# Determinants : brief history, geometric interpretation, properties, calculation, applications

P. Sam Johnson

**National Institute of Technology Karnataka (NITK)**
**Surathkal, Mangalore, India**

# Introduction

Determinant is a scalar associated with a square matrix.

- most important uses of determinants within "Linear Algebra" is in **the study of eigenvalues**.
- occur in **Cramer's rule** for solving linear equations
- can be used to give **a formula for the inverse of a non-singular matrix**
- in calculus of several variables, the **Jacobian** used in transforming a multiple integral uses determinant.

# Brief historical introduction

Based on two simultaneous quadratic equations

$$a_{11}x^2 + a_{12}x + c_{12} = 0,$$

$$a_{21}x^2 + a_{22}x + c_{22} = 0,$$

**a famous Japanese mathematician**, **Seki Kowa** (1642-1708) first discovered the idea of $2 \times 2$ determinant in 1683.

Eliminating $x^2$, we get $(a_{11}a_{22} - a_{21}a_{12})x + (a_{11}c_{22} - c_{21}a_{12}) = 0$. Seki Kowa then used the coefficient of $x$ to introduce a $2 \times 2$ determinant.

In the same year 1683, a famous **German mathematician**, Gottfried Wilhelm **Leibniz** (1646-1716) independently discovered the same idea of determinant in Europe. He wrote to de **L'Hospital** (1661-1740) explaining his idea in solving a system of linear equations.

Both **Seki Kowa and Leibniz** not only discovered the idea of determinant as unique scalar, but also knew many elementary properties of determinants including how to expand a determinant using any row or column what is now known as the **Laplace expansion** and how to determine which terms in the expansion are positive and which terms are negative.

Thus, the idea of determinant appeared **in Japan and Germany** at almost exactly the same time, but Seki Kowa in Japan certainly first published his work on determinant.

Figure: Seki Kowa

**Seki Kowa** was born on the same year in which **Sir Isaac Newton** (1642-1727) was born in England. Seki Kowa has been described as **Japan's "Newton".** He is frequently compared with **Archimedes** (287 - 212 B.C), **Newton** and Carl Friedrich **Gauss** (1777-1855).

In 1730s, Colin **Maclaurin** (1698-1746), a brilliant English mathematician, wrote **Treatise of Algebra** that was published in 1748, two years after his death. It contained first published results on determinants including **Cramer's rule** of solving two and three simultaneous linear equations, and indicating how to solve the four simultaneous linear equations in four unknowns by the fourth-order determinant.

Indeed, in 1750, **Cramer** gave an elegant general rule for solving $n$ linear simultaneous equations in $n$ unknowns $x_1, x_2, \ldots, x_n$. Cramer's solutions are given in terms of determinants. Cramer also published his rule in his treatise **Introduction to the Analysis of Algebraic Curves** in 1750.

Cramer's method is analytically exact, but computationally inefficient for all but small systems of linear equations because it involves computation of determinants of large orders. So, the computation of $\det A$ of order $n$ from its definition as a major problem of computating $n!$ terms.

Subsequently, other iterative methods and numerical techniques have replaced the use of determinants for solving linear systems and computing eigenvalues of matrices. Among other methods, Carl Gustave **Jacobi's** (1804-1851) **method** and its refinement known as the **Gauss-Seidel method** published by Phillip Ludwig (1821-1896) in 1874 have been successful as direct methods for solving linear system of equations.

Given a system of $n$ linear equations in $n$ unknowns, we solve the first equation in $x_1$, in the second for $x_2$ and so on. Thus, **starting with an initial approximation**, we use these new equations to iteratively update each unknown. **Jacobi's method uses all of the values at the $r$th iteration to compute the $(r+1)$th iterate, wheras the Gauss-Seidel method always uses the most recent value of each unknown in every calculation.**

Naturally, the quesion of convergence will arise about these iterative methods. Indeed, there are examples in which one of the methods converges and the other diverges. However, if either of these methods converges, then it must converge to the solution – it cannot converge to some other point. **These iterative methods are widely used when working with large determinants.**

In 1801, it was **Gauss**, universally known as the greatest German mathematician, who first used the properties of determinant in his **Disquisitiones Arithmeticae** to develop the theory of quadratic forms. Subsequently, Augustin Louis **Cauchy** (1789-1857), one of the greatest French mathematicians, made significant contributions to the theory of determinants in the modern sense in the context of quadratic forms in $n$ variables.

In 1812, Cauchy proved the multiplication theorem for determinants for the first time, and proved results on diagonalization of a matrix in the context of converting a form to the sum of squares. He reproved many earlier results, broadened, deepened and generalized them with rigorous mathematical style. In the context of determinants, he discovered formulae for volumes of parallelepied, tetrahedron and several other solid polyhedra.

