# Chapter 1

# Multilinear Algebra

## 1.1  Notes

• Motivation for the course and an overview of Guillemin and Haine (2018).

• Plan:

  – More (multi)linear algebra.

• Dual spaces.

• Let $V$ be an $n$-dimensional real vector space.

• **Hom $(V, \mathbb{R})$**: The set of all homomorphisms (i.e., linear maps) from $V$ to $\mathbb{R}$. *Also known as $V^*$.*

• **Dual basis** (for $V^*$): The set of linear transformations from $V$ to $\mathbb{R}$ defined by

$$e_j \mapsto \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases}$$

  where $e_1, \ldots, e_n$ is a basis of $V$. *Denoted by $e_1^*, \ldots, e_n^*$.*

• Check: $e_1^*, \ldots, e_n^*$ are a basis for $V^*$.

  – Are they linearly independent? Let $c_1 e_1^* + \cdots + c_n e_n^* = 0 \in \text{Hom}(V, \mathbb{R})$. Then

$$c_i = (c_1 e_1^* + \cdots + c_n e_n^*)(e_i) = 0 \in \mathbb{R}$$

  as desired.

  – Span? Let $\varphi \in \text{Hom}(V, \mathbb{R})$. Then we can verify that

$$\varphi(e_1)e_1^* + \cdots + \varphi(e_n)e_n^* = \varphi$$

  ■ We prove this by verifying the previous statement on the basis of $V$ (if two linear transformations have the same action on the basis of a vector space, they are equal).

• With a choice of basis for $V$, we obtain an isomorphism $\varepsilon : V \to V^*$ with the mapping $e_i \mapsto e_i^*$ for all $i$.

• The dual space is known as such because $(V^*)^* \cong V$, where $\cong$ is **canonical** (no choice of basis is needed).

• One more property of dual spaces: **functoriality**.

- – Given a linear transformation $A : V \to W$, we know that $A^* : W^* \to V^*$ where $A^*$ is the transpose of $A$. In particular, if $\varphi \in W^*$, then $\varphi \circ A : V \to \mathbb{R}$.

  – Claim: $A^*$ is linear.

- **Functoriality**: If $A : V \to W$ and $B : W \to U$, then $B^* : U^* \to W^*$ and $A^* : W^* \to V^*$. The functoriality statement is that $(B \circ A)^* = A^* \circ B^*$.

- $A^*$ is the **pullback** (or transpose) of $A$.

- Let $v_1, \ldots, v_n$ be a basis for $V$ and $w_1, \ldots, w_m$ be a basis for $W$. Then $[A]_{v_1,\ldots,v_n}^{w_1,\ldots,w_m} = A$ is the matrix of the linear transformation $A$ with respect to these bases. Then if $v_1^*, \ldots, v_n^*$ and $w_1^*, \ldots, w_m^*$ are the corresponding dual bases, then $[A^*]_{v_1^*,\ldots,v_n^*}^{w_1^*,\ldots,w_m^*} = A^T$. We can and should verify this for ourselves.

- This is over the real numbers, so $A^*$ is just the transpose because there are no complex numbers of which to take the conjugate!

- A generalization: Tensors.

- **$k$-tensor**: A **multilinear** map
$$T : \underbrace{V \times \cdots \times V}_{k \text{ times}} \to \mathbb{R}$$

- **Multilinear** (map $T$): A function $T$ such that
$$T(v_1, \ldots, v_i^1 + v_i^2, \ldots, v_k) = T(v_1, \ldots, v_i^1, \ldots, v_k) + T(v_1, \ldots, v_i^2, \ldots, v_k)$$
$$T(v_1, \ldots, \lambda v_i, \ldots, v_k) = \lambda T(v_1, \ldots, v_i, \ldots, v_k)$$

  for all $(v_1, \ldots, v_k) \in V^k$.

- The determinant is an $n$-tensor!

- 1-tensors are just covectors.

- **$L^k(V)$**: The vector space of all $k$-tensors on $V$.

- Calculating $\dim L^k(V)$. (Answer not given in this class.)

- Let $A : V \to W$. Then $A^* : L^k(W) \to L^k(V)$.

  – Check $(A \circ B)^* = B^* \circ A^*$.

