4 Determinant. Properties

Let me start with a system of two linear equation:

$$a_{11}x_1 + a_{12}x_2 = b_1,$$
$$a_{21}x_1 + a_{22}x_2 = b_2.$$

I multiply the first equation by a_{22}, second by a_{12} and subtract the second one from the first one. I get

$$x_1(a_{11}a_{22} - a_{21}a_{12}) = b_1a_{22} - b_2a_{12}.$$

Now I multiply the second equation by a_{11}, first by a_{21}, and subtract the first one from the second one:

$$x_2(a_{11}a_{22} - a_{21}a_{12}) = a_{11}b_2 - a_{21}b_1.$$

Assuming that $a_{11}a_{22} - a_{21}a_{12} \neq 0$ I get the solution

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}},$$
$$x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}.$$

The expressions in numerator and denominator have certain symmetry and I am tempted to introduce a definition for such objects. Namely, I define *the determinant* of a 2×2 matrix $\boldsymbol{A} = [a_{ij}]$ as

$$\det \boldsymbol{A} = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Then, if I also introduce the matrices

$$\boldsymbol{B}_1 = \begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix}, \qquad \boldsymbol{B}_2 = \begin{bmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{bmatrix},$$

then I can rewrite my solution in the following elegant form

$$x_1 = \frac{\det \boldsymbol{B}_1}{\det \boldsymbol{A}}, \quad x_2 = \frac{\det \boldsymbol{B}_2}{\det \boldsymbol{A}}.$$

It turns out that it is also possible to define the determinant for square matrices of an arbitrary size but any definition by a formula does not really look motivated (in part, due to the reasons that all the general formulas for the determinant of an arbitrary size square matrix look really complicated), and this served in part as a reason to abandon determinants from courses of linear algebra. On the other hand determinants are indispensable for many theoretical concepts in mathematics and basically unavoidable[1]. Hence, as a compromise, I will introduce the determinant by its properties and only after it I will present several (not extremely useful computationally) formulas for the determinant.

Math 329: *Intermediate Linear Algebra* by Artem Novozhilov©
e-mail: artem.novozhilov@ndsu.edu. Spring 2017

[1]It is impossible, for example, to learn how to make a change of variables in multiple integral without some basic understanding of determinants

The properties of the determinant are motivated by the fact that the determinant of a 2×2 matrix, how I defined it above, has a very simple geometric meaning.

Let $\boldsymbol{A} = [a_{ij}]_{2\times 2}$ and I introduce two column vectors

$$\boldsymbol{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}, \quad \boldsymbol{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix},$$

such that my matrix is $\boldsymbol{A} = [\boldsymbol{a}_1 \mid \boldsymbol{a}_2]$. These two vectors define a parallelogram on the plane, which has as its two sides these two vectors. The value $|\det \boldsymbol{A}| = |a_{11}a_{22} - a_{12}a_{21}|$ is actually the area of this parallelogram (see Fig. 4.1 for two examples).

Exercise 1. Can you prove the last claim?

