Section 3: Multilinear forms

Matthew Macauley

School of Mathematical & Statistical Sciences
Clemson University
http://www.math.clemson.edu/~macaule/

Math 8530, Advanced Linear Algebra

Overview

One of the goals of this section is to understand the concept of the determinant in a basis-free manner.

Formally, the determinant is the *unique normalized alternating n-linear form* satifying a particular "universal property".

To get there, we'll explore the concept of a multilinear, or k-linear form.

This actually generalizes several familiar concepts:

- A 1-linear form is just a scalar function $X \to K$.
- A 2-linear form is just a bilinear function $X \times X \to K$.

We'll have to understand various types of multilinear forms: symmetric, skew-symmetric, and alternating.

Before we can do this, we will cover two prerequesites:

- an overview as to what the determinant means geometrically (for motivation)
- a crash course on permutations.

Later on, we'll see related concepts such as the trace and tensors.

What is a determinant?

Definition (unofficial)

The determinant of $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is the signed volume of $T([0,1]^n)$, the image of the unit n-cube.

Permutations

Definition

Let $[n] := \{1, \dots, n\}$. A permutation is a bijection $\pi\colon [n] \to [n]$. The set of all $n!$ permutations is the symmetric group, S_n.

Definition

The discriminant of variables $x_1, \dots, x_n$ is

$$P(x_1, \dots, x_n) = \prod_{i<j} (x_i - x_j).$$

Permuting variables only changes the sign of the discriminant:

$$P\big(\pi(x_1, \dots, x_n)\big) = \prod_{i<j} (x_{\pi(i)} - x_{\pi(j)}) = \underbrace{\operatorname{sgn}(\pi)}_{\pm 1} \prod_{i<j} (x_i - x_j).$$

We call $\operatorname{sgn}(\pi)$ the sign of the permutation π.

Transpositions

A transposition is a permutation $\tau \in S_n$ that swaps two entries and fixes the rest. That is,

$$\tau(i) = j, \quad \tau(j) = i, \qquad \tau(k) = k, \text{ if } k \neq i, j.$$

We write this as (ij).

Proposition (HW)

(i) $\operatorname{sgn}(\pi_1 \circ \pi_2) = \operatorname{sgn}(\pi_1)\operatorname{sgn}(\pi_2)$

(ii) $\operatorname{sgn}(\tau) = -1$ for any transposition

(iii) every $\pi \in S_n$ can be written as a composition of transpositions: $\pi = \tau_k \circ \cdots \circ \tau_1$

(iv) the parity of this decomposition is unique

(v) if $\pi = \tau_k \circ \cdots \circ \tau_1$, then $\operatorname{sgn}(\pi) = (-1)^k$.

Multilinearity

Loosely speaking, linearity means we can pull apart sums and constants. We have seen:

1. Dual vectors: linear scalar functions $X \to K$
2. Scalar products: bilinear functions $U' \times X \to K$

n. Determinants: functions on n (row or column) vectors where we can break apart certain sums and pull out constants.

These are all examples of multilinear functions.

The determinant is actually a property of a linear map, not a matrix. In this section, we will define and study the determinant in this more abstract context.

The set of k-linear forms $X \times \cdots \times X \to K$ is a vector space of dimension n^k.

The following subclasses of k-linear forms are important subspaces:

- symmetric
- skew-symmetric
- alternating

k-linear forms

Definition

A *k-linear form* is a function $f : X_1 \times \cdots \times X_k \to K$ that is linear in each coordinate.

That is, if we fix $k-1$ inputs, it is linear in the remaining input.

Unless otherwise stated, we will assume that $X := X_1 = \cdots = X_k$.

1. 1-linear forms are linear functions in $X \to K$.
2. 2-linear forms are bilinear forms $X \times X \to K$.
3. A 3-linear form is a function $X \times X \times X \to K$.

The vector space of multilinear forms

Proposition

Let $\dim X = n$. The set of k-linear forms $X \times \cdots \times X \to K$ is a vector space of dimension n^k.

Symmetric and skew-symmetric multilinear forms

Let $f \colon X \times \cdots \times X \to K$ be a k-linear form.