- **Multi-index of $n$ of length $k$**: A $k$-tuple $(i_1, \ldots, i_k)$ where each $i_j \in \mathbb{N}$ satisfies $1 \leq i_j \leq n$ $(j = 1, \ldots, k)$. *Denoted by $I$.*

- Let $e_1, \ldots, e_n$ be a basis for $V$.

- **Tensor product** (of $T_1 \in L^k(V)$, $T_2 \in L^l(V)$): The function from $V^{k+l}$ to $\mathbb{R}$ defined by
$$(v_1, \ldots, v_{k+l}) \mapsto T_1(v_1, \ldots, v_k)T_2(v_{k+1}, \ldots, v_{k+l})$$

  *Denoted by $T_1 \otimes T_2$.*

- Claims:

  1. $T_1 \otimes T_2 \in L^{k+l}(V)$.
  2. $A^*(T_1 \otimes T_2) = A^*(T_1) \otimes A^*(T_2)$.

- **$e_I^*$**: The function $e_{i_1}^* \otimes \cdots \otimes e_{i_k}^*$, where $I = (i_1, \ldots, i_k)$ is a multi-index of $n$ of length $k$.

- Claim: Letting $I$ range over all $n^k$ multi-indices of $n$ of length $k$, the $e_I^*$ are a basis for $L^k(V)$.

- If $V = \mathbb{R}$, then $V = \mathbb{R}e_1$. If $V = \mathbb{R}^2$, then $V = \mathbb{R}e_1 \oplus \mathbb{R}e_2$.

- We know that $L^1(V) = V^* = Re_1^*$. Thus, $e_1^* \otimes e_2^* : V \times V \to \mathbb{R}$. Thus, for example,

$$(e_1^* \otimes e_2^*)((1,2),(3,4)) = e_1^*(1,2) \cdot e_2^*(3,4) = 1 \cdot 4 = 4$$

4/1:
- Plan: More multilinear algebra.

  - Properties of the tensor product.
  - Sign of a permutation.
  - Alternating tensors (lead into differential forms down the road).

- Recall: $V$ is an $n$-dimensional vector space over $\mathbb{R}$ with basis $e_1, \ldots, e_n$. $\mathcal{L}^k(V)$ is the vector space of $k$-tensors on $V$. $\{e_I^* \mid I \text{ a multiindex of } n \text{ of length } k\}$ is a basis for $\mathcal{L}^k(V)$.

- For example, if $V = \mathbb{R}^2$ and $T \in \mathcal{L}^2(V)$, then

$$T(a_1e_1 + a_2e_2, b_1e_1 + b_2e_2) = a_1b_1T(e_1,e_1) + a_1b_2T(e_1,e_2) + a_2b_1T(e_2,e_1) + a_2b_2T(e_2,e_2)$$

  - A basis of $\mathcal{L}^2(V)$ is
$$\{e_1^* \otimes e_1^*, e_1^* \otimes e_2^*, e_2^* \otimes e_1^*, e_2^* \otimes e_2^*\}$$

  - Recall that some basic properties are

$$e_1^* \otimes e_2^*((1,2),(3,4)) = 1 \cdot 4 = 4 \qquad\qquad e_2^* \otimes e_1^*((1,2),(3,4)) = 2 \cdot 3 = 6$$

  - It follows by the initial decomposition of $T$ that

$$T = a_1b_1e_1^* \otimes e_1^* + a_1b_2e_1^* \otimes e_2^* + a_2b_1e_2^* \otimes e_1^* + a_2b_2e_2^* \otimes e_2^*$$

- Important consequence: To know the action of $T$ on an arbitrary pair of vectors, you need only know its action on the basis; a higher-dimensional generalization of the earlier property.

- Note that
$$e_I^*(e_J) = \delta_{IJ} = \begin{cases} 1 & I = J \\ 0 & I \neq J \end{cases}$$

- Basic properties of the tensor product.

