Determinants

A. Eremenko

September 3, 2024

In the previous lectures, an algorithm was described for determining whether a matrix is invertible, and of finding the inverse if it is. However it is frequently desirable to have a *formula* rather than an algorithm. In this lecture we define a polynomial in the matrix elements which is zero if and only if the matrix is singular, and an explicit formula for the inverse.

The guiding idea is that n vectors are linearly dependent if and only if the parallelepiped built on these vectors (one vertex of the parallelepiped is at the origin, and sides meeting at this vertex are our vectors) has zero volume. We will define the volume of a parallelepiped built on n vectors axiomatically. Let us denote this volume by $f(\mathbf{a}_1, \ldots, \mathbf{a}_n)$. Let us begin with a statement of properties which a volume of a parallelogram should satisfy:

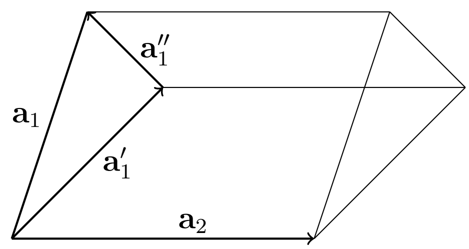
(i) f is linear with respect to each $\mathbf{a}_j$. This means that whenever $\mathbf{a}_j = \mathbf{a}'_j + \mathbf{a}''_j$ we have

$$f(\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n) = f(\mathbf{a}_1, \ldots, \mathbf{a}'_j, \ldots, \mathbf{a}_n) + f(\mathbf{a}_1, \ldots, \mathbf{a}''_j, \ldots, \mathbf{a}_n),$$

and

$$f(\mathbf{a}_1, \ldots, c\mathbf{a}_j, \ldots, a_n) = cf(a_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_n).$$

(ii) If some two vectors coincide then $f = 0$.
(iii) $f(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$. This is a normalization condition; we choose the unit cube as the unit of volume.



Using (i) and (ii) we obtain

$$\begin{aligned} 0 &= f(\ldots,\mathbf{a}_k+\mathbf{a}_j,\ldots,\mathbf{a}_k+\mathbf{a}_j,\ldots) \\ &= f(\ldots,\mathbf{a}_k,\ldots,\mathbf{a}_k,\ldots)+f(\ldots,\mathbf{a}_k,\ldots,\mathbf{a}_j,\ldots) \\ &+ f(\ldots,\mathbf{a}_j,\ldots,\mathbf{a}_k\ldots)+f(\ldots,\mathbf{a}_j,\ldots,\mathbf{a}_j,\ldots) \\ &= f(\ldots,\mathbf{a}_k,\ldots,\mathbf{a}_j\ldots)+f(\ldots,\mathbf{a}_j,\ldots,\mathbf{a}_k\ldots). \end{aligned}$$

This shows that our function f should change sign when two vectors are interchanged. Of course, one expects a volume to be positive, so we should define the volume not as f but as $|f|$. Function f satisfying (i)-(iii) will be called the *oriented volume*.

We can represent every vector as a linear combination of of the standard basis vectors $\mathbf{e}_j$. So using linearily of f with respect to each column, we obtain

$$f(\mathbf{a}_1,\ldots,\mathbf{a}_n)=\sum_{j_1,\ldots,j_n} a_{1,j_1}\ldots a_{n,j_n} f(e_{j_1},e_{j_2},\ldots,e_{j_n}), \tag{1}$$

where summation is over *all* sequences $(j_1,\ldots,j_n)$ with entries from $\{1,\ldots,n\}$. Many of the terms of this sum are zero, namely those for which $j_k=j_\ell$ for some $k\neq\ell$. The remaining sequences $(j_1,\ldots,j_n)$ are *permutations* of the set $\{1,\ldots,n\}$. For a permutation $(j_1,\ldots,j_n)$, the function $f(j_1,\ldots,j_n)$ is either 1 or -1, according to the property (ii). So we discuss first of all how to determine the sign of our function f for each permutation.

