Lecture 9: Determinants

Let $A \in M_{n\times n}$ be a square matrix with real or complex entries

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \tag{1}$$

The determinant of A is the real or complex number

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A^{ij}) \qquad \text{(for every fixed } i\text{)}, \tag{2}$$

where a_{ij} are the entries of A, and A^{ij} is a matrix obtained from A by crossing out the i-th row and the j-th column[1].

The expression (2) is called *Laplace expansion* of the determinant along the i-th row. As we will see hereafter $\det(A) = \det(A^T)$ and therefore there exists an equivalent *Laplace expansion* along the j-th column, which is

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A^{ij}) \qquad \text{(for every fixed } j\text{)}. \tag{4}$$

Remark: The fact that we can arbitrarily choose the row or the column along which develop the determinant and always obtain the same result suggests that the determinant is a rather special function. From a technical viewpoint it can be shown that (2) is the a unique alternating multilinear function[2]

$$\Lambda(\cdot) : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ times}} \to \mathbb{R}$$
$$(a_1, \ldots, a_n) \to \Lambda(a_1, \ldots, a_n) \tag{6}$$

satisfying

$$\Lambda(e_1, \ldots, e_n) = 1, \tag{7}$$

where $\{e_1, \ldots, e_n\}$ is the canonical basis of $\mathbb{R}^n$. In other words, we have

$$\det(A) = \Lambda(a_1, \ldots, a_n), \tag{8}$$

where a_j is the j-th column of A.

---

[1] The number

$$c^{ij} = (-1)^{i+j} \det(A^{ij}) \tag{3}$$

is often called co-factor of a_{ij} in the determinant expansion.

[2] An alternating multilinear function is a function $\Lambda(a_1, \ldots, a_n)$ that is linear in each argument a_j, e.g.,

$$\Lambda(a_1, a_2 + b_2, a_3) = \Lambda(a_1, a_2, a_3) + \Lambda(a_1, b_2, a_3),$$

and changes sign if we interchange a_j with a_i. For instance,

$$\Lambda(a_1, a_2, a_3) = -\Lambda(a_2, a_1, a_3) = \Lambda(a_3, a_1, a_2). \tag{5}$$

If $A = a$ is a number then we set $\det(A) = a$. Note that the determinant of a matrix is a *nonlinear* function of the matrix entries which is defined recursively in terms of determinants of the matrices A^{ij}, which have smaller dimension.

Examples: Let us provide a few examples of calculation of determinants

1. $A = a$ (real number). In this case we have $\det(A) = a$.

2. $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$. We develop the determinant along the first row, i.e., set $i = 1$ in (2).

   This yields
   $$\det(A) = (-1)^{1+1} a_{11} \det(A^{11}) + (-1)^{1+2} a_{12} \det(A^{12}), \tag{9}$$
   where
   $$A^{11} = a_{22}, \qquad A^{12} = a_{21}. \tag{10}$$
   Therefore we obtain
   $$\det(A) = a_{11}a_{22} - a_{12}a_{21}. \tag{11}$$
   Note that we obtain exactly the same formula if we develop the determinant along the second row, the first column or the second column.

3. $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. We develop the determinant along the first row, i.e., set $i = 1$ in (2).

   This yields
   $$\det(A) = (-1)^{1+1} a_{11} \det(A^{11}) + (-1)^{1+2} a_{12} \det(A^{12}) + (-1)^{1+3} a_{13} \det(A^{13}) \tag{12}$$
   where
   $$A^{11} = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \qquad A^{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} \qquad A^{11} = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}. \tag{13}$$
   Computing the determinants of A^{11}, A^{12} and A^{13} yields the formula
   $$\det(A) = a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}). \tag{14}$$
   Note that we obtain exactly the same formula if we develop the determinant along any other row or column.