# Notation for determinant.

Determinants arose independently of matrices in the solution of many practical problems, and the theory of determinants was well developed almost two centuries before the discovery of matrices.

**Arthur Cayley** (1821-1895), a British mathematician first introduced the notation of two vertical lines on either side of the array to denote the determinant which has become standard.

# Geometric interpretation of the determinant of a square matrix of size 2

Consider $\mathbb{R}^2$. Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ be any two nonzero vectors. We want to compute the area of the parallelogram spanned by these vectors. We can rearrange the given set into a rectangle by cutting away a right-angled triangle and pasting it on the opposite side. Hence the area of the parallelogram is "the base times the height". After calculation, the area of the parallelogram is $|x_1 y_2 - x_2 y_1|$. We may thus think of $\det(x, y)$ as the **signed area** of the parallelogram spanned by $x$ and $y$.

Let $A$ be the area of the unit square $S$ spanned by $\{e_1, e_2\}$. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map such that $Te_1 = v_1$ and $Te_2 = v_2$. Then the area of the parallelogram spanned by $v$ and $w$ is given by $\texttt{Area}[v, w] = |\det(T)| = |\det(T)|\ \texttt{Area}[e_1, e_2]$.

Thus the linear map $T$ distorts the area of $S$ by the factor $|\det(T)|$; that is, $det(T)$ is the factor by which the area of the unit square in $\mathbb{R}^2$ is multiplied to get the area of $T(S)$.

# Generalization to a square matrix of order $n$

Let $V$ be any vector space of dimension $n$. We wish to define determinant as a function which attaches to any $n$-tuple of vectors $(v_1, v_2, \ldots, v_n) \in V^n$ a real number. This number is to be thought of as **the signed volume of the parallelepied spanned by $v_i$'s.**

In any kind of measurement we need a unit against which others are measured. In our case this means that we have to make a choice of a parallelepied and declare its volume as 1.

Will any $n$ vectors $[v_1, v_2, \ldots, v_n]$ do ? No ! For, if they are linearly dependent the parallelepied lies in a vector subspace of dimension at most $n-1$ and hence its $n$-dimensional volume must be zero. Thus, a basis $\{e_i\}_{i=1}^n$ of $V$ must be fixed.

Based on our geometric intuition, we expect this function $f : V^n \to \mathbb{R}$ to possess the following geometric properties:

- **Magnified by $\alpha$ :**
  For all $\alpha \in \mathbb{R}$, and for all $i$,
  $$f(v_1, v_2, \ldots, \alpha v_i, \ldots, v_n) = \alpha f(v_1, v_2, \ldots, v_i, \ldots, v_n).$$
- **Area/Volume unaltered by cutting and rearranging :**
  For all $i \neq j$,
  $$f(v_1, v_2, \ldots, v_i, \ldots, v_j \ldots, v_n) = f(v_1, v_2, \ldots, v_i + v_j, \ldots, v_j, \ldots, v_n).$$
- **Normalization condition :**
  If $\{e_1, e_2, \ldots, e_n\}$ is the chosen basis of $V$,
  $$f(e_1, e_2, \ldots, \ldots, e_n) = 1.$$

Assume that there exists a function $f : V^n \to \mathbb{R}$ satisfying the above (basic) properties. Then

1. $f(v_1, v_2, \ldots, v_n) = 0 \texttt{ if } v_i = 0 \texttt{ for some } i$ with $1 \leq i \leq n$.
2. For $i \neq j$ and $\alpha \in \mathbb{R}$,
   $f(v_1, v_2, \ldots, v_i, \ldots, v_j \ldots, v_n) = f(v_1, v_2, \ldots, v_i + \alpha v_j, \ldots, v_j, \ldots, v_n)$.
3. More generally, we see that, for any $\alpha_j \in \mathbb{R}$ and $i \neq j$,
   $f(v_1, v_2, \ldots, v_i, \ldots, v_j \ldots, v_n) =$
   $f(v_1, v_2, \ldots, v_i + \sum_{i \neq j} \alpha_j v_j, \ldots, v_j, \ldots, v_n)$.
4. $f(v_1, v_2, \ldots, v_n) = 0$ if $\{v_1, v_2, \ldots, v_n\}$ is linearly dependent.
5. For any $j \in \{1, 2, \ldots, n\}$ and for any $v_j', v_j''$ we have
   $f(v_1, v_2, \ldots, v_i' + v_i'', \ldots, v_n) =$
   $f(v_1, v_2, \ldots, v_i', \ldots, v_n) + f(v_1, v_2, \ldots, v_i'', \ldots, v_n)$.
6. $f(v_1, v_2, \ldots, v_i, \ldots, v_j, \ldots, v_n) = -f(v_1, v_2, \ldots, v_j, \ldots, v_i, \ldots, v_n)$.