  1. *Right-distributive*: If $T_1 \in \mathcal{L}^k(V)$ and $T_2, T_3 \in \mathcal{L}^\ell(V)$, then

  $$T_1 \otimes (T_2 + T_3) = T_1 \otimes T_2 + T_1 \otimes T_3$$

  2. *Left-distributive*: If $T_1, T_2 \in \mathcal{L}^k(V)$ and $T_3 \in \mathcal{L}^\ell(V)$, then

  $$(T_1 + T_2) \otimes T_3 = T_1 \otimes T_3 + T_2 \otimes T_3$$

  3. *Associative*: If $T_1 \in \mathcal{L}^k(V)$, $T_2 \in \mathcal{L}^\ell(V)$, and $T_3 \in \mathcal{L}^m(V)$, then

  $$T_1 \otimes (T_2 \otimes T_3) = (T_1 \otimes T_2) \otimes T_2 = T_1 \otimes T_2 \otimes T_3$$

  4. *Scalar multiplication*: If $T_1 \in \mathcal{L}^k(V)$, $T_2 \in \mathcal{L}^\ell(V)$, and $\lambda \in \mathbb{R}$, then

  $$(\lambda T_1) \otimes T_2 = \lambda(T_1 \otimes T_2) = T_1 \otimes (\lambda T_2)$$

- Note that the tensor product is not commutative.

- Aside: Defining the sign of a permutation.

- $S_A$: The set of all automorphisms of $A$ (bijections from $A$ to $A$), where $A$ is a set.

- $S_n$: The set $S_{[n]}$.

- Given $\sigma_1, \sigma_2 \in S_n$, $\sigma_1 \circ \sigma_2 \in S_n$.

  - Thus, $S_n$ is a **group**.

- **Transposition**: A function $\tau \in S_n$ such that

$$\tau(k) = \begin{cases} j & k = i \\ i & k = j \\ k & k \neq i, j \end{cases}$$

  for some $i, j \in [n]$. *Denoted by* $\boldsymbol{\tau_{i,j}}$.

- Theorem: An element of $S_n$ can be written as the product of transpositions (i.e., for all $\sigma \in S_n$, there exist $\tau_1, \ldots, \tau_m \in S_n$ such that $\sigma = \tau_1 \circ \cdots \circ \tau_m$).

- **Sign** (of $\sigma \in S_n$): The number (mod 2) of transpositions whose product equals $\sigma$. *Denoted by* $\boldsymbol{(-1)^{\sigma}}$, $\mathbf{sign}\,\boldsymbol{(\sigma)}$.

- Theorem: The sign of $\sigma$ is well-defined. Additionally,

$$(-1)^{\sigma_1 \sigma_2} = (-1)^{\sigma_1} \cdot (-1)^{\sigma_2}$$

- Example: Consider the identity permutation. $(-1)^{\sigma} = +1$. We can think of this as the product of zero transpositions or, for instance, as the product of the two transpositions $\tau_{1,2} \circ \tau_{1,2}$. Another example would be $\tau_{2,3} \circ \tau_{1,2} \circ \tau_{1,2} \circ \tau_{2,3}$.

- Theorem: Let $X_i$ be a rational or polynomial function for each $i \in \mathbb{N}$. Then

$$(-1)^{\sigma} = \prod_{i<j} \frac{X_{\sigma(i)} - X_{\sigma(j)}}{X_i - X_j}$$

- Example: For the permutation $\sigma = (1, 2, 3)$, we have

$$\begin{aligned} (-1)^{\sigma} &= \frac{X_{\sigma(1)} - X_{\sigma(2)}}{X_1 - X_2} \cdot \frac{X_{\sigma(1)} - X_{\sigma(3)}}{X_1 - X_3} \cdot \frac{X_{\sigma(2)} - X_{\sigma(3)}}{X_2 - X_3} \\ &= \frac{X_2 - X_3}{X_1 - X_2} \cdot \frac{X_2 - X_1}{X_1 - X_3} \cdot \frac{X_3 - X_1}{X_2 - X_3} \\ &= \frac{-(X_1 - X_2)}{X_1 - X_2} \cdot \frac{-(X_1 - X_3)}{X_1 - X_3} \cdot \frac{X_2 - X_3}{X_2 - X_3} \\ &= -1 \cdot -1 \cdot 1 \\ &= +1 \end{aligned}$$

  which squares with the fact that $\sigma = \tau_{1,2} \circ \tau_{2,3}$.