Since we can equivalently expand the determinant along arbitrary rows orcolumns of A it is convenient to choose the row or the column with the largest number of zeros. This minimizes the number of calculations when computing the determinant using (2) or (4). For example, it is clear that it is convenient to compute the determinant of the following matrix along the second column:

$$\det\left(\begin{bmatrix} 3 & 1 & 4 \\ -1 & 0 & 2 \\ 1 & 0 & 5 \end{bmatrix}\right) = -1 \det\left(\begin{bmatrix} -1 & 2 \\ 1 & 5 \end{bmatrix}\right) = -1(-5-2) = 7. \tag{15}$$

Properties of the determinant. The determinant of any $n \times n$ matrix A satisfies the following important properties:

1. $\det(A) = \det\left(A^T\right)$
2. $\det(A)$ is a *linear function* of the columns (or the rows) of the matrix A. In other words, if we denote by a_i the i-th column of A and B a column vector of the same length of a_i then
   (a) $\det\left(\begin{bmatrix} a_1 \cdots (a_i + b) \cdots a_n \end{bmatrix}\right) = \det\left(\begin{bmatrix} a_1 \cdots a_i \cdots a_n \end{bmatrix}\right) + \det\left(\begin{bmatrix} a_1 \cdots b \cdots a_n \end{bmatrix}\right)$,
   (b) $\det\left(\begin{bmatrix} a_1 \cdots ca_i \cdots a_n \end{bmatrix}\right) = c\det\left(\begin{bmatrix} a_1 \cdots a_i \cdots a_n \end{bmatrix}\right)$,

   where a_i is the i-th column of A.
3. If the columns or the rows of A are linearly dependent then $\det(A) = 0$. If the columns or the rows or A are linearly independent (i.e., A is full rank) then $\det(A) \neq 0$.
4. If a multiple of one row (or one column) is added or subtracted to another row (or column) then the determinant does not change (this follows from property 2 by setting with $B = cA_j$, and property 3).
5. If two rows (or two columns) are interchanged then the determinant changes sign.

These properties can be easily verified for 2×2 and 3×3 matrices. The proof of these properties for general $n \times n$ matrices can be found in the book.

Note that from property 2(b) it follows that for any number c and any $n \times n$ matrix A:

$$\det(cA) = c^n \det(A). \tag{16}$$

In fact, the matrix cA has all columns (n in total) multiplied by c.

Example: Let us show properties 1. to 5. for the simple matrix

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, \qquad \det(A) = 1. \tag{17}$$

We have:

1. $\det\left(A^T\right) = \det\left(\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}\right) = 1 = \det(A)$.
2. (a) $\det\left(\begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}\right) + \det\left(\begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}\right) = 3 - 2 = 1.$
   (b) Multiply the second column of A by 3. This yields $\det\left(\begin{bmatrix} 1 & 3 \\ 2 & 9 \end{bmatrix}\right) = 3 = 3\underbrace{\det(A)}_{=1}$.
3. A has rank 2 and therefore its columns are linearly independent. However, if we consider the rank 1 matrix $\begin{bmatrix} 1 & 3 \\ 1 & 3 \end{bmatrix}$ then

$$\det\left(\begin{bmatrix} 1 & 3 \\ 1 & 3 \end{bmatrix}\right) = 3 - 3 = 0. \tag{18}$$

4. Multiply the first row of A by 2 and add it to the second row to obtain
$$\det\left(\begin{bmatrix}1 & 2\\ 3 & 7\end{bmatrix}\right) = 7-6 = 1 = \det(A). \tag{19}$$

5. Interchange the first and the second row of A to obtain
$$\det\left(\begin{bmatrix}1 & 3\\ 1 & 2\end{bmatrix}\right) = 2-3 = -1 = -\det(A). \tag{20}$$

From Property 3. at page 3 it follows that:

Theorem 1. Let A be a $n \times n$ matrix, b a $n \times 1$ column vector. Then

1. A is invertible $\Leftrightarrow$ $\det(A) \neq 0$.
2. A has linearly independent rows/columns $\Leftrightarrow$ $\det(A) \neq 0$.
3. The linear system of equations $Ax = b$ has a unique solution $\Leftrightarrow$ $\det(A) \neq 0$.

Determinant of matrix products and matrix inverses. By using the definition of determinant (2) it can be shown that for every $A, B \in M_{n\times n}$ we have
$$\det(AB) = \det(A)\det(B). \tag{21}$$
Clearly, this implies that
$$\det(AB) = \det(BA) \quad \text{and} \quad \det(A^p) = \det(A)^p \quad \text{for all } p \in \mathbb{N}. \tag{22}$$
By using these identities it is straightforward to show, e.g.,
$$\det(AB^TACAB) = \det(A)^3\det(B)^2\det(C), \tag{23}$$
where A, B and C are three $n \times n$ matrices. Moreover, if A is an invertible matrix then
$$1 = \det(AA^{-1}) = \det(A)\det(A^{-1}) \quad \Rightarrow \quad \det(A^{-1}) = \frac{1}{\det(A)}, \tag{24}$$
i.e. the determinant of the inverse matrix is the inverse of the determinant.

