{L}^0(V) = \mathbb{R}. \tag{4.18}$$

Definition 4.3. Let $T_i \in \mathcal{L}^{k_i}, i = 1, 2$, and define $k = k_1 + k_2$. We define the *tensor product* of T_1 and T_2 to be the tensor $T_1 \otimes T_2 : V^k \to \mathbb{R}$ defined by

$$T_1 \otimes T_2(v_1, \ldots, v_k) = T_1(v_1, \ldots, v_{k_1}) T_2(v_{k_1+1}, \ldots, v_k). \tag{4.19}$$

We can conclude that $T_1 \otimes T_2 \in \mathcal{L}^k$.

We can define more complicated tensor products. For example, let $T_i \in \mathcal{L}^{k_i}, i = 1, 2, 3$, and define $k = k_1 + k_2 + k_3$. Then we have the tensor product

$$\begin{aligned} &T_1 \otimes T_2 \otimes T_3(v_1, \ldots, v_k) \\ &\qquad = T_1(v_i, \ldots, v_{k_1}) T_2(v_{k_1+1}, \ldots, v_{k_1+k_2}) T_3(v_{k_1+k_2+1}, \ldots, v_k). \end{aligned} \tag{4.20}$$

Then $T_1 \otimes T_2 \otimes T_3 \in \mathcal{L}^k$. Note that we could have simply defined

$$\begin{aligned} T_1 \otimes T_2 \otimes T_3 &= (T_1 \otimes T_2) \otimes T_3 \\ &= T_1 \otimes (T_2 \otimes T_3), \end{aligned} \tag{4.21}$$

where the second equality is the associative law for tensors. There are other laws, which we list here.

- Associative Law: $(T_1 \otimes T_2) \otimes T_3 = T_1 \otimes (T_2 \otimes T_3)$.
- Right and Left Distributive Laws: Suppose $T_i \in \mathcal{L}^{k_i}, i = 1, 2, 3$, and assume that $k_1 = k_2$. Then
  - Left: $(T_1 + T_2) \otimes T_3 = T_1 \otimes T_3 + T_2 \otimes T_3$.
  - Right: $T_3 \otimes (T_1 + T_2) = T_3 \otimes T_1 + T_3 \otimes T_2$.
- Let λ be a scalar. Then

$$\lambda(T_1 \otimes T_2) = (\lambda T_1) \otimes T_2 = T_1 \otimes (\lambda T_2). \tag{4.22}$$

Now we look at an important class of k-tensors. Remember that $\mathcal{L}^1(V) = V^*$, and take any 1-tensors $\ell_i \in V^*, i = 1, \dots, k$.

Definition 4.4. The tensor $T = \ell_1 \otimes \cdots \otimes \ell_k$ is a *decomposable k-tensor*.

By definition, $T(v_1, \dots, v_k) = \ell_1(v_1) \dots \ell_k(v_k)$. That is, $\ell_1 \otimes \cdots \otimes \ell_k(v_1, \dots, v_k) = \ell_1(v_1) \dots \ell_k(v_k)$.

Now let us go back to considering $\mathcal{L}^k = \mathcal{L}^k(V)$.

Theorem 4.5.

$$\dim \mathcal{L}^k = n^k. \tag{4.23}$$

Note that for $k = 1$, this shows that $\mathcal{L}^1(V) = V^*$ has dimension n.

Proof. Fix a basis $e_1, \dots, e_n$ of V. This defines a dual basis $e_i^*, \dots, e_n^*$ of V^*, $e_i^* : V \to \mathbb{R}$ defined by

$$e_i^*(e_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{4.24}$$

Definition 4.6. A *multi-index* I of length k is a set of integers $(i_1, \dots, i_k), 1 \leq i_r \leq n$. We define

$$e_I^* = e_{i_1}^* \otimes \cdots \otimes e_{i_k}^* \in \mathcal{L}^k. \tag{4.25}$$

Let $J = (j_1, \dots, j_k)$ be a multi-index of length k. Then

$$e_I^*(e_{j_1}, \dots, e_{j_k}) = e_{i_1}^*(e_{j_1}) \dots e_{i_k}^*(e_{j_k}) = \begin{cases} 1 & \text{if } I = J, \\ 0 & \text{if } I \neq J. \end{cases} \tag{4.26}$$

Claim. *The k-tensors e_I^* are a basis of $\mathcal{L}^k$.*

Proof. To prove the claim, we use the following lemma.

Lemma 4.7. *Let T be a k-tensor. Suppose that $T(e_{i_1}, \ldots, e_{i_k}) = 0$ for all multi-indices I. Then $T = 0$.*

Proof. Define a $(k-1)$-tensor $T_i : V^{k-1} \to \mathbb{R}$ by setting

$$T_i(v_1, \ldots, v_{k-1}) = T(v_1, \ldots, v_{k-1}, e_j), \tag{4.27}$$

and let $v_k = \sum a_i e_i$. By linearity, $T(v_1, \ldots, v_k) = \sum a_i T_i(v_1, \ldots, v_{k-1})$. So, if the lemma is true for the T_i's, then it is true for T by an induction argument (we leave this to the student to prove). □

With this lemma we can prove the claim.

