

Multilinearity and the Determinant

Jason Kenyon

November 8, 2022

1 Multilinearity

Here, we define the determinant as the unique multilinear function that is alternating, and one when given the identity matrix. We then deduce some familiar properties thereof. Finally, we investigate other familiar multilinear functions, as well as introduce sesquilinearity—(multi)linear(ish) functions.

Henceforth, arguments of functions will be assumed to be vectors, not scalars, of an arbitrary space over an arbitrary field, unless we state otherwise. To begin, we define the determinant of an $n \times n$ matrix A .

Definition 1. $\det(A) = \sum_{i=1}^n (-1)^i A_{ij} \det(A_{*ij})$ where A_{*ij} is the matrix found by eliminating the i^{th} row and j^{th} column of A .

The above is the standard computational definition of the determinant, from which its main properties may be derived. We shall have recourse to these properties without proving them. In particular, multiplicativity shall be used extensively. The reader is assumed to be familiar with the basics of the determinant.

Definition 2. An n -linear function $f : X \rightarrow Y$ is a function such that for all $1 \leq i \leq n$

$$f(x_1, \dots, ax + y, \dots, x_n) = af(x_1, \dots, x, \dots, x_n) + f(x_1, \dots, y, \dots, x_n)$$

where $ax + y$ is the i^{th} argument of f .

Now that we have our primary object of concern, let's establish the properties our function must have:

Definition 3. An n linear function $f : X \rightarrow Y$ is alternating if

$$f(x_1, \dots, x_n) = 0$$

whenever $x_i = x_{i+1}$ for $1 \leq i \leq n$

Furthermore, suppose that

$$f(x_1, \dots, x_n) = 1 \text{ whenever } \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = I$$

Theorem 1. Suppose f satisfies the above properties. Let $A = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and

$$B = \begin{pmatrix} x_1 \\ \vdots \\ x_{i+1} \\ x_i \\ \vdots \\ x_n \end{pmatrix} \text{ for } 1 \leq i \leq n. \text{ Then } f(B) = -f(A).$$

Proof. By the alternating condition and multilinearity of f , we have

$$0 = f \begin{pmatrix} x \\ \vdots \\ x+y \\ x+y \\ \vdots \\ x_n \end{pmatrix} \quad (1)$$

$$= f \begin{pmatrix} x_1 \\ \vdots \\ x \\ x \\ x_n \end{pmatrix} + f \begin{pmatrix} x_1 \\ \vdots \\ y \\ y \\ x_n \end{pmatrix} + f(B) + f(A) \quad (2)$$

$$= f(B) + f(A) \quad (3)$$

□

Theorem 2. Let $A = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ with $x_i = x_j$ for $1 \leq i, j \leq n$ and $i \neq j$. Then

$$f(A) = 0$$

Proof. Suppose that $x_i = x_j$ and $i \neq j$. By the previous theorem,

$$-f \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ x_{i+1} \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} = f \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ x_j \\ \vdots \\ x_{i+1} \\ \vdots \\ x_n \end{pmatrix} = 0 \quad \square$$

Theorem 3. Let $A = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $B = \begin{pmatrix} x_1 \\ \vdots \\ x_i + ax_j \\ \vdots \\ x_n \end{pmatrix}$ for $1 \leq j \leq n$.

Proof. By the previous theorem and multilinearity of f :

$$f(B) = f(A) + af \begin{pmatrix} x_1 \\ \vdots \\ x_j \\ \vdots \\ x_j \\ \vdots \\ x_n \end{pmatrix} \quad (4)$$

$$= f(A) \quad (5)$$

□

Theorem 4. Let $A = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. If $\text{rank}(A) < n$, then $f(A) = 0$.

Proof. Since $\text{rank}(A) < n$ we may deduce $f(A) = f \begin{pmatrix} e_1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ with zero vectors

beginning at the $\text{rank}(A)^{\text{th}}$ row. Therefore, we may rewrite

$$f(A) = \begin{pmatrix} e_1 \\ \vdots \\ e_{\text{rank}(A)+i} - e_{\text{rank}(A)+i} \\ \vdots \\ e_n - e_n \end{pmatrix}$$

for $0 \leq i \leq n - \text{rank}(A)$. By multilinearity, it is clear that we may decompose

$$\text{this into permutations } f(A) = \sum_{i=1}^{n-\text{rank}(A)} f \begin{pmatrix} e_1 \\ \vdots \\ e_{\text{rank}(A)+i} \\ \vdots \\ e_n \end{pmatrix} - f \begin{pmatrix} e_1 \\ \vdots \\ e_{\text{rank}(A)+i} \\ \vdots \\ e_n \end{pmatrix} = 0$$

□

Theorem 5. For elementary matrices E_1, E_2 , and E_3 of type 1, 2, and 3 respectively $f(E_1) = -1$, $f(E_2) = k$, and $f(E_3) = 1$.

Each of the above follow directly from the multilinearity of f and the assumption that $f(I) = 1$. Hence, we shall omit the proof.

Theorem 6. $f(AB) = \det(A)f(B)$.

Proof. We start by proving the theorem for elementary matrices then extend that to arbitrary ones. By the previous theorem, $f(E_1) = -1$ and $f(E_1A) = -f(A)$ by theorem 1. Thus, $f(E_1A) = \det(E_1)f(A)$ since E_1 is the result of swapping a row in the identity matrix. Next, $f(E_2) = k$ by the previous theorem and $\det(E_2) = k\det(I) = k$. Since $f(E_2A) = kf(A)$, we have $f(E_2A) = \det(E_2)f(A)$. Finally, by the previous theorem, we have $f(E_3) = 1$ and $\det(E_3) = \det(I) + k\det(E)$ for some matrix E with two equivalent rows. Therefore, $f(E_3A) = f(A) = \det(E_3)f(A)$.