Computing the determinant of a matrix efficiently. How do we actually compute the determinant of a matrix? We have seen that one possibility is to use the definition (2), i.e., the Laplace rule. However, this not really computationally efficient if the dimension of the matrix is even moderately high, e.g., larger than 10 or 20. In fact, it can be shown that the number of operations to compute (2) is exactly
$$p = \lfloor n!e \rfloor - 2. \tag{25}$$
In this formula, $e = 2.7183...$ is the Napier number and the symbol $\lfloor n!e \rfloor$ denotes the nearest integer number smaller or equal than $n!e$, where $n!$ is the factorial of n. For instance, if $n = 2$ we have
$$p = \lfloor 2!e \rfloor - 2 = \lfloor 5.4366 \rfloor - 2 = 5 - 2 = 3.$$

In fact, as we see from equation (11), to compute the determinant we need two multiplications and one subtraction. Similarly, for 3×3 matrices ($n = 3$) we need

$$p = \lfloor 3!e \rfloor - 2 = \lfloor 16.3097 \rfloor - 2 = 16 - 2 = 14$$

operations. In fact, as we see from equation (12), to compute the determinant we need 9 multiplications and 5 subtractions. The number of operations increases exponentially fast as we increase the dimension of the matrix. For example, for a 20×20 matrix the Laplace rule (2) requires

$$p = \lfloor 20!e \rfloor - 2 \simeq 6.61 \times 10^{18} \qquad \text{operations.}$$

The 2022 Apple M2 Max processor is capable of 13.6 Teraflops in single precision (32 bits), i.e., 13.6×10^{12} single-precision floating point operations per second. Hence, to compute the determinant of a 20×20 matrix by using the Laplace rule on the latest MacBook Pro with M2 Max processor/GPU we need to let our laptop run for approximately

$$\frac{6.61 \times 10^{18} \text{ operations}}{13.6 \times 10^{12} \text{ flops}} = 4.86 \times 10^5 \text{ seconds} \simeq 5.62 \text{ days} \tag{26}$$

to complete the calculation. Repeating a similar calculation for a 22×22 would require

$$\frac{30.554 \times 10^{20} \text{ operations}}{13.6 \times 10^{12} \text{ flops}} = 2.2632 \times 10^8 \text{ seconds} \simeq 7.18 \text{ years.} \tag{27}$$

Fortunately, there is a more efficient algorithm to compute the determinant of a matrix. In fact, by using elementary row operations we known that we can reduce the matrix A to the following matrix in row-echelon form

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{12} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix} \tag{28}$$

The matrix U has the same determinant of A, up to a sign (i.e., + or $-$) determined by how many times we interchange rows in the Gauss elimination with pivoting-by-row process. If we denote by s the number of row permutations we take in the Gauss elimination process we have

$$\det(A) = (-1)^s \prod_{k=1}^{n} u_{kk}. \tag{29}$$

In fact, the determinant of an upper-triangular (or a lower-triangular) matrix is simply the product of the diagonal elements. The total number of operations to transform an $n \times n$ matrix A into the upper triangular form U is

$$\frac{2}{3}n^3 - \frac{n^2}{2} - \frac{n}{6}. \tag{30}$$

The number of products in (29) is n, while taking the exponential is one operation. Hence, the total number of operations to compute the determinant with Gauss elimination is

$$\frac{2}{3}n^3 - \frac{n^2}{2} - \frac{n}{6} + n + 1 = \frac{2}{3}n^3 - \frac{1}{2}n^2 + \frac{5}{6}n + 1 \tag{31}$$

For a 22×22 matrix we get 6876 operations. If we use a 2022 Apple M2 Max processor, this requires

$$\frac{6876 \text{ operations}}{13.6 \times 10^{12} \text{ flops}} = 5.06 \times 10^{-10} \text{ seconds.} \tag{32}$$

Example: Consider the matrix

$$A = \begin{bmatrix} 1 & -1 & 3 \\ 1 & -2 & 1 \\ -1 & 1 & 3 \end{bmatrix} \quad \to \quad \text{Gauss elimination} \quad \to \quad U = \begin{bmatrix} 1 & -1 & 3 \\ 0 & -1 & -2 \\ 0 & 0 & 6 \end{bmatrix} \tag{33}$$

Since we did not perform any permutation we have $s = 0$ in (29) and therefore

$$\det(A) = \det(U) = -6.$$

.

