26. Determinants I

26.1 Prehistory
26.2 Definitions
26.3 Uniqueness and other properties
26.4 Existence

Both as a careful review of a more pedestrian viewpoint, and as a transition to a coordinate-independent approach, we roughly follow Emil Artin's rigorization of determinants *of matrices* with entries in a *field*. Standard properties are derived, in particular *uniqueness*, from simple assumptions. We also prove *existence*. Soon, however, we will want to develop corresponding intrinsic versions of ideas about *endomorphisms*. This is *multilinear algebra*. Further, for example to treat the Cayley-Hamilton theorem in a forthright manner, we will want to consider modules over commutative rings, not merely vector spaces over fields.

---

1. *Prehistory*

Determinants arose many years ago in *formulas* for solving linear equations. This is **Cramer's Rule**, described as follows. [1] Consider a system of n linear equations in n unknowns $x_1, \ldots, x_n$

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & c_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & c_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & \ldots & + & a_{3n}x_n & = & c_3 \\ \vdots & & & & \ddots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \ldots & + & a_{nn}x_n & = & c_n \end{array}$$

[1] We will prove Cramer's Rule just a little later. In fact, quite contrary to a naive intuition, the proof is very easy from an only slightly more sophisticated viewpoint.

Let A be the matrix with $(i,j)^{th}$ entry a_{ij}. Let $A^{(\ell)}$ be the matrix A with its ℓ^{th} column replaced by the c_is, that is, the $(i,\ell)^{th}$ entry of $A^{(\ell)}$ is c_ℓ. Then Cramer's Rule asserts that

$$x_\ell = \frac{\det A^{(\ell)}}{\det A}$$

where det is determinant, at least for $\det A \neq 0$. It is implicit that the *coefficients* a_{ij} and the *constants* c_ℓ are in a *field*. As a practical method for solving linear systems Cramer's Rule is far from optimal. Gaussian elimination is much more efficient, but is less *interesting*.

Ironically, in the context of very elementary mathematics it seems difficult to give an intelligible definition or formula for determinants of arbitrary sizes, so typical discussions are limited to very small matrices. For example, in the 2-by-2 case there is the palatable formula

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

Thus, for the linear system

$$\begin{aligned} ax + by &= c_1 \\ cx + dy &= c_2 \end{aligned}$$

by Cramer's Rule

$$x = \frac{\det \begin{pmatrix} c_1 & b \\ c_2 & d \end{pmatrix}}{\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}} \qquad y = \frac{\det \begin{pmatrix} a & c_1 \\ c & c_2 \end{pmatrix}}{\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}}$$

In the 3-by-3 case there is the still-barely-tractable formula (reachable by a variety of elementary mnemonics)

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$= \left(a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}\right) - \left(a_{31}a_{22}a_{13} + a_{32}a_{23}a_{11} + a_{31}a_{21}a_{12}\right)$$

Larger determinants are defined ambiguously by induction as *expansions by minors.* [2]

Inverses of matrices are expressible, inefficiently, in terms of determinants. The **cofactor matrix** or **adjugate matrix** A^{adjg} of an n-by-n matrix A has $(i,j)^{th}$ entry

$$A^{\mathrm{adjg}}_{ij} = (-1)^{i+j} \det A^{(ji)}$$

where A^{ji} is the matrix A with j^{th} row and i^{th} column removed. [3] Then

$$A \cdot A^{\mathrm{adjg}} = (\det A) \cdot 1_n$$

where 1_n is the n-by-n identity matrix. That is, if A is invertible,

$$A^{-1} = \frac{1}{\det A} \cdot A^{\mathrm{adjg}}$$

[2] We describe expansion by minors just a little later, and prove that it is in fact unambiguous and correct.

[3] Yes, there is a reversal of indices: the $(ij)^{th}$ entry of A^{adjg} is, up to sign, the determinant of A with j^{th} row and i^{th} column removed. Later discussion of *exterior algebra* will clarify this construction/formula.