- Claims to verify with the above formula:

  1. $\text{sign}(\sigma) \in \{\pm 1\}$.
  2. $\text{sign}(\tau_{i,j}) = -1$.
  3. $\text{sign}(\sigma_1 \sigma_2) = \text{sign}(\sigma_1)\,\text{sign}(\sigma_2)$.

4/4:
- Plan:

  - More multilinear algebra.

- Alternating $k$-tensors — 2 views:
    1. As a subspace of $\mathcal{L}^k(V)$.
    2. As a quotient of $\mathcal{L}^k(V)$.
- Next time: Operators as alternating tensors.
    - Wedge products.
    - Interior products.
    - Pullbacks.

- Recall: $\dim V = n$, $e_1, \ldots, e_n$ a basis, $\mathcal{L}^k(V)$ the space of $k$-tensors, $\sigma \in S_k$ implies $(-1)^\sigma \in \{\pm 1\}$, key property: $(-1)^{\sigma_1 \sigma_2} = (-1)^{\sigma_1}(-1)^{\sigma_2}$.

- $\boldsymbol{T^\sigma}$: The $k$-tensor over $V$ defined by

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

  where $T \in \mathcal{L}^k(V)$, $\sigma \in S_k$, and $\bar{\sigma}$ denotes the inverse of $\sigma$.

- Example: $n = 2$, $k = 2$. Let $T = e_1^* \otimes e_2^* \in \mathcal{L}^2(V)$. Let $\sigma = \tau_{1,2}$. Then $T^\sigma = e_2^* \otimes e_1^*$.

- Another property is $e_I^\sigma = e_{\sigma(I)}^*$.

- Properties:

    1. $T^{\sigma_1 \sigma_2} = (T^{\sigma_1})^{\sigma_2}$.
    2. $(T_1 + T_2)^\sigma = T_1^\sigma + T_2^\sigma$.
    3. $(cT)^\sigma = cT^\sigma$.

- Thus, you can view $\sigma : \mathcal{L}^k(V) \to \mathcal{L}^k(V)$ as a linear map!

- **Alternating $k$-tensor**: A tensor $T \in \mathcal{L}^k(V)$ such that $T^\sigma = (-1)^\sigma T$ for all $\sigma \in S_k$.

    - Equivalently, $T^\tau = -T$ for all $\tau \in S_k$.

- An example of an alternating 2-tensor when $\dim V = 2$ is $T = e_1^* \otimes e_2^* - e_2^* \otimes e_1^*$.

    - Naturally, $T_{1,2}^\tau = -T$, and $\tau_{1,2}$ is the unique transposition in $S_2$.

- $e_1^* \otimes e_2^*$ is *not* an alternating 2-tensor since $(e_1^* \otimes e_2^*)^\tau = e_2^* \otimes e_1^* \neq (-1)^\tau (e_1^* \otimes e_2^*)$.

- We can look at $n = 2$, $k = 1$ for ourselves.

- Note: If $T - 1, T_2$ are both alternating $k$-tensors, then $T_1 + T_2$ is also alternating, as is $cT_1$ for all $c \in \mathbb{R}$.

- $\boldsymbol{\mathcal{A}^k(V)}$: The vector space of alternating $k$-tensors.

- **Alt $(\boldsymbol{T})$**: The function $\mathrm{Alt} : \mathcal{L}^k(V) \to \mathcal{L}^k(V)$ defined by

$$\mathrm{Alt}(T) = \sum_{\sigma \in S_k} (-1)^\sigma T^\sigma$$

- Properties:

    1. $\mathrm{im}(\mathrm{Alt}) = \mathcal{A}^k(V)$.
    2. $\mathcal{L}^k(V)/\ker(\mathrm{Alt}) = \Lambda^k(V^*)$ is isomorphic to $\mathcal{A}^k(V)$.
    3. $\mathrm{Alt}(T)^\sigma = (-1)^\sigma \mathrm{Alt}(T)$.

