5 Determinant. Formal definition

In this section I first show that actually there are a lot of formulas that satisfy the three defining properties of determinant, and after it I will prove that the determinant is unique and therefore all these formulas in the end give the same answer.

5.1 Cofactor expansion

In the following theorem I provide $2n$ formulas that satisfy the three key properties of the determinant function.

Theorem 5.1. *Let $\boldsymbol{A}$ be an $n \times n$ matrix and $\boldsymbol{A}_{ij}$ denote the matrices that are obtained from $\boldsymbol{A}$ by deleting row i and column j. Then*

$$\det \boldsymbol{A} = \sum_{k=1}^{n} a_{j,k}(-1)^{j+k} \det \boldsymbol{A}_{j,k}, \quad j = 1, \ldots, n,$$

and

$$\det \boldsymbol{A} = \sum_{k=1}^{n} a_{k,j}(-1)^{j+k} \det \boldsymbol{A}_{k,j}, \quad j = 1, \ldots, n.$$

Remark 5.2. The first formula is called the expansion of the determinant through the j-th row, and the second one is the expansion of the determinant through the j-th column. Since I can always exchange two rows several times to put the required row on the top and keeping everything else intact, and since $\det \boldsymbol{A} = \det \boldsymbol{A}^\top$ then it is not difficult to show that all these expression lead to the same answer. I will leave it as an exercise.

Exercise 1. Using the properties of the determinant function, prove that all the formulas above lead to the same answer.

Proof. To make my notations simpler, I will work only with the first formula, very similar reasoning can be applied to the rest of them.

I need to show that the expression

$$\sum_{k=1}^{n} a_{1,k}(-1)^{1+k} \det \boldsymbol{A}_{1,k}$$

satisfies (1) the linearity with respect to any column, (2) equals to zero if there are two identical columns, and (3) gives 1 if $\boldsymbol{A} = \boldsymbol{I}$.

I will use induction to prove this theorem. If the student is not familiar with a proof by induction, they should wait for the next section and after it return to this theorem.

All three properties are true for $n = 2$, since my formula gives me in this case that

$$\det \boldsymbol{A} = a_{11}a_{22} - a_{21}a_{12}.$$

Now I assume that they hold for $n - 1$.

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

To show that the first property holds for n I need to check it for each term $a_{1,k} \det \boldsymbol{A}_{1,k}$. If I have that the i-th column is a sum of two vectors and $i \neq k$ then $a_{1,k}$ is a constant and $\det \boldsymbol{A}_{1,k}$ depends linearly on the i-th column of matrix $\boldsymbol{A}$ by the induction assumption, since it is a determinant of $n-1 \times n-1$ matrix. If $k = i$ then $\det \boldsymbol{A}_{1,k}$ is constant and $a_{1,k}$ depends linearly.

Assume now that columns i and j coincide in $\boldsymbol{A}$. Then the formula yields

$$\det \boldsymbol{A} = a_{1,i}(-1)^{1+i} \det \boldsymbol{A}_{1,i} + a_{1,j}(-1)^{1+j} \det \boldsymbol{A}_{1,j}$$

since in all other cases $\boldsymbol{A}_{1,k}$ will contain two identical columns and therefore are zero by the induction assumption. Assume, without loss of generality, that $i < j$. Then $\boldsymbol{A}_{1,i}$ can be obtained from $\boldsymbol{A}_{1,j}$ by switching columns i and $i+1$, then $i+1$ and $i+2$, etc, $j-2$ and $j-1$. Therefore, $\det \boldsymbol{A}_{1,j}(-1)^{j-i-1} = \det \boldsymbol{A}_{1,i}$ and hence $\det \boldsymbol{A} = 0$.

Finally, the third (normalization) property is clear: $\det \boldsymbol{I}_n = 1 \cdot \det \boldsymbol{I}_{n-1} = 1$. ∎

Definition 5.3. *The numbers*

$$C_{i,j} = (-1)^{i+j} \det \boldsymbol{A}_{i,j}$$

are called cofactors. Matrix $\boldsymbol{C} = [C_{i,j}]_{n\times n}$ is called the cofactor matrix of $\boldsymbol{A}$.