In the 2-by-2 case this formula is useful:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Similarly, a matrix (with entries in a field) is invertible if and only if its determinant is non-zero. [4]

The **Cayley-Hamilton theorem** is a widely misunderstood result, often given with seriously flawed proofs. [5] The **characteristic polynomial** $P_T(x)$ of an n-by-n matrix T is defined to be

$$P_T(x) = \det(x \cdot 1_n - T)$$

The assertion is that

$$P_T(T) = 0_n$$

where 0_n is the n-by-n zero matrix. The main use of this is that the **eigenvalues** of T are the roots of $P_T(x) = 0$. However, except for very small matrices, this is a suboptimal computational approach, *and* the **minimal polynomial** is far more useful for demonstrating qualitative facts about endomorphisms. Nevertheless, because there is a *formula* for the characteristic polynomial, it has a substantial popularity.

The easiest false proof of the Cayley-Hamilton Theorem is to apparently compute

$$P_T(T) = \det(T \cdot 1_n - T) = \det(T - T) = \det(0_n) = 0$$

The problem is that the substitution $x \cdot 1_n \longrightarrow T \cdot 1_n$ is not legitimate. The operation cannot be any kind of *scalar multiplication* after T is substituted for x, nor can it be composition of endomorphisms (nor multiplication of matrices). Further, there are interesting fallacious explanations of this incorrectness. For example, to say that we cannot substitute the *non-scalar* T for the *scalar variable* x fails to recognize that this is exactly what happens in the assertion of the theorem, *and* fails to see that the real problem is in the notion of the scalar multiplication of 1_n by x. That is, the *correct* objection is that $x \cdot 1_n$ is no longer a matrix with entries in the original field k (whatever that was), but in the polynomial ring $k[x]$, or in its field of fractions $k(x)$. But then it is much less clear what it might mean to substitute T for x, if x has become a kind of scalar.

Indeed, Cayley and Hamilton only proved the result in the 2-by-2 and 3-by-3 cases, by direct computation.

Often a correct argument is given that invokes the (existence part of the) structure theorem for finitely-generated modules over PIDs. A little later, our discussion of *exterior algebra* will allow a more direct argument, using the adjugate matrix. More importantly, the exterior algebra will make possible the long-postponed *uniqueness* part of the proof of the structure theorem for finitely-generated modules over PIDs.

[4] We prove this later in a much broader context.

[5] We give two different correct proofs later.

2. *Definitions*

For the present discussion, a **determinant** is a function D of *square matrices* with entries in a field k, taking values in that field, satisfying the following properties.

• **Linearity** as a function of each column: letting $C_1, \ldots, C_n$ in k^n be the columns of an n-by-n matrix C, for each $1 \le i \le n$ the function

$$C_i \longrightarrow D(C)$$

is a k-linear map $k^n \longrightarrow k$. [6] That is, for scalar b and for two columns C_i and C_i'

$$D(\ldots, bC_i, \ldots) = b \cdot D(\ldots, C_i, \ldots)$$

$$D(\ldots, C_i + C_i', \ldots) = D(\ldots, C_i, \ldots) + D(\ldots, C_i', \ldots)$$

• **Alternating property:** [7] If two adjacent columns of a matrix are equal, the determinant is 0.

• **Normalization:** The determinant of an identity matrix is 1:

$$D \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix} = 1$$

That is, as a function of the columns, if the columns are the standard basis vectors in k^n then the value of the determinant is 1.

3. *Uniqueness and other properties*

• If two columns of a matrix are interchanged the value of the determinant is multiplied by -1. That is, writing the determinant as a function of the columns

$$D(C) = D(C_1, \ldots, C_n)$$

we have

$$D(C_1, \ldots, C_{i-1}, C_i, C_{i+1}, C_{i+2}, \ldots, C_n) = -D(C_1, \ldots, C_{i-1}, C_{i+1}, C_i, C_{i+2}, \ldots, C_n)$$

Proof: There is a little trick here. Consider the matrix with $C_i + C_j$ at both the i^{th} and j^{th} columns. Using the linearity in both i^{th} and j^{th} columns, we have

$$0 = D(\ldots, C_i + C_j, \ldots, C_i + C_j, \ldots)$$

[6] Linearity as a function of several vector arguments is called **multilinearity**.