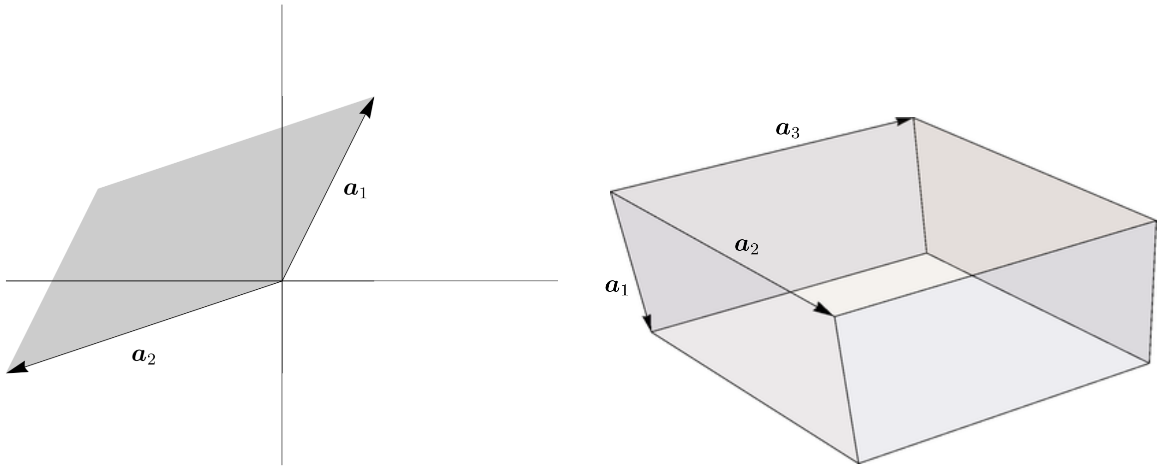

Figure 4.1: Geometric meaning of the determinant. $\mathbf{R}^2$ is on the left and $\mathbf{R}^3$ is on the right.

Therefore, I will think of the properties of a determinant of an arbitrary square matrix as the properties that should be in place for the volume of a parallelogram defined by the columns of this matrix (Fig. 4.1). I note that I allow negative values for the determinant and hence talk about the *signed* volume.

First, the determinant must be linear in each argument. This technically means that I postulate that (all the bold small letters denote column vectors with n components).

$$\det[\boldsymbol{a}_1 \mid \ldots \mid \alpha\boldsymbol{a}_k + \boldsymbol{b}_k \mid \ldots \mid \boldsymbol{a}_n] = \alpha \det[\boldsymbol{a}_1 \mid \ldots \mid \boldsymbol{a}_k \mid \ldots \mid \boldsymbol{a}_n] + \det[\boldsymbol{a}_1 \mid \ldots \mid \boldsymbol{b}_k \mid \ldots \mid \boldsymbol{a}_n]. \quad (4.1)$$

Here α is some scalar. The "volume justification" is that the usual formula tells me that the volume is equal to the area times height, and if my vector multiplied by a constant α then, intuitively, the height is also multiplied by the same constant; if a vector is the sum of two vectors then the corresponding height is also the sum of two heights.

Intuitively, the volume of a parallelogram, which is built on some vectors such that two of them coincide, should be zero (think, e.g., of a three dimensional case). That is

$$\det[\boldsymbol{a}_1 \mid \ldots \mid \alpha\boldsymbol{a}_k \mid \ldots \mid \boldsymbol{a}_k \mid \ldots \mid \boldsymbol{a}_n] = 0. \quad (4.2)$$

Note that the properties (4.1) and (4.2) together imply that

$$\det[\boldsymbol{a}_1 \mid \ldots \mid \boldsymbol{a}_j + \alpha \boldsymbol{a}_k \mid \ldots \mid \boldsymbol{a}_k \mid \ldots \boldsymbol{a}_n] = \det[\boldsymbol{a}_1 \mid \ldots \mid \boldsymbol{a}_j \mid \ldots \mid \boldsymbol{a}_k \mid \ldots \mid \boldsymbol{a}_n].$$

In words, an elementary column operation of the third type does not change the determinant.

Another important implication is that if I exchange the order of two columns in the matrix the determinant will change its sign. Here is a simple proof of this fact. Let me fix all the columns in the matrix except columns i and j and consider the determinant as a function of these two columns, which I will denote $f(\boldsymbol{a}, \boldsymbol{b})$. Now, using (4.1) and (4.2), I get

$$0 = f(\boldsymbol{a}+\boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b}) = f(\boldsymbol{a}, \boldsymbol{b}) + f(\boldsymbol{a}, \boldsymbol{b}) + f(\boldsymbol{b}, \boldsymbol{a}) + f(\boldsymbol{b}, \boldsymbol{b}) = f(\boldsymbol{a}, \boldsymbol{b}) + f(\boldsymbol{b}, \boldsymbol{a}).$$

And finally, I will need a normalization. Namely (and quite naturally), I postulate that

$$\det \boldsymbol{I} = 1. \tag{4.3}$$

It turns out that (4.1), (4.2), and (4.3) define a unique function on the collection of all possible square matrices. A proof that this function is unique will be given later, but now I will show that a lot of other properties can be deduced from my assumptions (4.1), (4.2), and (4.3) (which, by the way, are easily checked for the 2×2 case, I invite the reader to do these computations).