Theorem 5.4. *Let $\boldsymbol{A}$ be an invertible matrix and $\boldsymbol{C} = [C_{i,j}]$ be its cofactor matrix. Then*

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

Proof. Let me consider the product $\boldsymbol{A}\boldsymbol{C}^\top$. By the definition of matrix multiplication

$$(\boldsymbol{A}\boldsymbol{C}^\top)_{j,j} = a_{j,1}C_{j,1} + \ldots + a_{j,n}C_{j,n} = \det \boldsymbol{A}$$

for any j.

For the off-diagonal entries I have

$$(\boldsymbol{A}\boldsymbol{C}^\top)_{k,j} = a_{k,1}C_{j,1} + \ldots + a_{k,n}C_{j,n} = \det \boldsymbol{B},$$

where the matrix $\boldsymbol{B}$ has two identical rows (take row j and replace it with row k and after it use the expansion with respect to row j) and hence $\det \boldsymbol{B} = 0$. Therefore,

$$\boldsymbol{A}\boldsymbol{C}^\top = (\det \boldsymbol{A})\boldsymbol{I},$$

which means that $\boldsymbol{C}^\top / \det \boldsymbol{A}$ is a right inverse and hence is the inverse. ∎

Now I can prove the formulas for the solution of the linear system, which I used to motivate the appearance of determinants.

Corollary 5.5 (Cramer's rule)**.** *Consider the system of linear algebraic equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ with a square $\boldsymbol{A}$ and assume that $\det \boldsymbol{A} \neq 0$. Then each entry of the vector of solutions is given by*

$$x_k = \frac{\det \boldsymbol{B}_k}{\det \boldsymbol{A}},$$

where $\boldsymbol{B}_k$ are the matrices obtained from $\boldsymbol{A}$ by replacing the k-th column with the vector $\boldsymbol{b}$.

Proof. Since I have that in my case $\boldsymbol{A}$ is invertible hence

$$\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b} = \frac{1}{\det \boldsymbol{A}}\boldsymbol{C}^{\top}\boldsymbol{b},$$

from where the conclusion follows. ∎

Exercise 2. Expand on the proof above.

Exercise 3. For the $n \times n$ matrix

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & a_0 \\ -1 & 0 & 0 & \dots & 0 & a_1 \\ 0 & -1 & 0 & \dots & 0 & a_2 \\ \vdots & & & \ddots & & \\ 0 & 0 & 0 & \dots & 0 & a_{n-2} \\ 0 & 0 & 0 & \dots & -1 & a_{n-1} \end{bmatrix}$$

form the matrix $\boldsymbol{A} - \lambda \boldsymbol{I}$, where λ an arbitrary scalar. Compute $\det(\boldsymbol{A} - \lambda\boldsymbol{I})$. You should get a nice expression involving $a_0, a_1, \dots, a_{n-1}$ and λ. Row expansion and induction is probably the best way to go.

Answer: $(-1)^n\lambda^n + (-1)^{n-1}a_{n-1}\lambda^{n-1} + \dots - a_1\lambda + a_0$.

5.2 Uniqueness of the determinant

Before presenting the general computations and proofs, let me give you an example for a 2×2 matrix. Let $\boldsymbol{A} = [\boldsymbol{a}_1 \mid \boldsymbol{a}_2]$, where

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

Next, I will use the standard unit vectors

$$\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

to write

$$\boldsymbol{a}_1 = a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2,$$

and a similar expression for the second vector $\boldsymbol{a}_2$.

Now, using the properties of determinant,

$$\begin{aligned} \det \boldsymbol{A} = \det[\boldsymbol{a}_1 \mid \boldsymbol{a}_2] &= \det[a_{11}\boldsymbol{e}_1 + a_{21}\boldsymbol{e}_2 \mid a_{12}\boldsymbol{e}_1 + a_{22}\boldsymbol{e}_2] = \\ &= a_{11}a_{22}\det[\boldsymbol{e}_1 \mid \boldsymbol{e}_2] + a_{11}a_{12}\det[\boldsymbol{e}_1 \mid \boldsymbol{e}_1] + a_{21}a_{12}\det[\boldsymbol{e}_2 \mid \boldsymbol{e}_1] + a_{21}a_{22}\det[\boldsymbol{e}_2 \mid \boldsymbol{e}_2] = \\ &= a_{11}a_{22}\det[\boldsymbol{e}_1 \mid \boldsymbol{e}_2] + a_{21}a_{12}\det[\boldsymbol{e}_2 \mid \boldsymbol{e}_1]. \end{aligned}$$