For any permutation $\pi \in S_k$, define the k-linear form πf by

$$(\pi f)(x_1, \dots, x_k) = f\left(x_{\pi_1}, \dots, x_{\pi_k}\right).$$

Definition

A k-linear form is:

1. symmetric if $\pi f = f$ for every permutation $\pi \in S_k$
2. skew-symmetric if $\tau f = -f$ for every transposition $\tau \in S_k$.

Symmetric, skew-symmetric, and alternating forms

Recall that a k-linear form $f\colon X \times \cdots \times X \to K$ is:

- symmetric if $\pi f = f$ for all $\pi \in S_k$,
- skew-symmetric if $\tau f = -f$ for all transpositions $\tau \in S_k$.

Definition

A k-linear form is alternating if $f(x_1, \ldots, x_k) = 0$ whenever $x_i = x_j$ $(i \neq j)$.

It is easy to show that the set of alternating (respectively, symmetric or skew-symmetric) k-linear forms is a subspace of $\mathcal{T}^k(X')$.

Alternating vs. skew-symmetric

Proposition 3.1

Every alternating form is skew-symmetric.

Corollary 3.2

If $1 + 1 \neq 0$, then every skew-symmetric form is alternating.

Alternating forms and linear dependence

Proposition 3.3

If f is alternating and $y_1, \ldots, y_k$ is linearly dependent, then $f(y_1, \ldots, y_k) = 0$.

Alternating forms and linear independence

Proposition 3.4

If f is a nonzero alternating n-linear form and $e_1, \dots, e_n$ a basis, then $f(e_1, \dots, e_n) \neq 0$.

Corollary 3.5

Any two alternating n-linear forms are linearly dependent.

Symmetric, skew-symmetric, and alternating forms

Throughout, $\dim X = n < \infty$. Recall that a k-linear form $f\colon X \times \cdots \times X \to K$ is:

- symmetric if $\pi f = f$ for all $\pi \in S_k$
- skew-symmetric if $\tau f = -f$ for all transpositions $\tau \in S_k$
- alternating if $f(x_1, \dots, x_k) = 0$ whenever $x_i = x_j$ $(i \neq j)$.

All of these are subspaces of $\mathcal{T}^k(X')$, the space of k-linear forms. *What are their dimensions?*

Goal

Show that the subspace of alternating n-linear forms is 1-dimensional, by verifying

- any two alternating n-linear forms are linearly dependent (see previous lecture)
- there is a non-zero alternating n-linear form.

The determinant of $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is the unique alternating n-linear form satisfying $T(e_1, \dots, e_n) = 1$.

But we'd still like a definition that doesn't refer to the choice of basis. . .

The dimension of the subspace of alternating n-linear forms is ≥ 1

Proposition 3.5

There is a nonzero alternating n-linear form.

Determinants, at last

Let $T\colon X \to X$ be linear. For an alternating n-linear f, define a new alternating n-linear form

$$\bar{T}f\colon X^n \longrightarrow K, \qquad (\bar{T}f)(x_1, \dots, x_n) = f(Tx_1, \dots, Tx_n).$$

That is, T induces a map $\bar{T}$ on the (1-dimensional) space of alternating n-linear forms:

$$f \longmapsto \bar{T}f.$$

But any linear map on a 1-dimensional space is just scalar multiplication, $x \mapsto \lambda x$. Therefore,

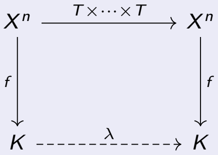
$$\bar{T}\colon f \longmapsto \lambda f.$$

The scalar λ is called the determinant of T. It satisfies the following.

Universal property of the determinant

Given a linear map $T\colon X \to X$, there exists a unique scalar λ such that for every alternating n-linear form f,

$$f(Tx_1, \dots, Tx_n) = \lambda f(x_1, \dots, x_n).$$



A few basic properties of determinants

If $Tx = cx$, then

$$(\bar{T}f)(x_1, \dots, x_n) = f(Tx_1, \dots, Tx_n) = f(cx_1, \dots, cx_n) = c^n f(x_1, \dots, x_n).$$

Thus, $\det T = c^n$.

It follows that $\det 0 = 0$ and $\det(\mathrm{Id}) = 1$.