Proposition 4.1. *Let* $\boldsymbol{A}$ *be a square matrix.*

1. *If* $\boldsymbol{A}$ *has a zero column then* $\det \boldsymbol{A} = 0$.
2. *If* $\boldsymbol{A}$ *is a diagonal matrix then* $\det \boldsymbol{A} = a_{11} \ldots a_{nn}$.
3. *If* $\boldsymbol{A}$ *is a triangular matrix then* $\det \boldsymbol{A} = a_{11} \ldots a_{nn}$.
4. $\boldsymbol{A}$ *is invertible if and only if* $\det \boldsymbol{A} \neq 0$.
5. $\det(\alpha \boldsymbol{A}) = \alpha^n \det \boldsymbol{A}$.

Proof. 1. By (4.1) $\det[\boldsymbol{a}_1 \mid \ldots \mid \boldsymbol{0} \mid \ldots \mid \boldsymbol{a}_n] = 0 \det[\boldsymbol{a}_1 \mid \ldots \mid \boldsymbol{0} \mid \ldots \mid \boldsymbol{a}_n] = 0$.

2. Since for the diagonal matrix I can factor out all the diagonal elements and get the identity matrix, the conclusion follows.

3. If a triangular matrix has a zero on the main diagonal then, by the elementary column operations, I can always reduce it to a matrix with a zero column, and hence the determinant will be zero, due to Claim 1. If there are no zero elements on the main diagonal then by the third type column operations I can reduce my matrix to the diagonal one with the same diagonal elements. These operations does not change the determinant. Now applying already proved Claim 2 I get the desired result.

The reasonings above actually give a computational recipe to calculate the determinant. Namely, using the elementary column operations of the first and third type (if you are not very comfortable with column operations, do row operations on $\boldsymbol{A}^\top$ and then transpose the final result again) find the column echelon form, which is a triangular matrix, keeping track of the number of times you switched two columns, say, k times. The determinant of the triangular matrix is the product of the diagonal elements, and to get the determinant of the original matrix one just needs to multiply it by $(-1)^k$. By the way, recall that a square matrix is invertible if and only if its row (or, similarly, column) reduced echelon form is the identity matrix, and hence this computational procedure proves Claim 4.

The last point follows from (4.1). ∎

Exercise 2. Carefully write down a proof for claim 4 in the proposition.

Example 4.2. Compute

$$\det \begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix}.$$

Note that

$$\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & -3 \\ 2 & 3 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -3 \\ 3 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -3 \\ 0 & 2 & -6 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & -3 \\ 0 & 0 & -12 \end{bmatrix}.$$

The determinant of the last matrix is $1 \cdot (-1) \cdot (-12) = 12$. I recall that while I was row reducing the matrix, I switch columns once. Therefore the determinant of the original matrix is -12.

Corollary 4.3. *The system* $\boldsymbol{Ax} = \boldsymbol{b}$ *with a square matrix* $\boldsymbol{A}$ *has a unique solution if and only if* $\det \boldsymbol{A} \neq 0$

Lemma 4.4. *Let* $\boldsymbol{A}$ *be a square matrix and* $\boldsymbol{E}$ *be an elementary matrix of the same size. Then*

$$\det(\boldsymbol{AE}) = \det \boldsymbol{A} \det \boldsymbol{E}.$$

Proof. Since $\boldsymbol{E} = \boldsymbol{IE}$ and $\det \boldsymbol{I} = 1$, I know determinants of any elementary matrix. (If the previous sentence is not clear: Take, e.g., the elementary column operation of the first type; I know that the multiplication from the right by an elementary matrix of type 1 changes two columns, therefore, using the properties of the determinant, I conclude that the determinant of this elementary matrix must be -1; analogously for other elementary matrices.) Since the multiplication by an elementary matrix amounts to an elementary column operation, the conclusion follows. ■

Corollary 4.5. *For elementary matrices* $\boldsymbol{E}_1, \ldots, \boldsymbol{E}_k$

$$\det(\boldsymbol{AE}_1 \ldots \boldsymbol{E}_k) = \det \boldsymbol{A} \det \boldsymbol{E}_1 \ldots \det \boldsymbol{E}_k.$$

Theorem 4.6. *For square matrices* $\boldsymbol{A}$, $\boldsymbol{B}$*:*

1. $\det \boldsymbol{A} = \det \boldsymbol{A}^\top$.