Let $V$ be any vector space. An $n$-**linear map** is a function $f : V^n \to \mathbb{R}$ such that for each $i$, $1 \leq i \leq n$, the following are true:
$f(v_1, \dots, v_i + w_i \dots, v_n) = f(v_1, \dots, v_i \dots, v_n) + f(v_1, \dots, w_i \dots, v_n)$
$f(v_1, \dots, \alpha v_i \dots, v_n) = \alpha f(v_1, \dots, v_i \dots, v_n)$ for all $v_i, w_i \in V$ and $\alpha \in \mathbb{R}$.

Let $S_n$ denote the set of permutations (bijections) of the set $\{1, 2, \dots, n\}$. Let $f : V^n \to \mathbb{R}$ be an $n$-linear map. $f$ is said to be **skew-symmetric** if $f(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(n)}) = \texttt{sign}\ (\sigma) f(v_1, v_2, \dots, v_n)$ for all $\sigma \in S_n$.

If $f : V^n \to \mathbb{R}$ satisfying the above two basic properties, then $f$ is $n$-linear and skew-symmetric.

Let $V$ be a vector space. Fix a basis $\{e_1, e_2, \ldots, e_n\}$ of $V$. Then **there exists a unique function** $g : V^n \to \mathbb{R}$ such that

1. $g$ is $n$-linear.
2. $g$ is skew-symmetric.
3. $g(e_1, e_2, \ldots, e_n) = 1$.

If $g$ is defined by setting

$$g(v_1, v_2, \ldots, v_n) = \sum_{\sigma \in S_n} \mathtt{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

where $v_i = \sum_{j=1}^{n} a_{ij} e_j$, then $g$ has the above three properties.

We call the $g$ as the **determinant** and denote it by det.

**General form $n$-linear and skew-symmetric:** Any map $f : V^n \to \mathbb{R}$ which is $n$-linear and skew-symmetric is of the form

$$f = f(e_1, e_2, \ldots, e_n)\ det.$$

That is,

$$f(B) = f(e_1, e_2, \ldots, e_n)\ det(B).$$

Moreover, the function det satisfies the following properties:

- det is $n$-linear
- If one interchanges $v_i$ and $v_j$ for $i \neq j$, then the determinants are of opposite sign. More generally,

$$\det(v_{\sigma(1)}, v_{\sigma(2)}, \ldots, v_{\sigma(n)}) = \texttt{ sign } (\sigma) \det(v_1, v_2, \ldots, v_n)$$

  for any $\sigma \in S_n$.
- $\det(e_1, e_2, \ldots, e_n) = 1$.
- $\det(v_1, v_2, \ldots, v_n) = 0$ if $v_i$ are linearly dependent.

# Determinant function on $M(n, \mathbb{R})$, the set of $n \times n$ matrices with real entries

Let $A \in M(n, \mathbb{R})$ and write $A(C_1, \dots, C_n)$ where $C_i$ is the $i$th column of $A$.

We then define

$$\det A := \det(C_1, \dots, C_n)$$

where det is the $n$-linear skew-symmetric function on $\mathbb{R}^n \times \mathbb{R}^n \times \cdots \mathbb{R}^n \to \mathbb{R}$ with $\det(e_1, \dots, e_n) = 1$.

Note that $\det A = \det(Ae_1, \dots, Ae_n)$ for $A \in M(n, \mathbb{R})$.

If $A = (a_{ij})$ is a square matrix of order $n$, the **determinant** of $A$ is

$$\det A = \sum \mathtt{sign}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \tag{1}$$

where $\sigma$ runs over the $n!$ permutations of $(1, 2, \ldots, n)$.

The summation, therefore, extends over $n!$ permutations of $(1, 2, \ldots, n)$, half of which are even and half odd. Thus (1) is a **formal definition of a determinant of order** $n$ and **seldom used in practice for evaluating a determinant of higher order** $n > 3$. However, it is effectively used for proving the properties of determinants with applications.

Each term in the above expression contains exactly one element from each row and exactly one element from each column of the matrix.

# Properties of determinants

- The determinant of $A$ is 0 if $A$ has either a null row or a null column.
- The determinant of $A$ is 0 if any two rows (or columns) are identical.
- The determinant of a triangular matrix is the product of the diagonal elements.
- The interchange of any two rows (or columns) of a matrix changes the sign of its determinants.
- The determinant of a matrix and its transpose are equal, that is, $\det A = \det A^T$.