– Proof:

$$\text{Alt}(T)^{\sigma'} = \left( \sum_{\sigma \in S_k} (-1)^\sigma T^\sigma \right)^{\sigma'}$$
$$= \sum_{\sigma \in S_k} (-1)^\sigma T^{\sigma\sigma'}$$
$$= (-1)^{\sigma'} \sum_{\sigma \in S_k} (-1)^{\sigma'} (-1)^\sigma T^{\sigma\sigma'}$$
$$= (-1)^{\sigma'} \sum_{\sigma \in S_k} (-1)^{\sigma\sigma'} T^{\sigma\sigma'}$$
$$= (-1)^{\sigma'} \text{Alt}(T)$$

– The last equality holds because summing over all $\sigma$ is the same as summing over all $\sigma' \circ \sigma$.
– This implies $\text{im}(\text{Alt}) \leq \mathcal{A}^k(V)$.

4. If $T \in \mathcal{A}^k(T)$, $\text{Alt}(T) = k!T$.

– We have

$$\text{Alt}(T) = \sum_{\sigma \in S_k} (-1)^\sigma T^\sigma$$
$$= \sum_{\sigma \in S_k} (-1)^\sigma (-1)^\sigma T$$
$$= \sum_{\sigma \in S_k} T$$
$$= k!T$$

where $T^\sigma = (-1)^\sigma T$ since $T \in \mathcal{A}^k(V)$ by definition.
– This implies that $\text{im}(\text{Alt}) = \mathcal{A}^k(V)$: $\text{Alt}(\frac{1}{k!}T) = T \in \mathcal{A}^k(V)$.

5. $\text{Alt}(T^\sigma) = \text{Alt}(T)^\sigma$.

6. $\text{Alt} : \mathcal{L}^k(V) \to \mathcal{L}^k(V)$ is linear.

- Warning: Some people take $\text{Alt}(T) = \frac{1}{k!} \sum_{\sigma \in S_k} (-1)^\sigma T^{\sigma}$[1].

- Example: $n = k = 2$. We have

$$\text{Alt}(e_1^* \otimes e_2^*) = e_1^* \otimes e_2^* - e_2^* \otimes e_1^*$$

- **Non-repeating** (multi-index $I$): A multi-index $I$ such that $i_{j_1} \neq i_{j_2}$ for all $j_1 \neq j_2$.

- **Increasing** (multi-index $I$): A multi-index $I$ such that $i_1 < \cdots < i_k$.

- Claim: $\{\text{Alt}(e_I^*)\}$ where $I$ is non-repeating and increasing is a basis for $\mathcal{A}^k(V)$. There are $\binom{n}{k}$ of these; thus, $\dim \mathcal{A}^k(V) = \binom{n}{k}$.

## 1.2 Chapter 1: Multilinear Algebar

*From Guillemin and Haine (2018).*

3/31:
- Guillemin and Haine (2018) defines real vector spaces, the operations on them, their basic properties, and the zero vector.

---

[1]Klug prefers this convention, but the text takes the other one.

- **Linearly independent** (vectors $v_1, \ldots, v_k$): A finite set of vectors $v_1, \ldots, v_k \in V$ such that the map from $\mathbb{R}^k$ to $V$ defined by $(c_1, \ldots, c_k) \mapsto c_1 v_1 + \cdots + c_k v_k$ is injective.

- **Spanning** (vectors $v_1, \ldots, v_k$): We require that the above map is surjective.

- Guillemin and Haine (2018) defines basis, finite-dimensional vector space, dimension, subspace, linear map, and kernel.

- **Image** (of $A : V \to W$): The range space of $A$, a subspace of $W$. *Also known as* $\mathbf{im}\,(\boldsymbol{A})$.

- Guillemin and Haine (2018) defines the matrix of a linear map.

- **Inner product** (on $V$): A map $B : V \times V \to \mathbb{R}$ with the following three properties.

    - *Bilinearity*: For vectors $v, v_1, v_2, w \in V$ and $\lambda \in \mathbb{R}$, we have

    $$B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w)$$

    and
    $$B(\lambda v, w) = \lambda B(v, w)$$

    - *Symmetry*: For vectors $v, w \in V$, we have $B(v, w) = B(w, v)$.
    - *Positivity*: For every vector $v \in V$, we have $B(v, v) \geq 0$. Moreover, if $v \neq 0$, then $B(v, v) > 0$.

- **$\boldsymbol{W}$-coset**: A set of the form $\{v + w \mid w \in W\}$, where $W$ is a subspace $V$ and $v \in V$. *Denoted by* $\boldsymbol{v + W}$.