2. $\det(\boldsymbol{AB}) = \det \boldsymbol{A} \det \boldsymbol{B}$.

Proof. 1. I will rely on two facts. First, for any elementary matrix $\boldsymbol{E}$ I have $\det \boldsymbol{E} = \det \boldsymbol{E}^\top$, which can be checked by direct calculations. Second, if $\boldsymbol{A}$ is not invertible then $\boldsymbol{A}^\top$ is also not invertible and both determinants are zero, hence it is sufficient to prove the claim only for the invertible matrices.

Let $\boldsymbol{A}$ be invertible, then it can be represented as a product of elementary matrices

$$\boldsymbol{A} = \boldsymbol{E}_1 \ldots \boldsymbol{E}_k.$$

Taking the transpose of both sides, using Corollary 4.5, and the fact that $\det \boldsymbol{E} = \det \boldsymbol{E}^\top$ conclude the proof.

2. If $\boldsymbol{B}$ is invertible, the proof follows from the fact that $\boldsymbol{B}$ can be represented as a product of elementary matrices and Corollary 4.5. If $\boldsymbol{B}$ is not invertible then the product $\boldsymbol{AB}$ is also not invertible (why?) and I get $0 = 0$. ■

Exercise 3. Prove that if $\boldsymbol{B}$ is not invertible then $\boldsymbol{AB}$ for any $\boldsymbol{A}$ is also not invertible.

Remark 4.7. The first point in the last theorem actually tells us that everything which was said about properties of determinants with respect to elementary column operations is true for elementary row operations.

Remark 4.8. Since the properties of determinants are so important, let me list them again, all together. Here $\boldsymbol{A}$ is a square matrix.

1. Determinant is linear in each row (column) when the other rows or columns are kept fixed.
2. $\det(\alpha \boldsymbol{A}) = \alpha^n \det \boldsymbol{A}$.
3. If two rows (columns) are switched the determinant changes its sign.
4. If there is a zero column or row then $\det \boldsymbol{A} = 0$.
5. For a triangular matrix $\det \boldsymbol{A}$ is equal to the product of the elements on the main diagonal. In particular, $\det \boldsymbol{I} = 1$.
6. $\det \boldsymbol{A} = 0$ if and only if $\boldsymbol{A}$ is not invertible (or, equivalently, $\det \boldsymbol{A} \neq 0$ if and only if $\boldsymbol{A}$ is invertible).
7. $\boldsymbol{Ax} = \boldsymbol{b}$ has a unique solution if and only if $\det \boldsymbol{A} \neq 0$.
8. $\det \boldsymbol{A}$ does not change is we perform elementary row or column operations of the third type (take a row or a column, multiply by a constant, and add to another row or column).
9. $\det \boldsymbol{A} = \det \boldsymbol{A}^\top$.
10. $\det(\boldsymbol{AB}) = \det \boldsymbol{A} \det \boldsymbol{B}$.
11. $\det \boldsymbol{A}^{-1} = (\det \boldsymbol{A})^{-1}$.

So, we know a lot properties of determinants, the row or column operations allow us to calculate the determinants. Do we need anything else? Actually, yes. We are still lacking a proof of uniqueness of a determinant (note that existence is actually guaranteed by the finite number of elementary operations required to put a given matrix into a reduced row or column echelon form). We still have no proof that if we perform the elementary operations in different orders we'll end up with the same answer. For this we need to show that our property (4.1)–(4.3) define a unique function, which would guarantee that the determinant is unique.