We know that $\text{rank}(AB) \leq \text{rank}(A)$ by definition of matrix multiplication, and $f(A) = 0$ if $\text{rank}(A) < n$ by theorem 4. Thus if $\text{rank}(A) < n$, $f(AB) = 0$ and similarly, $\det(A) = 0$. So $f(AB) = \det(A)f(B) = 0$. If A is full rank, then we may represent $A = E_1E_2 \dots E_n$ as full rank implies that A may be completely row reduced. This means that

$$f(AB) = f(E_1 \dots E_n B) \quad (6)$$

$$= \det(E_1)f(E_2 \dots E_n B) \quad (7)$$

$$= \dots \quad (8)$$

$$= \det(E_1) \dots \det(E_n)f(B) \quad (9)$$

$$= \det(E_1 \dots E_n)f(B) \quad (10)$$

$$= \det(A)f(B) \quad (11)$$

□

Theorem 7. $f(A) = \det(A)$

Proof. If $\text{rank}(A) < n$ then $f(A) = 0$ by theorem 4. Similarly, $\det(A) = 0$. Otherwise, just as before we may represent $A = E_1 \dots E_n$. Hence

$$f(A) = f(E_1 \dots E_n) \quad (12)$$

$$= \det(E_1)f(E_2 \dots E_n) \quad (13)$$

$$= \dots \quad (14)$$

$$= \det(E_1) \dots \det(E_n) \quad (15)$$

$$= \det(E_1 \dots E_n) \quad (16)$$

$$= \det(A) \quad (17)$$

□

Therefore we have shown that the determinant is nothing but the alternating, multilinear function that evaluates to 1 given the identity.

2 Inner Products and Definiteness

Inner products are of the the primary tools used in analysis and geometry. Their properties correspond directly to our idea of space, and are therefore of utmost import.

Definition 4. An inner product is a function $\langle \cdot, \cdot \rangle : V \times V \rightarrow F$ for a given vector space V over a field F that satisfies the following:

1. $\langle v, w \rangle = \overline{\langle w, v \rangle}$
2. $\langle rv + q, w \rangle = r\langle v, w \rangle + \langle q, w \rangle$
3. $\langle v, v \rangle > 0$

for all $v, w, q \in V$ and $r \in F$, with $v \neq \mathbf{0}$.

Note that F is either \mathbb{R} or \mathbb{C} . Also, if $F = \mathbb{R}$ then linearity in the first argument, coupled with symmetry, implies linearity in the second. The standard inner product on \mathbb{C}^n is defined as follows:

Definition 5. $v \cdot w = \sum_{i=1}^n \bar{v}_i w_i$

As the conjugate of a real number is itself, this definition reduces to the familiar dot product on \mathbb{R}^n . Additionally, it should be clear that this definition satisfies the criteria defined previously for inner products. Now we define what it means for a matrix to be positive definite:

Definition 6. An $n \times n$ matrix A with entries in either \mathbb{R} or \mathbb{C} is called positive definite if $v^* A v > 0$ for all nonzero $v \in F^n$.

Now we show that any positive definite $n \times n$ self-adjoint matrix defines an inner product.

Theorem 8. Let F be either the real or complex field and A an $n \times n$ matrix with entries thereof. If $A^* = A$, then $\langle x, y \rangle := x^* A y$ is an inner product.

Proof. We begin by demonstrating conjugate symmetry:

$$\langle x, y \rangle = x^* A y \tag{18}$$

$$= x \cdot A y \tag{19}$$

$$= \overline{A y \cdot x} \tag{20}$$

$$= \overline{(A y)^* x} \tag{21}$$

$$= \overline{y^* A^* x} \tag{22}$$

$$= \overline{y^* A x} \tag{23}$$

$$= \overline{\langle y, x \rangle} \tag{24}$$

Next, linearity in the first argument:

$$\langle ax + b, y \rangle = (ax + b)^* Ay \quad (25)$$

$$= ((ax)^* + b^*) Ay \quad (26)$$

$$= (ax)^* Ay + b^* Ay \quad (27)$$

$$= \bar{a}x^* Ay + \langle b, y \rangle \quad (28)$$

$$= \bar{a}x \cdot Ay + \langle b, y \rangle \quad (29)$$

$$= (\bar{a}x)^* Ay + \langle b, y \rangle \quad (30)$$

$$= x^* a Ay + \langle b, y \rangle \quad (31)$$

$$= ax^* Ay + \langle b, y \rangle \quad (32)$$

$$= a\langle x, y \rangle + \langle b, y \rangle \quad (33)$$

And finally, $\langle x, x \rangle = x^* Ax > 0$ when x is nonzero, by assumption. Furthermore, $\mathbf{0}A\mathbf{0} = 0$. Thus $\langle x, x \rangle = 0$ if and only if $x = \mathbf{0}$. \square

Additionally, it should be easy to see that conjugate symmetry and linearity in the first argument imply that $\langle x, ay + b \rangle = \bar{a}\langle x, y \rangle + \langle x, b \rangle$ when $F \neq \mathbb{R}$. That is, the inner product is sesquilinear in general, but reduces to a bilinear function in the case of \mathbb{R} .

With this result now proven, we see that the standard inner product on \mathbb{C}^n is the inner product defined by the identity matrix.