[7] The etymology of *alternating* is somewhat obscure, but does have a broader related usage, referring to rings that are *anti-commutative*, that is, in which $x \cdot y = -y \cdot x$. We will see how this is related to the present situation when we talk about *exterior algebras*. Another important family of alternating rings is *Lie algebras*, named after Sophus Lie, but in these the product is written $[x, y]$ rather than $x \cdot y$, both by convention and for functional reasons.

$$= D(\ldots, C_i, \ldots, C_i, \ldots) + D(\ldots, C_i, \ldots, C_j, \ldots)$$
$$+ D(\ldots, C_j, \ldots, C_i, \ldots) + D(\ldots, C_j, \ldots, C_j, \ldots)$$

The first and last determinants on the right are also 0, since the matrices have two identical columns. Thus,

$$0 = D(\ldots, C_i, \ldots, C_j, \ldots) + D(\ldots, C_j, \ldots, C_i, \ldots)$$

as claimed. ///

[3.0.1] **Remark:** If the characteristic of the underlying field k is not 2, then we can replace the requirement that equality of two columns forces a determinant to be 0 by the requirement that interchange of two columns multiplies the determinant by -1. But this latter is a strictly weaker condition when the characteristic is 2.

• For any permutation π of $\{1, 2, 3, \ldots, n\}$ we have

$$D(C_{\pi(1)}, \ldots, C_{\pi(n)}) = \sigma(\pi) \cdot D(C_1, \ldots, C_n)$$

where C_i are the columns of a square matrix and σ is the *sign* function on S_n.

Proof: This argument is completely natural. The *adjacent transpositions* generate the permutation group S_n, and the sign function $\sigma(\pi)$ evaluated on a permutation π is $(-1)^t$ where t is the number of adjacent transpositions used to express π in terms of adjacent permutations. ///

• The value of a determinant is *unchanged* if a multiple of one column is added to another. That is, for indices $i < j$, with columns C_i considered as vectors in k^n, and for $b \in k$,

$$D(\ldots, C_i, \ldots, C_j, \ldots) = D(\ldots, C_i, \ldots, C_j + bC_i, \ldots)$$
$$D(\ldots, C_i, \ldots, C_j, \ldots) = D(\ldots, C_i + bC_j, \ldots, C_j, \ldots)$$

Proof: Using the linearity in the j^{th} column,

$$D(\ldots, C_i, \ldots, C_j + bC_i, \ldots) = D(\ldots, C_i, \ldots, C_j, \ldots) + b \cdot D(\ldots, C_i, \ldots, C_i, \ldots)$$
$$= D(\ldots, C_i, \ldots, C_j, \ldots) + b \cdot 0 = D(\ldots, C_i, \ldots, C_j, \ldots)$$

since a determinant is 0 if two columns are equal. ///

• Let

$$C_j = \sum_i b_{ij} A_i$$

where b_{ij} are in k and $A_i \in k^n$. Let C be the matrix with i^{th} column C_i, and let A the the matrix with i^{th} column A_i. Then

$$D(C) = \left(\sum_{\pi \in S_n} \sigma(\pi)\, b_{\pi(1),1} \ldots, b_{\pi(n),n} \right) \cdot D(A)$$

and also

$$D(C) = \left(\sum_{\pi \in S_n} \sigma(\pi)\, b_{1,\pi(1),1} \ldots, b_{n,\pi(n)} \right) \cdot D(A)$$

Proof: First, expanding using (multi-) linearity, we have

$$D(\ldots, C_j, \ldots) = D(\ldots, \textstyle\sum_i b_{ij} A_i, \ldots) = \displaystyle\sum_{i_1,\ldots,i_n} b_{i_1,1} \ldots b_{i_n,n}\, D(A_{i_1}, \ldots, A_{i_n})$$

where the ordered n-tuple $i_1, \ldots, i_n$ is summed over all choices of ordered n-tupes with entries from $\{1, \ldots, n\}$. If any two of i_p and i_q with $p \neq q$ are equal, then the matrix formed from the A_i will have two identical columns, and will be 0. Thus, we may as well sum over *permutations* of the ordered n-tuple $1, 2, 3, \ldots, n$. Letting π be the permutation which takes ℓ to i_ℓ, we have

$$D(A_{i_1}, \ldots, A_{i_n}) = \sigma(\pi) \cdot D(A_1, \ldots, A_n)$$

Thus,

$$D(C) = D(\ldots, C_j, \ldots) = \left(\sum_{\pi \in S_n} \sigma(\pi)\, b_{\pi(1),1} \ldots, b_{n,\pi(n)} \right) \cdot D(A)$$

as claimed. For the second, complementary, formula, since multiplication in k is commutative,

$$b_{\pi(1),1} \ldots b_{\pi(n),n} = b_{1,\pi^{-1}(1)} \ldots b_{n,\pi^{-1}(n)}$$