# Properties of determinants

- If $B$ is obtained from $A$ by multiplying any one row or column by $k$, $\det B = k \det A$. The determinant of a matrix $kA$ of order $n$ is $k^n \det A$.
- For a fixed $k$, let the $k$-th row of $A$ be the sum of two row vectors $x^T$ and $y^T$. Then $\det A = \det B + \det C$ where $B$ (resp. $C$) is obtained from $A$ by replacing the $k$-th row by $x^T$ (resp. $y^T$).
  Thus from the above two points, the determinant is a linear function of the $k$th row when the other rows are kept fixed.
- If a scalar multiple of one row (resp. column) is added to another row (resp. column) of a matrix, the determinant of the matrix is not altered.

## Properties of determinants

- **An efficient method for evaluating** $\det A$**:**
  We reduce $A$ to an upper triangular matrix $B$ using elementary row operations. Suppose $\alpha_1, \alpha_2, \ldots, \alpha_p$ are the scalars used in the row operations of the type 'multiplying a row by a non-zero scalar' and suppose $q$ interchanges of rows are used. Then clearly

$$\det A = (-1)^q \frac{b_{11} b_{22} \cdots b_{nn}}{\alpha_1 \alpha_2 \cdots \alpha_p}.$$

- Since the $\alpha_i$'s are non-zero, $\det A \neq 0$ iff all the diagonal elements of $B$ are nonzero.
- A square matrix $A$ is non-singular iff $\det A \neq 0$.
- $\det A^{-1} = 1/(\det A)$.

# Geometric interpretation of the determinant of a square matrix of size 3

For a $3 \times 3$ matrix $A$, $\det A$ is the volume of the parallelepied formed with $OP, OQ$ and $OR$ as three edges, where $O$ is the origin and $P, Q$ and $R$ are the points of $\mathbb{R}^3$ corresponding to the three rows of $A$. By convention, the volume is positive iff $OP, OQ$ and $OR$ from a right-handed system.

That is, the following happens: imagine a right-handed screw placed perpendicular to the plane $OPQ$ with its tip at the origin. If the screw is rotated from $OP$ to $OQ$ by the smaller angle, the line $OR$ lies on the side of the plane $OPQ$ into which the screw advances.

# A recursive method for evaluating determinants by expressing a determinant of order $n$ in terms of determinants of order $n-1$.

Let $A$ be a matrix of order $n \geq 2$. Then the **cofactor** of $a_{ij}$ in $A$ as $A_{ij} = (-1)^{i+j} \det(M_{ij})$ where $M_{ij}$ is the matrix obtained from $A$ by deleting the $i$-th row and the $j$-th column. The $\det M_{ij}$ is called the **minor** of $a_{ij}$.

**A determinant can be expanded by any row.** Let $A$ be a square matrix of order $n$ and $k$ an integer such that $1 \leq k \leq n$. In 1772, Pierre Simon **Laplace** (1749-1827), a famous French mathematician, proved the expansion of determinant of order $n$ in terms of minors or cofactors along the $i$-th row in the form

$$\det A = \sum_{j=1}^{n} a_{kj} A_{kj}.$$

The expansion is universally known as the **Laplace Expansion**.

**A determinant can be expanded by any column.** We can similarly expand $\det A$ by the $k$-th column as follows:

$$\det A = \sum_{i=1}^{n} a_{ik} A_{ik}.$$

Clearly it is best to expand a determinant by a row or a column having the maximum number of zeros.

Let $A$ be an $n \times n$ matrix. Suppose now $I = \{i_1, \ldots, i_k\}$ and $J = \{j_1, \ldots, j_k\}$. The submatrix $A(I|J)$ is obtained by deleting $\overline{I}$ rows and $\overline{J}$ columns, where $\overline{I}$ and $\overline{J}$ are the complements of $I$ and $J$ respectively in $\{1, 2, \ldots, n\}$. Its determinant is called a $k$-**rowed minor** of $A$. Then the **cofactor** of $A(I|J)$ in $A$ is

$$A_{IJ} = (-1)^{i_1 + \cdots + i_k + j_1 + \cdots + j_k} |A(\overline{I}|\overline{J})|.$$

**Laplace Expansion.** Let $A$ be an $n \times n$ matrix and let $I = \{i_1, \ldots, i_k\}$. Then

$$\begin{aligned} |A| &= \sum_J |A(I|J)|\ A_{IJ} \\ &= \sum_J (-1)^{i_1+\cdots+i_k+j_1+\cdots+j_k} |A(I|J)||A(\bar{I}|\bar{J})| \end{aligned}$$

where $J$ runs over all subsets of $\{1, 2, \ldots, n\}$ with size $k$ and $\bar{I}, \bar{J}$ are the complements of $I$ and $J$ respectively in $\{1, 2, \ldots, n\}$.