1. **Permutations**. Suppose that $(j_1,j_2,\ldots,j_n)$ is a *permutation* of

$$(1,2,\ldots,n),$$

that is each j_k is one of the integers between 1 and n, and every such integer occurs exactly once. Every permutation can be obtained starting from

$(1, 2, \ldots, n)$ by consecutive interchanges of pairs of numbers. Such interchange of two numbers is called a *transposition*. For example, to obtain $(5, 3, 2, 4, 1)$ we start with $(1, 2, 3, 4, 5)$, and then

- interchange 1 and 5 to obtain $(5, 2, 3, 4, 1)$, and then
- interchange 2 and 3 to obtain $(5, 3, 2, 4, 1)$.

The number of required transpositions to obtain a given permutation may depend on the way we do it, but the *parity* of this number depends only on this given permutation. Thus a permutation is called *even* if an even number of transpositions is required, and *odd* otherwise. For example,

- the identity permutation $(1, 2, \ldots, n)$ is even (it is obtained using 0 transpositions),
- every transposition itself is odd,
- $(5, 3, 2, 4, 1)$ is even (because we obtained it above with two transpositions).

Here is the complete list of even permutations on 3 elements:

$$(1, 2, 3), (2, 3, 1), (3, 1, 2).$$

The remaining 3 are transpositions, so they are odd.

The total number of permutations of n elements is $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$, exactly half of them are even and the rest are odd.

In practice, to find the sign of permutation, one breaks it into cycles. Start with any element, and see where it goes under the permutation, then where this second element goes, and so on, until one returns to the original element. We obtain a cycle. Then start with some element which does not belong to this cycle, and obtain the second cycle, disjoint from the first one. Continue this procedure until one obtains a union of disjoint cycles. For example, $(2, 3, 1)$ consists of one cycle $1 \mapsto 3 \mapsto 2 \mapsto 1$, and $(2, 1, 3)$ consists of two cycles $1 \mapsto 2 \mapsto 1$ and $3 \mapsto 3$. The number of distinct elements of a cycle is the *length* of the cycle. Each cycle of length k is a product of $k - 1$ transpositions.

There is a second method of determining the sign of a permutation: one counts all pairs (j_k, j_ℓ) where $k < \ell$ but $j_k > j_\ell$. Such pairs are called inversions. The parity of a permutation is equal to the parity of the number of inversion. For example, $(2, 3, 1)$ has two inversions: $(2, 1)$ and $(3, 1)$.

2. **Definition of a determinant**. Formula (1) and discussion of signs of permutations suggest that the only function f (oriented volume) which has three properties (i), (ii), (iii) stated in the beginning must be defined as follows.

To each *square* matrix A (whose columns are $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ as above) corresponds a number, called *determinant* of A, and denoted by $|A|$ or $\det A$. For an $n \times n$ matrix $A = (a_{ij})$ we define

$$\det A = f(\mathbf{a}_1, \dots, a_n) = \sum_{(j_1,\dots,j_n)} \pm a_{1j_1} a_{2j_2} \dots a_{nj_n},$$

where the summation is over all permutations $(j_1, \dots, j_n)$ of $(1, 2, \dots, n)$, and the $+$ sign is chosen for each even permutation, and the $-$ sign for each odd permutation. For example,

if $n = 1$, then $\det A = a_{11}$,
if $n = 2$, then $\det A = a_{11}a_{22} - a_{12}a_{21}$,
if $n = 3$, then $\det A =$

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Thus there are $n!$ summands in the definition of determinant of a $n \times n$ matrix. Fortunately there are simpler ways to evaluate determinants then applying the definition directly.

It is easy to check that with this definition, the determinant indeed has all three properties of the oriented volume f.

3. **Why do we need them?** There are two main uses of determinants.

a) A matrix A is singular if and only if $\det A = 0$. Thus we have an analytic criterion for this important property of a matrix.

b) The volume of the parallelepiped generated by n vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ in $\mathbf{R}^n$ is equal to $|\det A|$, where $A = [\mathbf{x}_1, \dots, \mathbf{x}_n]$ is the matrix, whose columns are these vectors. (It should be clear what does it mean a "parallelogram, generated by two vectors in $\mathbf{R}^2$", and a "parallelepiped, generated by three vectors in $\mathbf{R}^3$". The general definition, not appealing to geometric intuition and thus applicable in any dimension, is "the set of all linear combinations $\sum c_j \mathbf{x}_j$, with $0 \le c_j \le 1$.)