The two determinants that are left are easy to calculate

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1, \quad \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1,$$

and therefore my computations show that no matter what, the determinant of a 2×2 matrix must be calculated as

$$\det \boldsymbol{A} = a_{11}a_{22} - a_{21}a_{12},$$

that is this formula is *equivalent* to the properties

(i) determinant is linear with respect to each column if the other columns are fixed;

(ii) determinant is equal to zero if we have two identical columns;

(iii) determinant changes its sign if two columns are exchanged (which is a consequence of the two previous properties);

(iv) finally, $\det \boldsymbol{I} = 1$.

Note that the second indices in each term in the formula above are $1, 2$, but the first indices are $1, 2$ for the first term and $2, 1$ in the second. If we have a sequence of n positive integers from 1 to n, with the property that no two of them are equal, then this sequence is called a *permutation of degree* n. For example for $n = 2$ there are only two permutations $(1, 2)$ and $(2, 1)$. For $n = 3$ we already have 6 different permutations. In general, there are $n! = 2 \cdot \ldots \cdot n$ permutations of degree n (prove it).

Now I switch to the general case.

I will consider matrix $\boldsymbol{A} = [a_{ij}]_{n\times n}$, which can be represented as the matrix composed of column-vectors $\boldsymbol{A} = [\boldsymbol{a}_1 \mid \ldots \mid \boldsymbol{a}_n]$. I introduce the standard unit vectors $\boldsymbol{e}_i$ as the column-vectors with all the entries equal to zero and only the i-th entry equal to 1. I clearly have $\boldsymbol{a}_j = \sum_{i=1}^{n} a_{ji}\boldsymbol{e}_i$, that is each column of my matrix can be represented as a linear combination of the standard unit vectors. First, I apply the linearity property (i) only to one column, say to the first one:

$$\begin{aligned}\det \boldsymbol{A} &= \det \left[\sum_{i=1}^{n} a_{i1}\boldsymbol{e}_1 \mid \ldots \mid \boldsymbol{a}_n\right] = \\ &= \sum_{i=1}^{n} \det\left[a_{i1}\boldsymbol{e}_i \mid \ldots \mid \boldsymbol{a}_n\right] = \\ &= \sum_{i=1}^{n} a_{i1} \det[\boldsymbol{e}_i \mid \ldots \mid \boldsymbol{a}_n].\end{aligned}$$

I can continue to do the same thing with the second column, but I will quickly understand that sooner or later the letters which I use to denote indices will be over, therefore I decide to use double index notation (I continue from the above with slight modification of indices):

$$
\begin{aligned}
&= \sum_{i_1=1}^{n} a_{i_1,1} \det \left[\boldsymbol{e}_{i_1} \mid \sum_{i_2=1}^{n} a_{i_2,2} \boldsymbol{e}_{i_2} \mid \ldots \mid \boldsymbol{a}_n \right] = \\
&= \sum_{i_1=1}^{n} a_{i_1,1} \sum_{i_2=1}^{n} \det \left[\boldsymbol{e}_{i_1} \mid a_{i_2,2} \boldsymbol{e}_{i_2} \mid \ldots \mid \boldsymbol{a}_n \right] = \\
&= \sum_{i_1=1}^{n} a_{i_1,1} \sum_{i_2=1}^{n} a_{i_2,2} \det \left[\boldsymbol{e}_{i_1} \mid \boldsymbol{e}_{i_2} \mid \ldots \mid \boldsymbol{a}_n \right] = \\
&= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} a_{i_1,1} a_{i_2,2} \det \left[\boldsymbol{e}_{i_1} \mid \boldsymbol{e}_{i_2} \mid \ldots \mid \boldsymbol{a}_n \right].
\end{aligned}
$$

In the last line I used the property of the summation, convince yourself that this is indeed true. Continue in a similar way, I end up with the ugly expression

$$
= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \ldots \sum_{i_n=1}^{n} a_{i_1,1} \ldots a_{i_n,2} \det \left[\boldsymbol{e}_{i_1} \mid \boldsymbol{e}_{i_2} \mid \ldots \mid \boldsymbol{e}_{i_n} \right],
$$

which has n^n terms. Note that I did not use anything except (i) yet.