    - If $v_1 - v_2 \in W$, then $v_1 + W = v_2 + W$.
    - It follows that the distinct $W$-cosets decompose $V$ into a disjoint collection of subsets of $V$.

- **Quotient space** (of $V$ by $W$): The set of distinct $W$-cosets in $V$, along with the following definitions of vector addition and scalar multiplication.

    $$(v_1 + W) + (v_2 + W) = (v_1 + v_2) + W \qquad\qquad \lambda(v + W) = (\lambda v) + W$$

    *Denoted by* $\boldsymbol{V/W}$.

- **Quotient map**: The linear map $\pi : V \to V/W$ defined by

    $$\pi(v) = v + W$$

    - $\pi$ is surjective.
    - Note that $\ker(\pi) = W$ since for all $w \in W$, $\pi(w) = w + W = 0 + W$, which is the zero vector in $V/W$.

- If $V, W$ are finite dimensional, then

    $$\dim(V/W) = \dim(V) - \dim(W)$$

- Proposition 1.2.9: Let $A : V \to U$ be a linear map. If $W \subset \ker(A)$, then there exists a unique linear map $A^\sharp : V/W \to U$ with the property that $A = A^\sharp \circ \pi$, where $\pi : V \to V/W$ is the quotient map.

    - This proposition rephrases in terms of quotient spaces the fact that if $w \in W$, then $A(v+w) = Av$.

- **Dual space** (of $V$): The set of all linear functions $\ell : V \to \mathbb{R}$, along with the following definitions of vector addition and scalar multiplication.

    $$(\ell_1 + \ell_2)(v) = \ell_1(v) + \ell_2(v) \qquad\qquad (\lambda \ell)(v) = \lambda \cdot \ell(v)$$

    *Denoted by* $V^*$.

- **Dual basis** (of $e_1, \ldots, e_n$ a basis of $V$): The basis of $V^*$ consisting of the $n$ functions that take every $v = c_1 e_1 + \cdots + c_n e_n$ to one of the $c_i$. *Denoted by* $\boldsymbol{e_1^*}, ..., \boldsymbol{e_n^*}$. *Given by*

$$e_i^*(v) = c_i$$

  for all $v \in V$.

- Claim 1.2.12: If $V$ is an $n$-dimensional vector space with basis $e_1, \ldots, e_n$, then $e_1^*, \ldots, e_n^*$ is a basis of $V^*$.

  *Proof.* We will first prove that $e_1^*, \ldots, e_n^*$ spans $V^*$. Let $\ell \in V^*$ be arbitrary. Set $\lambda_i = \ell(e_i)$ for all $i \in [n]$. Define $\ell' = \sum_{i=1}^n \lambda_i e_i^*$. Then

$$\ell'(e_j) = \sum_{i=1}^n \lambda_i e_i^*(e_j) = \lambda_j \cdot 1 = \ell(e_j)$$

  for all $j \in [n]$. Therefore, since $\ell, \ell'$ take identical values on the basis of $V$, $\ell = \ell'$, as desired.

  We now prove that $e_1^*, \ldots, e_n^*$ spans $V^*$. Let $\sum_{i=1}^n \lambda_i e_i^* = 0$. Then for all $j \in [n]$,

$$\lambda_j = \left( \sum_{i=1}^n \lambda_i e_i^* \right)(e_j) = 0$$

  as desired. $\qquad\square$

- **Transpose** (of $A$): The map from $W^*$ to $V^*$ defined by $\ell \mapsto A^* \ell = \ell \circ A$ for all $\ell \in W^*$.

- Claim 1.2.15: If $e_1, \ldots, e_n$ is a basis of $V$, $f_1, \ldots, f_m$ is a basis of $W$, $e_1^*, \ldots, e_n^*$ and $f_1^*, \ldots, f_m^*$ are the corresponding dual bases, and $[a_{i,j}]$ is the $m \times n$ matrix of $A$ with respect to $\{e_i\}, \{f_i\}$, then the linear map $A^*$ is defined in terms of $\{f_i^*\}, \{e_i^*\}$ by the transpose matrix $(a_{j,i})$.