4. **Properties of determinants**. Besides the three fundamental properties (i), (ii), (iii) in the definition, the determinant has these useful properties:

(iv) $\det A = \det A^T$. This follows from the fact that transposed permutation matrix is inverse to the original, and thus their parity is the same.

It follows that there are properties analogous to (i) and (ii) for the rows.

(v) Determinant is a linear function with respect to each row.

(vi) If two rows are equal, the determinant is zero. If two rows are interchanged, the determinant changes sign.

From (v) and (vi) follows:

(vii) If a multiple of a row is added to another row, the determinant does not change.

It follows from (iv) that operations on *columns* similar to (vii), (v), (vi) have similar effect to the row operations.

Finally we have this important properties: A is singular if and only if its determinant equals to zero. To see this consider the row echelon form of A whose determinant differs from that of A only by sign. The last property we state as a theorem:

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

Proof. If A is singular, then AB is singular, and both sides in this case are zero. If A is non-singular, consider the function

$$f(\mathbf{b}_1, \ldots, \mathbf{b}_n) = \frac{\det(AB)}{\det(A)},$$

Here $\mathbf{b}_j$ are the columns of B. This function has all properties (i), (ii), (iii) of the oriented volume. Therefore it must be equal to $\det(B)$. This completes the proof.

One can compute determinants using (v), (vii), or similar properties of column operations. For example

$$\begin{vmatrix} 1 & 3 \\ 2 & -1 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ 0 & -7 \end{vmatrix} = (-7)\begin{vmatrix} 1 & 3 \\ 0 & 1 \end{vmatrix} = (-7)\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = -7.$$

(If one arrives in the end to something else than the identity matrix, the determinant is zero.) One derives from (iv) that if some row consists entirely of zeros, then the determinant is zero.

5. **Row and column expansions**. Suppose that A is a $n \times n$ matrix. If we remove some $n-m$ rows and $n-m$ columns, where $m < n$, what remains is a new matrix of smaller size $m \times m$. Determinants of such matrices are called *minors* of order m of A.

The following special case is important. Take an element a_{ij} and remove the i-th row and j-th column from A. The minor we obtain is called the *minor of* a_{ij} and denoted by M_{ij}. (It is the determinant of some matrix of size $(n-1) \times (n-1)$.) Now the *cofactor of* a_{ij} is defined as $(-1)^{i+j} M_{ij}$. The following row expansion and column expansion formulas hold:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \quad \text{for every} \quad i = 1, \ldots, n,$$

and

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij} \quad \text{for every} \quad j = 1, \ldots, n.$$

Your book has plenty of examples illustrating these formulas. They are convenient to use for evaluating determinants of matrices which have many zeros in some row or in some column. For a generic matrix the method based on row operations is faster.

6. **More applications of determinants**. Determinants permit to write simple explicit formulas for solutions of the problems we considered before in this course. We already mentioned the criterion of singularity $\det A = 0$. Here is more:

a) The inverse B of a non-singular matrix A is given by

$$b_{ij} = \frac{1}{|A|} (-1)^{i+j} M_{ji}.$$

Notice that the indices of the minor are in the opposite order! In words: to compute the inverse, we first replace each element of a matrix A by its cofactor, then transpose the result, and then divide it by $\det A$. We can divide by $\det A$ exactly because A is non-singular.

b) Solution of a system of linear equations

$$A\mathbf{x} = \mathbf{b}$$

with a non-singular matrix A. We know that in this case a solution exists and it is unique. For $j = 1, \ldots, n$, let A_j be the matrix obtained by replacing the j-th column of A with the column $\mathbf{b}$, that is with the RHS of the equation. Then the coordinates of the solution vector $\mathbf{x}$ are given by

$$x_j = \frac{\det A_j}{\det A}.$$

Again, these formulas have sense, because we assume A to be non-singular, so $\det A \neq 0$. This formula is called the Cramer Rule.

Explicit formulas in a) and b) are rarely used in numerical computation because there are much better algorithms, for example, the row operations. Still we will see that in many cases it is important to have an explicit formula for the solution of a problem.