Now it is time to use (ii): I note that most of the terms in the expression above have two identical columns (if you not convinced at this point, take a 3×3 matrix and perform necessary computations yourself, without looking into the notes), the only terms that are not zero are those that contain a permutation of the standard unit vectors in the right-most matrix, hence I can rewrite my expression in the simplified form

$$
= \sum_{\sigma \in \mathrm{Perm}}^{n} a_{\sigma(1),1} \ldots a_{\sigma(n),2} \det \left[\boldsymbol{e}_{\sigma(1)} \mid \ldots \mid \boldsymbol{e}_{\sigma(n)} \right],
$$

which now has only $n!$ terms. To understand the notations: I denote through Perm the set of all possible permutations of degree n, σ is the index that represent a particular permutation, and $\sigma(j)$ is the j-th integer in this permutation (e.g., if $\sigma = (1, 4, 3, 2)$ is a permutation of degree 4 then $\sigma(1) = 1, \sigma(2) = 4$, etc).

Using Property (iii), which is, as it was shown in class, is a consequence of (i) and (ii), I note that any matrix of the form $[\boldsymbol{e}_{\sigma(1)} \mid \ldots \mid \boldsymbol{e}_{\sigma(n)}]$ can be transformed into the identity matrix by a finite number on column exchanges, hence the determinant of this matrix is $(-1)^m$, where m is the number of column swaps required to go from the original matrix to the identity. Of course, one can pick different routs to the identity, but the key fact is that no matter how one chooses the columns to switch this number is either even, or odd. In the first case I say that the *sign* of the corresponding permutation is $+1$, and in the second one that it is -1 (the former permutation is even and the latter is odd).

Example 5.6. Consider a permutation $(1, 4, 3, 2)$. I can transform it into $(1, 2, 3, 4)$, e.g., as follows:

$$
(1, 4, 3, 2) \to (1, 2, 3, 4)
$$

or

$$
(1, 4, 3, 2) \to (1, 3, 4, 2) \to (1, 3, 4, 3) \to (1, 2, 3, 4).
$$

In both cases the number of changes is odd, and hence $\operatorname{sgn} \sigma = -1$ and this permutation is odd.

Proposition 5.7. *The number of steps to transform the permutation* $(1, 2, \dots, n)$ *into* $(\sigma(1), \dots, \sigma(n))$ *is either even or odd, and therefore* $\operatorname{sgn}\sigma$ *is well defined.*

Proof. Let me first give a different definition of the sign of a permutation. Namely, let $\operatorname{sgn}\sigma = (-1)^k$, where k is the number of *disorders* in σ. That is, number of situations when $\sigma(i) > \sigma(j)$ when $i < j$. For example, in $(1, 4, 3, 2)$ there are 3 disorders.

First I claim that if I exchange any two elements in a permutation its sign changes. It is obvious for two neighbor elements, for two arbitrary elements that are l elements apart, this can be done by $2l + 1$ neighbor exchanges, and hence the permutation changes sign. Also, $\operatorname{sgn}(1, \dots, n) = 1$, and therefore I proved that to transform $(1, \dots, n)$ into an even permutation (with sign equal to 1) I must have even number of steps, and into an odd permutation I must have odd number of steps, which concludes the proof. ∎

Therefore, now I have that

$$\det \boldsymbol{A} = \sum_{\sigma \in \operatorname{Perm} \mathrm{S}} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n} \det[\boldsymbol{e}_1 \mid \dots \mid \boldsymbol{e}_n], \tag{5.1}$$

which, after using (iv), gives me

Theorem 5.8. *Consider the function* det *on the set of all square matrices and assume that this function is linear with respect to each column, it equals zero if the matrix has two identical columns, and this function is equal to 1 if evaluated at the identity matrix, then this function is unique and its value can be calculated by* (5.2)*:*

$$\det \boldsymbol{A} = \sum_{\sigma \in \operatorname{Perm} \mathrm{S}} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \dots a_{\sigma(n),n}, \tag{5.2}$$

Exercise 4. Use (5.2) to write down the determinant of a 3×3 matrix.