Cramer's rule. It is possible to express the solution to a linear system of equations in terms of determinants. Specifically, let

$$Ax = b \tag{34}$$

be a system of n linear equations in n unknowns. Suppose that the system has a unique solution (i.e., $\det(A) \neq 0$). Then

$$x_i = \frac{1}{\det(A)} \det([a_1 \cdots b \cdots a_n]), \tag{35}$$

where $[a_1 \cdots b \cdots a_n]$ is a matrix obtained by replacing the i-th column of A (denoted by a_i) with the column vector b.

Example: Compute the solution to the following system of equations using Cramer's rule:

$$\underbrace{\begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ 2 \end{bmatrix}}_{b}. \tag{36}$$

We have $\det(A) = 1$, and therefore

$$x_1 = \det\left(\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}\right) = -1, \qquad x_2 = \det\left(\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\right) = 1. \tag{37}$$

An explicit formula for A^{-1} . Let A be an invertible matrix. By definition, the inverse of A is a square matrix (denoted as A^{-1}) with the following properties

$$AA^{-1} = I_n \qquad A^{-1}A = I_n, \tag{38}$$

where I_n is the $n \times n$ identity matrix. Let h_i be the columns of the matrix A^{-1}, i.e.,

$$A^{-1} = [h_1 \quad h_2 \quad \cdots \quad h_n] \qquad h_i \in M_{n\times 1} \quad i = 1, \dots, n. \tag{39}$$

By definition of matrix-vector product we have

$$AA^{-1} = [Ah_1 \quad Ah_2 \quad \cdots \quad Ah_n]. \tag{40}$$

At this point, define the following column vectors $e_i \in M_{n\times 1}$ $(i = 1, ..., n)$

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \qquad \cdots, \qquad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \tag{41}$$

Note that e_i is the i-th column of the identity matrix I_n. With this notation we can write the matrix equation $AA^{-1} = I_n$ as

$$[Ah_1 \quad Ah_2 \quad \cdots \quad Ah_n] = [e_1 \quad e_2 \quad \cdots \quad e_n]. \tag{42}$$

Hence, the n columns of the inverse matrix A^{-1}, i.e., h_1, ..., h_n are solutions to n linear systems

$$Ah_1 = e_1, \qquad Ah_2 = e_2, \qquad \ldots, \qquad Ah_n = e_n. \tag{43}$$

By using Cramer's rule we obtain that the i-th component of the column vector h_j is

$$h_{ji} = \frac{1}{\det(A)} \det\left([a_1 \cdots e_j \cdots a_n]\right) \tag{44}$$

where $[a_1 \cdots e_j \cdots a_n]$ is a matrix in which we replaced the i-th column a_i with e_j. By using the Laplace rule along the i-th column of $[a_1 \cdots e_j \cdots a_n]$ we obtain

$$\det\left([a_1 \cdots e_j \cdots a_n]\right) = (-1)^{i+j} \det(A^{ji}) = C^{ji} \qquad (j,i)\text{-cofactor}. \tag{45}$$

This yields the following expression

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} C^{11} & \cdots & C^{1n} \\ \vdots & \ddots & \vdots \\ C^{n1} & \cdots & C^{nn} \end{bmatrix}^T. \tag{46}$$