Proposition 3.6

For any two linear maps $A, B \colon X \to X$,

$$\det(AB) = (\det A)(\det B).$$

Corollary 3.7

If $A \colon X \to X$ is invertible, then $\det A^{-1} = (\det A)^{-1} \neq 0$.

The determinant of a 2×2 matrix

The determinant of an $n \times n$ matrix can be thought of as an alternating n-linear function of its column vectors.

Let's use bilinearity to find a formula for the determinant of $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$.

The determinant of a 3×3 matrix

Let's now apply this to finding the determinant of $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$.

The determinant of an $n \times n$ matrix

Proposition 3.8

The determinant of an $n \times n$ matrix $A = (a_{ij})$ is

$$\det A = \sum_{\pi \in S_n} a_{1,\pi(1)} a_{2,\pi(2)} \cdots a_{n,\pi(n)},$$

and by symmetry, $\det A = \det A^T$.

The trace of a matrix

Definition

The trace of an $n \times n$ matrix is $\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}$.

Proposition 3.9

(a) Trace is linear: $\operatorname{tr}(kA) = k(\operatorname{tr} A)$ and $\operatorname{tr}(A + B) = \operatorname{tr} A + \operatorname{tr} B$.

(b) Trace is "commutative": $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

(c) Similar matrices have the same determinant and trace.

Minors and cofactors

Lemma 3.10

Let $A = [c_1, \dots, c_n]$ be an $n \times n$ matrix, and define B by adding kc_i to the j^{th} column, for $i \neq j$. Then $\det A = \det B$.

Definition

Let A be an $n \times n$ matrix, and let A_{ij} be the $(n-1) \times (n-1)$ matrix formed by removing the i^{th} row and j^{th} column.

- The (i,j) minor of A is $M_{ij} := \det A_{ij}$.
- The (i,j) cofactor of A is $C_{ij} := (-1)^{i+j} \det A_{ij}$.

Lemma 3.11

Let A be an $n \times n$ matrix with first column e_1, i.e., $A = \begin{bmatrix} 1 & - \\ 0 & A_{11} \end{bmatrix}$. Then $\det A = C_{11}$.

Corollary 3.12

Let A be a matrix whose j^{th} column is e_i. Then

$$\det A = C_{ij}.$$

Laplace expansion

Recall: If the j^{th} column of A is e_i, then $\det A = C_{ij}$.

Theorem (Laplace expansion)

The determinant of A is

$$\det A = \sum_{i=1}^{n} a_{ij} C_{ij},$$

for any fixed $j = 1, \dots, n$.

Systems of equations

Consider an invertible matrix, written as an n-tuple of its column vectors:

$$A = (a_1, \dots, a_n) = (Ae_1, \dots, Ae_n).$$

The system of equations $Ax = u$, with $x = \sum_{j=1}^{n} x_j e_j$ can be written

$$\sum_{j=1}^{n} x_j a_j = u.$$

For each k, define the matrix

$$A_k = (a_1, \dots, a_{k-1}, u, a_{k+1}, \dots, a_n),$$

and let's compute its determinant.

A formula for A^{-1}

Theorem (Cramer's rule)

The solution to the system of equations $Ax = u$, with $x = \sum_{j=1}^{n} x_j e_j$ is

$$x_k = \frac{1}{\det A} \sum_{i=1}^{n} C_{ik} u_i .$$

Theorem 3.13

If A is invertible, then the (i,j)-entry of its inverse A^{-1} is

$$(A^{-1})_{ij} = \frac{C_{ji}}{\det A} .$$

The idea behind tensor products

Consider two vector spaces U, V over K, and say $\dim U = n$ and $\dim V = m$. Then

$$U \cong \{a_{n-1}x^{n-1} + \cdots + a_1x + a_0 \mid a_i \in K\}, \qquad V \cong \{b_{m-1}y^{m-1} + \cdots + b_1x + b_0 \mid b_i \in K\}.$$

The direct product $U \times V$ has basis

$$\{(x^{n-1}, 0), \ldots, (x, 0), (1, 0)\} \cup \{(0, y^{m-1}), \ldots, (0, y), (0, 1)\}.$$

An arbitrary element has the form

$$(a_{n-1}x^{n-1} + \cdots + a_1x + a_0,\ b_{m-1}y^{m-1} + \cdots + b_1y + b_0) \in U \times V.$$

Notice that $(3x^i, y^j) \neq (x^i, 3y^j)$ in $U \times V$.