Note that there are $nc_k$ terms in the sum and each term is a product of matrices of orders $k$ and $n-k$ respectively.

# Cramer's rule

Consider a system of three equations in three unknowns, $Ax = b$, where $x = (x_1, x_2, x_3)^T$. Let $A_1, A_2$ and $A_3$ be the columns of $A$. Then

$$\begin{aligned} x_1 \det(A_1, A_2, A_3) &= \det(x_1 A_1, A_2, A_3) \\ &= \det(x_1 A_1 + x_2 A_2 + x_3 A_3, A_2, A_3) \\ &= \det(b, A_2, A_3). \end{aligned}$$

Therefore $x_1 = \frac{\det(b, A_2, A_3)}{\det A}$.

**Generalization for any $n$.** Let $A \in M(n, \mathbb{R})$ with $\det A \neq 0$. Let $b \in \mathbb{R}^n$ be a column vector. Then the solution of $Ax = b$ is given by

$$x_j = \frac{\det(A_1, \ldots, b, \ldots, A_n)}{\det A}$$

where $b$ is in the $j$th place and $\det A = \det(A_1, \ldots, A_n)$ where $A_i$ is the $i$th column of $A$.

**Interpolating a polynomial of $(n-1)$ degree.** We next consider a polynomial of degree $n$ in the form

$$y = P(x) = a_1 + a_2 x + \cdots + a_n x^{n-1}$$

which passes through $n$ points $(x_i, y_i), i = 1, 2, \ldots, n$ so that $P(x_i) = y_i$ leads to a system of $n$ linear equations with $n$ variables $a_1, a_2, \ldots, a_n$. In matrix notation, this system reads

$$V_n \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \ldots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

The $n \times n$ coefficient matrix is also known as **Vandermonde's matrix** and the associated $n$-th order Vandermonde's determinant is given by

$$\begin{aligned} \det V_n &= \Pi_{1 \leq j < i \leq n}(x_i - x_j) \\ &= (x_2 - x_1) \cdots (x_n - x_1)(x_3 - x_2) \cdots (x_n - x_2) \cdots (x_n - x_{n-1}). \end{aligned}$$

Clearly, $V_n$ is non-singular iff $x_1, x_2, \ldots, x_n$ are distinct.

The Vandermonde matrix occurs in many contexts like polynomial interpolation, regression, weighing designs, coding theory, signal processing, secret codes in cryptography and error-correcting codes in digital communication.

In addition to major work on the theory of determinants by Cramer and Vandermonde, in 1764, Etienne Bezout (1730-1783) also made some significant contributions to determinants including the solution of $n$ homogeneous equations in $n$ unknowns. He proved that the nontrivial solutions of this system exist provided the determinant of the coefficient matrix is zero.

# Applications of determinants

Determinants were closely related to areas and volumes in geometry and can be used to produce equations of lines, planes and other curves.

- There exists a unique straight line passing through two distinct points $(x_1, y_1)$ and $(x_2, y_2)$ in a plane. If the equation of this line is $ax + by + c = 0$, then $ax_1 + by_1 + c = 0$ and $ax_2 + by_2 + c = 0$. Thus, we have a system of three linear equations with variables $a, b$ and $c$. So there exists a nontrivial solution for $a, b$ and $c$ provided the determinant of the coefficient matrix must be zero. That is,

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0.$$

  This is **the determinant form of the equation of a line passing through two points** $(x_1, y_1)$ **and** $(x_2, y_2)$.

Three points $(x_1, y_1), (x_2, y_2)$ and $(x_3, y_3)$ are **collinear** if and only if

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

Similarly, the equation of the **plane passing through three noncollinear points** $(x_1, y_1, z_1), (x_2, y_2, z_2)$ and $(x_3, y_3, z_3)$ in the $xyz$-space has the determinant form

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0.$$

# References


- L. Debnatha, "*A brief historical introduction to determinants with applications*", International Journal of Mathematical Education in Science and Technology, Volume 44, Issue **3**, 15 April 2013, 388-407.
- S. Kumaresan, "*Linear Algebra - A Geometric Approach*", PHI Learning Pvt. Ltd., 2011.
- A. Ramachandra Rao and P. Bhimasankaram, "*Linear Algebra*", Hindustan Book Agency, 2000.