Example: Compute the inverse of the following matrix

$$A = \begin{bmatrix} -2 & 4 \\ 4 & 3 \end{bmatrix}. \tag{47}$$

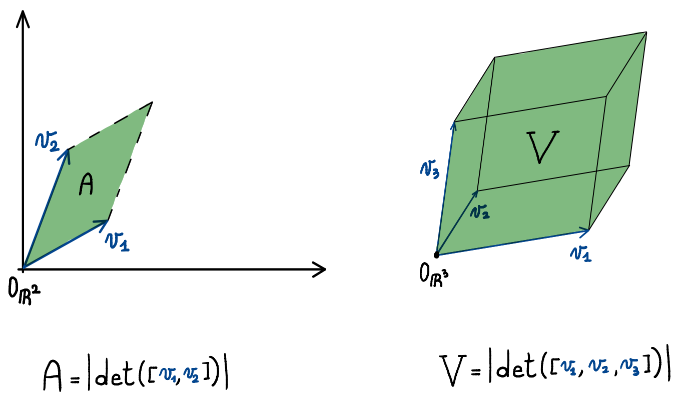
We have $\det(A) = -22$, and

$$\begin{aligned} C^{11} &= \det\left(\begin{bmatrix} e_1 & a_2 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 1 & 4 \\ 0 & 3 \end{bmatrix}\right) = 3, \\ C^{12} &= \det\left(\begin{bmatrix} a_1 & e_1 \end{bmatrix}\right) = \det\left(\begin{bmatrix} -2 & 1 \\ 4 & 0 \end{bmatrix}\right) = -4, \\ C^{21} &= \det\left(\begin{bmatrix} e_2 & a_2 \end{bmatrix}\right) = \det\left(\begin{bmatrix} 0 & 4 \\ 1 & 3 \end{bmatrix}\right) = -4, \\ C^{22} &= \det\left(\begin{bmatrix} a_1 & e_2 \end{bmatrix}\right) = \det\left(\begin{bmatrix} -2 & 0 \\ 4 & 1 \end{bmatrix}\right) = -2. \end{aligned}$$

Therefore,

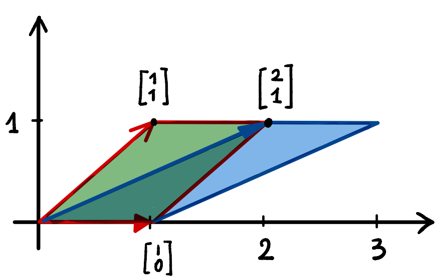
$$A^{-1} = \frac{1}{22} \begin{bmatrix} -3 & 4 \\ 4 & 2 \end{bmatrix}. \tag{48}$$

Volumes of parallelograms. The determinant of a matrix represents the volume enclosed by the vectors defined by the columns (or the rows) or the matrix.



At this point we notice that there are quite a lot of properties of A and V following from the properties of determinant. For example, if we add a scalar multiple of v_1 to v_2, then the area of the parallelogram defined by the two vectors does not change. This follows from the fact that the determinant is a linear function of the columns, and that the determinant of a matrix with linearly dependent columns is equal to zero. For example,

$$A = \left|\det\left(\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\right)\right| = \left|\det\left(\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}\right)\right|. \tag{49}$$



Of course, the green and blue areas are the same. Other properties of the area of a parallelogram can be derived from properties of the determinant. Next, consider an invertible transformation $F : \mathbb{R}^n$ $\mathbb{R}^n$, represented by $n \times n$ invertible matrix L. We know that if $\{v_1, \dots, v_n\}$ is a basis of $\mathbb{R}^n$ then

$$\begin{bmatrix}u_1 & \dots & u_n\end{bmatrix} = \begin{bmatrix}Lv_1 & \dots & Lv_n\end{bmatrix} = L\begin{bmatrix}v_1 & \dots & v_n\end{bmatrix} \tag{50}$$

is also a basis of $\mathbb{R}^n$. The volume of the parallelograms enclosed by $\{v_1, \dots, v_n\}$ and $\{u_1, \dots, u_n\}$ are

$$V_0 = \left|\det\left(\begin{bmatrix}v_1 & \dots & v_n\end{bmatrix}\right)\right|, \qquad V_1 = \left|\det\left(\begin{bmatrix}u_1 & \dots & u_n\end{bmatrix}\right)\right|. \tag{51}$$

By applying the determinant operator to equation (50), and using the fact that the determinant of the matrix product is the product of the matrix determinants we see that

$$V_1 = |\det(L)|\, V_0. \tag{52}$$

This formula is very important in a variety of fields ranging from multi-dimensional integration theory to continuum mechanics.