Also,

$$1 = \sigma(1) = \sigma(\pi \circ \pi^{-1}) = \sigma(\pi) \cdot \sigma(\pi^{-1})$$

And the map $\pi \longrightarrow \pi^{-1}$ is a bijecton of S_n to itself, so

$$\sum_{\pi \in S_n} \sigma(\pi)\, b_{\pi(1),1} \ldots, b_{n,\pi(n)} = \sum_{\pi \in S_n} \sigma(\pi)\, b_{1,\pi(1)} \ldots, b_{\pi(n),n}$$

which yields the second formula. ///

[3.0.2] **Remark:** So far we have *not* used the normalization that the determinant of the identity matrix is 1. Now we will use this.

• Let c_{ij} be the $(i,j)^{th}$ entry of an n-by-n matrix C. Then

$$D(C) = \sum_{\pi \in S_n} \sigma(\pi)\, c_{\pi(1),1} \ldots, c_{n,\pi(n)}$$

Proof: In the previous result, take A to be the identity matrix. ///

• (**Uniqueness**) There is at most one one determinant function on n-by-n matrices.

Proof: The previous formula is valid once we prove that determinants exist. ///

• The transpose $C^\top$ of C has the same determinant as does C

$$D(C^\top) = D(C)$$

Proof: Let c_{ij} be the $(i,j)^{th}$ entry of C. The $(i,j)^{th}$ entry $c_{ij}^\top$ of $C^\top$ is c_{ji}, and we have shown that

$$D(C^\top) = \sum_{\pi \in S_n} \sigma(\pi)\, c^\top_{\pi(1),1} \ldots c^\top_{\pi(n),n}$$

Thus,

$$D(C^\top) = \sum_{\pi \in S_n} \sigma(\pi)\, c_{\pi(1),1} \ldots c_{n,\pi(n)}$$

which is also $D(C)$, as just shown. ///

• (**Multiplicativity**) For two square matrices A, B with entries a_{ij} and b_{ij} and product $C = AB$ with entries c_{ij}, we have

$$D(AB) = D(A) \cdot D(B)$$

Proof: The j^{th} column C_j of the product C is the linear combination

$$A_1 \cdot b_{1,j} + \ldots + A_n \cdot b_{n,j}$$

of the columns $A_1, \ldots, A_n$ of A. Thus, from above,

$$D(AB) = D(C) = \left(\sum_\pi \sigma(\pi)\, b_{\pi(1),1} \ldots b_{\pi(n),1} \right) \cdot D(A)$$

And we know that the sum is $D(B)$. ///

• If two rows of a matrix are identical, then its determinant is 0.

Proof: Taking transpose leaves the determinant alone, and a matrix with two identical columns has determinant 0. ///

• **Cramer's Rule** Let A be an n-by-n matrix with j^{th} column A_j. Let b be a column vector with i^{th} entry b_i. Let x be a column vector with i^{th} entry x_i. Let $A^{(\ell)}$ be the matrix obtained from A by replacing the j^{th} column A_j by b. Then a solution x to an equation

$$Ax = b$$

is given by

$$x_\ell = \frac{D(A^{(\ell)})}{D(A)}$$

if $D(A) \neq 0$.

Proof: This follows directly from the alternating multilinear nature of determinants. First, the equation $Ax = b$ can be rewritten as an expression of b as a linear combination of the columns of A, namely

$$b = x_1 A_1 + x_2 A_2 + \ldots + x_n A_n$$

Then

$$D(A^{(\ell)}) = D(\ldots, A_{\ell-1}, \sum_j x_j A_j, A_{\ell+1}, \ldots) = \sum_j x_j \cdot D(\ldots, A_{\ell-1}, A_j, A_{\ell+1}, \ldots)$$

$$= x_\ell \cdot D(\ldots, A_{\ell-1}, A_\ell, A_{\ell+1}, \ldots) = x_\ell \cdot D(A)$$

since the determinant is 0 whenever two columns are identical, that is, unless $\ell = j$. ///

[3.0.3] **Remark:** In fact, this proof of Cramer's Rule does a little more than verify the formula. First, even if $D(A) = 0$, still

$$D(A^{(\ell)}) = x_\ell \cdot D(A)$$

Second, for $D(A) \neq 0$, the computation actually shows that the solution x is *unique* (since any solutions x_ℓs satisfy the indicated relation).

• An n-by-n matrix is invertible if and only if its determinant is non-zero.