Exercise 5. Give an estimate how many elementary operations, that is, additions and multiplications, you need to perform if you calculate the determinant by 1) row reduction, 2) summing through all possible permutations, formula (5.1), and 3) cofactor expansion.

It is interesting to note that I used (iv) only at the very end. Repeating the steps above, I immediately get

Theorem 5.9. *Let function G be defined on the set of all square matrices and satisfies properties (i) and (ii) above. Then for this function*

$$G(\boldsymbol{A}) = \det \boldsymbol{A}\, G(\boldsymbol{I}). \tag{5.3}$$

Exercise 6. Prove the theorem above.

I will need (5.3) in the following exercises.

Exercise 7. Consider block triangular matrix (this means that this matrix is composed of other matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{0}, *$ of a reasonable size)

$$\begin{bmatrix} \boldsymbol{A} & * \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix},$$

and the function

$$f(\boldsymbol{A}) = \det \begin{bmatrix} \boldsymbol{A} & * \\ \mathbf{0} & \boldsymbol{B} \end{bmatrix}.$$

(Here I fix $*$ and $\boldsymbol{B}$ and consider my expression as a function of just sub-block $\boldsymbol{A}$.)

Show that f satisfies (i) and (ii) and hence (5.3).

Solution: Let me check just property (ii), the first one is argued analogously. Assume that $\boldsymbol{A}$ has two identical columns. Then, since all the entries in the block matrix below these two columns of $\boldsymbol{A}$ are zero, the block-matrix also has two identical columns. I know that determinant of a matrix with two identical columns is 0, and hence I conclude that my function f satisfies (*ii*).

Exercise 8. Argue, using (i)–(iv), that

$$\det \begin{bmatrix} \boldsymbol{I} & * \\ \mathbf{0} & \boldsymbol{B} \end{bmatrix} = \det \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{bmatrix}.$$

Solution: Recall that from (i)–(iii) it follows that elementary column operations of the third type do not change matrix determinant. Since in each row of the upper half of the block matrix I have 1, I can always perform elementary column operations of the third type to guarantee that all other entries in these rows will be zero.

Exercise 9.

Show that the function

$$g(\boldsymbol{B}) = \det \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{bmatrix}$$

satisfies (i)–(iv) and hence $g(\boldsymbol{B}) = \det \boldsymbol{B}$.

Solution: Let me check only (iv): plug instead of $\boldsymbol{B}$ the identity matrix. The resulting matrix is

$$\det \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I} \end{bmatrix} = \det \boldsymbol{I},$$

where, clearly, the dimensions of $\boldsymbol{I}$ may be all different. I know that $\det \boldsymbol{I} = 1$ and hence $g(\boldsymbol{I}) = 1$, which means that g satisfies (iv).

Exercise 10.

Conclude, putting the exercises above together, that

$$\det \begin{bmatrix} \boldsymbol{A} & * \\ \mathbf{0} & \boldsymbol{B} \end{bmatrix} = \det \boldsymbol{A} \det \boldsymbol{B}.$$

Solution: I showed above that $f(\boldsymbol{A}) = \det \boldsymbol{A} f(\boldsymbol{I})$. I have (see above)

$$f(\boldsymbol{I}) = \det \begin{bmatrix} \boldsymbol{I} & * \\ \mathbf{0} & \boldsymbol{B} \end{bmatrix} = \det \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{B} \end{bmatrix} = \det \boldsymbol{B}.$$

Therefore,

$$\det \begin{bmatrix} \boldsymbol{A} & * \\ \mathbf{0} & \boldsymbol{B} \end{bmatrix} = \det \boldsymbol{A} \det \boldsymbol{B}.$$

Exercise 11. If you feel really comfortable working through the exercises in this subsection, give an alternative prove that

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

Hint: consider the following function

$$h(\boldsymbol{B}) = \det(\boldsymbol{AB})$$

and see which properties it satisfies.