What is determinant?

Indian Institute of Science Education & Research Kolkata

IISER KOLKATA

Definitions of determinants

Laplace : Cofactor expansion along i^{th} row

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\tilde{A}_{ij}).$$

- $\tilde{A}_{ij}$ is the $(n-1) \times (n-1)$ minor associated to a_{ij}.
- the expression is independent of the choice of row or column.

Leibniz : Using permutations

$$\det(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

- S_n denotes the group of permutations of $\{1, 2, \ldots, n\}$.
- Every permutation σ has a sign, taking values in $\{\pm 1\}$.

Remarks: Historical & computational

- Nomenclature

  1801 Gauss: Used the term in *Disquisitiones arithmeticae* but it is different from the modern notion.

  1812 Cauchy: Introduced the nomenclature and proved multiplicativity, i.e., $\det(AB) = \det(A)\det(B)$.

  1841 Cayley: Introduced the notation $|A|$ for determinants.

- Computational complexity

  Laplace's definition is computationally inefficient, typically requiring $O(n!)$ operations to compute $|A|$ for an $n \times n$ matrix.

  One may use LU decomposition to compute determinants in $O(n^3)$ as there are efficient algorithms for LU decomposition.

Determinant is multiplicative

Steps in order to show $\det(AB) = \det(A)\det(B)$.

- $\det(D(\lambda_1, \ldots, \lambda_n)) = \lambda_1 \cdots \lambda_n$ and $\det(I_n) = 1$.
- $\det(AE) = \det(EA) = \det(A)\det(E)$ for elementary matrix E.
- $\det(A) \neq 0$ if and only if A is invertible (has full rank).
- $\det(AB) = \det(A)\det(B)$

If $\operatorname{rank}(A) < n$, then $\operatorname{rank}(AB) \leq \operatorname{rank}(A) < n$. Thus,

$$\det(AB) = 0 = \det(A)\det(B).$$

Similar arguments apply if $\operatorname{rank}(B) < n$ (use row rank). If both A and B are invertible, then

$$A = E_1 \cdots E_k,\ B = E'_1 \cdots E'_l$$

can be expressed as products of elementary matrices (Gaussian elimination). As determinant is multiplicative for elementary matrices, we are done.

Remark The exponential of matrix A is defined to be

$$e^A = I_n + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

This is an invertible matrix and

$$\det(A) = e^{\mathrm{trace}(A)}.$$

In particular, e^{-A} is the inverse of e^A and

$$\det(e^A e^B) = e^{\mathrm{trace}(A+B)}$$

although $e^{A+B} \neq e^A e^B$.

Characterization of determinant

Theorem *If $\delta : M_n(\mathbb{R}) \to \mathbb{R}$ is a function such that*

(i) δ *is a linear function of a column if other columns are fixed;*

(ii) δ *changes sign if we swap two columns;*

(iii) $\delta(I_n) = 1$,

then δ is the determinant.

Remark The first property is called *multilinearity* in columns.

The second property is called *alternating*. This is related to the sign of permutations as a swap is a permutation with sign -1. This property is equivalent to $\delta(A) = 0$ if A has two identical columns.

The third property is called *normalization*. It is just to conform with our idea that the volume of a unit n-cube is 1 although it is rather arbitrary as there is no universal notion of unit.

If we start with Laplace's definition of cofactor expansion along row 1, then we observe (via induction) that det satisfies the hypothesis of the Theorem. Now start with cofactor expansion along any other row. This definition also satisfies (i) through (iii), whence must equal the earlier definition using row 1. Thus, Laplace's definition is independent of the choice of row.

We may also check that the defintions given by Laplace and Leibniz are the same. A typical proof involves analyzing permutations and their signs carefully.

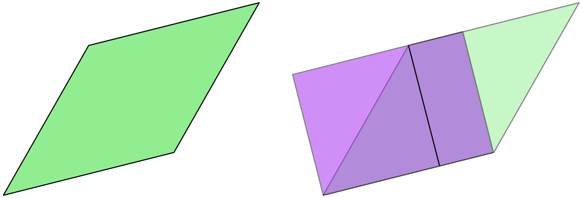
The motivations for conditions (i) through (iii) are arising from calculus, while studying (signed) volumes of polytopes. In general, if (iii) is omitted from the Theorem, then we may conclude that

$$\delta(A) = \delta(I_n)\det(A).$$

Determinant as volume

Given vectors $\{v_1, \ldots, v_n\}$ in $\mathbb{R}^n$, consider the parallelopiped P obtained from it.

Definition [volume] The volume of P is defined inductively as follows. For $n = 1$ it is the length of v_1. For a parallelogram, the volume



is the length of v_1 multiplied by the height, i.e., write $v_2 = \lambda v_1 + w_2$ where $w_2 \cdot v_1 = 0$ and height is the length of w_2.

Assume that we have defined inductively the volume of parallelopipeds obtained from $n-1$ vectors. Define the volume of P, obtained from $\{v_1, \ldots, v_n\}$ in $\mathbb{R}^n$, as the product of the volume associated to $\{v_1, \ldots, v_{n-1}\}$ and the length of w_n, where $v_n = u_n + w_n$ is chosen such that

$$w_n \cdot v_j = 0, \ j = 1, \ldots, n-1$$

and $u_n \in \text{span}\{v_1, \ldots, v_{n-1}\}$.

The volume is independent of the ordering of the vectors.

Theorem *Given an m-dimensional parallelopiped P in $\mathbb{R}^n$, if A denotes the $m \times n$ matrix with rows of A being the edges of P, then*

$$\text{volume}(P)^2 = \det(AA^t).$$

In particular, for a square matrix $|\det(A)| = \text{volume}(P)$.

Determinants in calculus

Determinants appear in change of variables. If we consider a map φ from $\mathbb{R}^n$ to $\mathbb{R}^n$ (typically locally defined, invertible and smooth), then write $(v_1, \ldots, v_n) = \varphi(u_1, \ldots, u_n)$. The relation between infinitesimal volume elements is given by

$$dv_1 dv_2 \cdots dv_n = |\det((D\varphi)(u_1, \ldots, u_n))| \, du_1 du_2 \cdots du_n$$

where $(D\varphi)(u_1, \ldots, u_n)$ is the matrix of partial derivatives. This matrix is called the Jacobian of φ.

If we use differential forms, then we may get rid of the modulus but then $du_1 du_2 \cdots du_n$ has to be replaced by $du_1 \wedge du_2 \wedge \cdots \wedge du_n$, a differential n-form.