Proof: If A has an inverse A^{-1}, then from $A \cdot A^{-1} = 1_n$ and the multiplicativity of determinants,

$$D(A) \cdot D(A^{-1}) = D(1_n) = 1$$

so $D(A) \neq 0$. On the other hand, suppose $D(A) \neq 0$. Let e_i be the i^{th} standard basis element of k^n, as a column vector. For each $j = 1, \ldots, n$ Cramer's Rule gives us a solution b_j to the equation

$$A\, b_j = e_j$$

Let B be the matrix whose j^{th} column is b_j. Then

$$AB = 1_n$$

To prove that also $BA = 1_n$ we proceed a little indirectly. Let T_M be the endomorphism of k^n given by a matrix M. Then

$$T_A \circ T_B = T_{AB} = T_{1_n} = \mathrm{id}_{k^n}$$

Thus, T_A is surjective. Since

$$\dim \mathrm{Im}\, T_A + \dim \ker T_A = n$$

necessarily T_A is also injective, so is an isomorphism of k^n. In particular, a right inverse is a left inverse, so also

$$T_{BA} = T_B \circ T_A = \mathrm{id}_{k^n}$$

The only matrix that gives the identity map on k^n is 1_n, so $BA = 1_n$. Thus, A is invertible. ///

[3.0.4] **Remark:** All the above discussion assumes *existence* of determinants.

---

4. *Existence*

The standard *ad hoc* argument for existence is ugly, and we won't write it out. If one must a way to proceed is to check directly by induction on size that an *expansion by minors* along any row or column meets the requirements for a determinant function. Then invoke uniqueness.

This argument might be considered acceptable, but, in fact, it is much less illuminating than the use above of the key idea of *multilinearity* to prove properties of determinants before we're sure they exist. With hindsight, the capacity to talk about a determinant function $D(A)$ which is *linear* as a function of each column (and is alternating) is very effective in proving properties of determinants.

That is, without the notion of *linearity* a derivation of properties of determinants is much clumsier. This is why high-school treatments (and 200-year-old treatments) are awkward.

By contrast, we need a more sophisticated viewpoint than basic linear algebra in order to give a conceptual reason for the *existence* of determinants. Rather than muddle through expansion by minors, we will wait until we have developed the exterior algebra that makes this straightforward.

Exercises

26.[4.0.1] Prove the *expansion by minors* formula for determinants, namely, for an n-by-n matrix A with entries a_{ij}, letting A^{ij} be the matrix obtained by deleting the i^{th} row and j^{th} column, for any fixed row index i,

$$\det A = (-1)^i \sum_{j=1}^{n} (-1)^j \, a_{ij} \, \det A^{ij}$$

and symmetrically for expansion along a column. (*Hint:* Prove that this formula is linear in each row/column, and invoke the uniqueness of determinants.)

26.[4.0.2] From just the most basic properties of determinants of matrices, show that the determinant of an upper-triangular matrix is the product of its diagonal entries. That is, show that

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ 0 & a_{22} & a_{23} & \ldots & a_{2n} \\ 0 & 0 & a_{33} & & \\ \vdots & & & \ddots & \vdots \\ 0 & \ldots & & 0 & a_{nn} \end{pmatrix} = a_{11}a_{22}a_{33}\ldots a_{nn}$$

26.[4.0.3] Show that determinants *respect block decompositions*, at least to the extent that

$$\det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \;=\; \det A \cdot \det D$$

where A is an m-by-n matrix, B is m-by-n, and D is n-by-n.

26.[4.0.4] By an example, show that it is *not* always the case that

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} \;=\; \det A \cdot \det D - \det B \cdot \det C$$

for blocks A, B, C, D.

26.[4.0.5] Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ be two orthonormal bases in a real inner-product space. Let M be the matrix whose ij^{th} entry is

$$M_{ij} \;=\; \langle x_i, y_j \rangle$$

Show that $\det M = 1$.

26.[4.0.6] For real numbers a, b, c, d, prove that

$$\left| \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right| \;=\; (\text{area of parallelogram spanned by } (a,b) \text{ and } (c,d))$$

26.[4.0.7] For real vectors $v_i = (x_i, y_i, z_i)$ with $i = 1, 2, 3$, show that

$$\left| \det \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} \right| \;=\; (\text{volume of parallelogram spanned by } v_1, v_2, v_3)$$