There is another way to "multiply" the vector spaces U and V together. It is easy to check that the following is a vector space:

$$\left\{ \sum_{j=0}^{m-1} \sum_{i=0}^{n-1} c_{ij} x^i y^j \mid c_{ij} \in K \right\}.$$

This is the idea of the tensor product, denoted $U \otimes V$.

Formalizing this is a bit delicate. For example, $3x^i \cdot y^j = x^i \cdot (3y^j) = 3(x^i \cdot y^j)$.

The tensor product in terms of bases

Though we are normally not allowed to "multiply" vectors, we can define it by inventing a special symbol.

Denote the formal "product" of two vectors $u \in U$ and $v \in V$ as $u \otimes v$.

Pick bases $u_1, \dots, u_n$ for U and $v_1, \dots, v_m$ for V.

Definition

The tensor product of U and V is the vector space with basis $\{u_i \otimes v_j\}$.

By definition, every element of $U \otimes V$ can be written uniquely as

$$\sum_{j=1}^{m} \sum_{i=1}^{n} c_{ij}(u_i \otimes v_j).$$

It is immediate that $\dim(U \otimes V) = (\dim U)(\dim V)$.

Remark

Not every multivariate polynomial in x and y factors as a product $p(x)q(y)$.

Not every element in $U \otimes V$ can be written as $u \otimes v$ – called a pure tensor.

A basis-free construction of the tensor product

Given vector spaces U and V, let $F_{U\times V}$ be the vector space with *basis* $U \times V$:

$$F_{U\times V} = \Big\{ \sum c_{uv} e_{u,v} \quad \mid \quad u \in U, \ \ v \in V \Big\}.$$

For all $u, u' \in U$ and $v, v' \in V$, we "need" the following to hold:

$$e_{u+u',v} = e_{u,v} + e_{u',v} \qquad e_{u,v+v'} = e_{u,v} + e_{u,v'} \qquad e_{cu,v} = ce_{u,v} \qquad e_{u,cv} = ce_{u,v}.$$

Consider the set of "*null sums*" from $F_{U\times V}$:

$$S = \Big[\bigcup_{\substack{u,u'\in U \\ v\in V}} e_{u+u',v} - e_{u,v} - e_{u',v} \Big] \cup \Big[\bigcup_{\substack{u\in U \\ v,v'\in V}} e_{u,v+v'} - e_{u,v} - e_{u,v'} \Big]$$
$$\cup \Big[\bigcup_{\substack{u\in U, v\in V \\ c\in K}} e_{cu,v} - ce_{u,v} \Big] \cup \Big[\bigcup_{\substack{u\in U, v\in V \\ c\in K}} e_{u,cv} - ce_{u,v} \Big].$$

Let $N_q = \text{Span}(S)$. Denote the equivalence class of $e_{u,v}$ mod N_q as $u \otimes v$.

Definition

The tensor product of U and V is the quotient space $U \otimes V := F_{U\times V}/N_q$.

Why this basis-free construction works

Let W be a vector space with basis $\{w_{ij} \mid 1 \leq i \leq n,\ 1 \leq j \leq m\}$. Define the linear map

$$\alpha\colon W \longrightarrow U \otimes V, \qquad \alpha\colon w_{ij} \longmapsto u_i \otimes v_j.$$

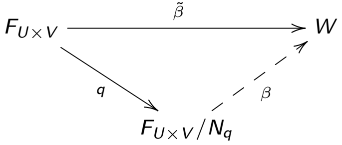
We'd like to define the (inverse) map $\beta\colon U \otimes V \to W$, but to do so, we need a basis for $U \otimes V$. What we *can* do is define a map

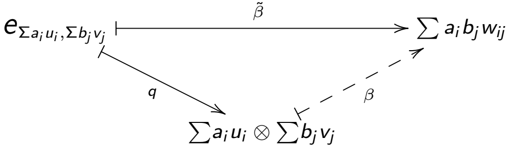
$$\tilde{\beta}\colon F_{U\times V} \longrightarrow W, \qquad \tilde{\beta}\colon e_{\Sigma a_i u_i, \Sigma b_j v_j} \longmapsto \sum_{i,j} a_i b_j w_{ij}.$$

Remark (exercise)

The nullspace of $\tilde{\beta}$ contains the nullspace of q.