First we show that the e_I^*'s are linearly independent. Suppose that

$$0 = T = \sum c_I e_I^*. \tag{4.28}$$

For any multi-index J of length k,

$$\begin{aligned} 0 &= T(e_{j_1}, \ldots, e_{j_k}) \\ &= \sum c_I e_I^*(e_{j_1}, \ldots, e_{j_k}) \\ &= c_J \\ &= 0. \end{aligned} \tag{4.29}$$

So the e_I^*'s are linearly independent.

Now we show that the e_I^*'s span $\mathcal{L}^k$. Let $T \in \mathcal{L}^k$. For every I let $T_I = T(e_{i_1}, \ldots, e_{i_l})$, and let $T' = \sum T_I e_I^*$. One can check that $(T - T')(e_{j_1}, \ldots, e_{j_k}) = 0$ for all multi-indices J. Then the lemma tells us that $T = T'$, so the e_I^*'s span $\mathcal{L}^k$, which proves our claim. □

Since the e_I^*'s are a basis of $\mathcal{L}^k$, we see that

$$\dim \mathcal{L}^k = n^k, \tag{4.30}$$

which proves our theorem. □

4.4 Pullback Operators

Let V, W be vector spacers, and let $A : V \to W$ be a linear map. Let $T \in \mathcal{L}^k(W)$, and define a new map $A^*T \in \mathcal{L}^k(V)$ (called the "pullback" tensor) by

$$A^*T(v_1, \dots, v_k) = T(Av_1, \dots, Av_k). \tag{4.31}$$

You should prove the following claims as an exercise:

Claim. *The map $A^* : \mathcal{L}^k(W) \to \mathcal{L}^k(V)$ is a linear map.*

Claim. *Let $T_i \in \mathcal{L}^{k_i}(W), i = 1, 2$. Then*

$$A^*(T_1 \otimes T_2) = A^*T_1 \otimes A^*T_2. \tag{4.32}$$

Now, let $A : V \to W$ and $B : W \to U$ be maps, where U is a vector space. Given $T \in \mathcal{L}^k(U)$, we can "pullback" to W by B^*T, and then we can "pullback" to V by $A^*(B^*T) = (B \circ A)^*T$.

4.5 Alternating Tensors

In this course we will be restricting ourselves to *alternating tensors*.

Definition 4.8. A permutation of order k is a bijective map

$$\sigma : \{1, \dots, k\} \to \{1, \dots, k\}. \tag{4.33}$$

The map is a bijection, so σ^{-1} exists.

Given two permutations σ_1, σ_2, we can construct the composite permutation

$$\sigma_1 \circ \sigma_2(i) = \sigma_1(\sigma_2(i)). \tag{4.34}$$

We define

$$S_k \equiv \text{ The set of all permutations of } \{1, \dots, k\}. \tag{4.35}$$

There are some special permutations. Fix $1 \le i < j \le k$. Let τ be the permutation such that

$$\tau(i) = j \tag{4.36}$$

$$\tau(j) = i \tag{4.37}$$

$$\tau(\ell) = \ell, \ell \neq i, j. \tag{4.38}$$

The permutation τ is called a *transposition*.

Definition 4.9. The permutation τ is an elementary transposition if $j = i + 1$.

We state without proof two very useful theorems.

Theorem 4.10. *Every permutation can be written as a product* $\sigma = \tau_1 \circ \tau_2 \circ \cdots \circ \tau_m$*, where each* τ_i *is an elementary transposition.*

Theorem 4.11. *Every permutation* σ *can be written either as a product of an even number of elementary transpositions or as a product of an odd number of elementary transpositions, but not both.*

Because of the second theorem, we can define an important invariant of a permutation: the *sign of the permutation*.

Definition 4.12. If $\sigma = \tau_1 \circ \cdots \circ \tau_m$, where the τ_i's are elementary transpositions, then the *sign of* σ is

$$\text{sign of } \sigma = (-1)^\sigma = (-1)^m. \tag{4.39}$$

Note that if $\sigma = \sigma_1 \circ \sigma_2$, then $(-1)^\sigma = (-1)^{\sigma_1}(-1)^{\sigma_2}$. We can see this by letting $\sigma_1 = \tau_1 \circ \cdots \circ \tau_{m_1}$, and $\sigma_2 = \tau'_1 \circ \cdots \circ \tau'_{m_2}$, and noting that $\sigma_1 \circ \sigma_2 = \tau_1 \circ \cdots \circ \tau_{m_1} \circ \tau'_1 \circ \cdots \circ \tau'_{m_2}$.