Since $N_q \subseteq N_{\tilde{\beta}}$, the map $\tilde{\beta}$ *factors through* $F_{U\times V}/N_q := U \otimes V$:





The maps α and β are inverses because $\alpha \circ \beta = \mathrm{Id}_{U\otimes V}$ and $\beta \circ \alpha = \mathrm{Id}_W$.

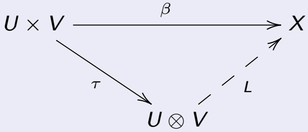
Universal property of the tensor product

Let $\tau\colon U \times V \to U \otimes V$ be the map $(u, v) \mapsto u \otimes v$.

The following says that every bilinear map from $U \times V$ can be "factored through" $U \otimes V$.

Theorem 3.14

For every bilinear $\beta\colon U \times V \to X$, there is a unique linear $L\colon U \otimes V \to X$ such that $\beta = L \circ \tau$.



The universal property can provide us with alternate proofs of some basic results, such as:

(i) $\{u_i \otimes v_j\}$ is linearly independent

(ii) $U \otimes V \cong V \otimes U$

(iii) $(U \otimes V) \otimes W \cong U \otimes (V \otimes W)$

(iv) $(U \times V) \otimes W \cong (U \otimes W) \times (V \otimes W)$.

Tensors as linear maps

Proposition 3.15

There is a natural isomorphism

$$U \otimes V \longrightarrow \mathrm{Hom}(U', V), \qquad u \otimes v \longmapsto \Big(\ell \mapsto (\ell, u)v\Big).$$

The following shows the linear map $\ell \overset{E_{ij}}{\longmapsto} (\ell, u_i)v_j$ in matrix form:

$$\underbrace{\begin{bmatrix} c_1 & \cdots & c_i & \cdots & c_n \end{bmatrix}}_{\ell = \sum c_i \ell_i \in U'} \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & 1 & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{E_{ij} := v_j^T u_i} = \underbrace{\begin{bmatrix} 0 & \cdots & c_i & \cdots & 0 \end{bmatrix}}_{c_i v_j \in V}$$

More generally:

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \otimes \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} = vu^T = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} = \begin{bmatrix} v_1 u_1 & v_1 u_2 & \cdots & v_1 u_n \\ v_2 u_1 & v_2 u_2 & \cdots & v_2 u_n \\ \vdots & \vdots & \ddots & \vdots \\ v_m u_1 & v_m u_2 & \cdots & v_m u_n \end{bmatrix}$$

Tensors as a way to extend an $\mathbb{R}$-vector space to a $\mathbb{C}$-vector space

Let X be an $\mathbb{R}$-vector space with basis $\{x_1, \dots, x_n\}$.

Note that $\mathbb{C}$ is a 2-dimensional $\mathbb{R}$-vector space, with basis $\{1, i\}$.

Suppose $A\colon X \to X$ is a linear map with eigenvalues $\lambda_{1,2} = \pm i$.

If v is an eigenvector v for $\lambda = i$, then $v \notin X$. But v should live in some "extension" of X.

In this bigger vector space, we want to have vectors like

$$zv, \qquad z \in \mathbb{C}, \quad v \in X.$$

What we really want is $\mathbb{C} \otimes X$, which has basis

$$\big\{1 \otimes x_1, \dots, 1 \otimes x_n, i \otimes x_1, \dots, i \otimes x_n\big\} \text{ "="} \big\{x_1, \dots, x_n, ix_1, \dots, ix_n\big\}.$$

Notice how the associativity that we would expect comes for free with the tensor product, and compare it to the other examples from this lecture:

$$(3i)v = i(3v), \qquad (3x^i)y^j = x^i(3y^j), \qquad (3u)v^T = u(3v^T), \qquad 3u \otimes v = u \otimes 3v.$$