3 Determinants

Definition

Let A be a *square* matrix of size n:

$$A = \begin{bmatrix} a_{11} & ... & a_{1n} \\ & ... & \\ a_{n1} & ... & a_{nn} \end{bmatrix}$$

Its **determinant** is a *scalar* $\det A$ defined by the formula

$$\det A = \sum_{\sigma} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} ... a_{n\sigma(n)}.$$

Here σ is a **permutation** of the indices $1, 2, ..., n$. A permutation σ can be considered as an invertible function $i \mapsto \sigma(i)$ from the set of n elements $\{1, ..., n\}$ to itself. We use the functional notation $\sigma(i)$ in order to specify the i-th term in the permutation $\sigma = \begin{pmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{pmatrix}$. Thus, each **elementary product** in the determinant formula contains exactly one matrix entry from each row, and these entries are chosen from n different columns. The sum is taken over all $n!$ ways of making such choices. The coefficient $\varepsilon(\sigma)$ in front of the elementary product equals 1 or -1 and is called the **sign** of the permutation σ.

We will explain the general rule of the signs after a few examples. In these examples, we begin using one more conventional notation for determinants. According to it, a square array of matrix entries placed between two vertical bars denotes the *determinant* of the matrix. Thus, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ denotes a *matrix*, but $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ denotes a *number* equal to the determinant of that matrix.

`Examples.` (1) For $n = 1$, the determinant $|a_{11}| = a_{11}$.

(2) For $n = 2$, we have: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$.

(3) For $n = 3$, we have $3! = 6$ summands

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$$

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

corresponding to permutations $\binom{123}{123}$, $\binom{123}{213}$, $\binom{123}{231}$, $\binom{123}{321}$, $\binom{123}{312}$, $\binom{123}{132}$.

The rule of signs for $n = 3$ is schematically shown on Figure 27.

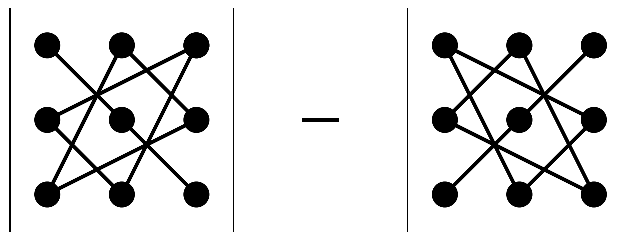

Figure 28

EXERCISES

201. Prove that the following determinant is equal to 0:

$$\begin{vmatrix} 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & c & d \\ 0 & 0 & 0 & e & f \\ p & q & r & s & t \\ v & w & x & y & z \end{vmatrix}. \qquad \lightning$$

202. Compute determinants:

$$\begin{vmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{vmatrix}, \begin{vmatrix} \cosh x & \sinh x \\ \sinh x & \cosh x \end{vmatrix}, \begin{vmatrix} \cos x & \sin y \\ \sin x & \cos y \end{vmatrix}. \qquad \checkmark$$

203. Compute determinants:

$$\begin{vmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix}, \begin{vmatrix} 0 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 6 \end{vmatrix}, \begin{vmatrix} 1 & i & 1+i \\ -i & 1 & 0 \\ 1-i & 0 & 1 \end{vmatrix}. \qquad \checkmark$$

Parity of Permutations

The general rule of signs relies on properties of permutations.

Let Δ_n denote the following polynomial in n variables $x_1, \ldots, x_n$:

$$\Delta_n(x_1, \ldots, x_n) = \prod_{1 \le i < j \le n} (x_i - x_j).$$

`Examples:` $\Delta_2 = x_1 - x_2$, $\Delta_3 = (x_1 - x_2)(x_1 - x_3)(x_2 - x_3)$. By definition, $\Delta_1 = 1$. In general, Δ_n is the product of all "n-choose-2" linear factors $x_i - x_j$ written in such a way that $i < j$.

Let σ be a permutation of $\{1, \ldots, n\}$. It acts on polynomials P in the variables $x_1, \ldots, x_n$ by permutation of the variables: $(\sigma P)(x_1, \ldots, x_n) := P(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$.

Example. Let $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$. Then

$$\sigma\Delta_3 = (x_3 - x_1)(x_3 - x_2)(x_1 - x_2) = (-1)^2(x_1 - x_3)(x_2 - x_3)(x_1 - x_2).$$

One says that σ **inverses** a pair of indices $i < j$ if $\sigma(i) > \sigma(j)$. The total number $l(\sigma)$ of pairs $i < j$ that σ inverses is called the **length** of the permutation σ. Thus, in the previous example, σ inverses the pairs $(1, 2)$ and $(1, 3)$, and has length $l(\sigma) = 2$.

Lemma. $\sigma\Delta_n = \varepsilon(\sigma)\Delta_n$, ***where*** $\varepsilon(\sigma) = (-1)^{l(\sigma)}$.

Proof. Indeed, a permutation of $\{1, \ldots, n\}$ also permutes all pairs $i \neq j$, and hence permutes all the linear factors in Δ_n. However, a factor $x_i - x_j$ is transformed into $x_{\sigma(i)} - x_{\sigma(j)}$, which occurs in the product Δ_n with the same sign whenever $\sigma(i) < \sigma(j)$, and with the opposite sign whenever $\sigma(i) > \sigma(j)$. Thus, $\sigma\Delta_n$ differs from Δ_n by the sign $(-1)^{l(\sigma)}$. $\square$

A permutation σ is called **even** or **odd** depending on the sign $\varepsilon(\sigma)$, i.e. when the length is even or odd respectively.

Examples. (1) The **identity permutation** id (defined by $\mathrm{id}(i) = i$ for all i) is even since $l(\mathrm{id}) = 0$.

(2) Consider a **transposition** τ, i.e. a permutation that swaps two indices, say $i < j$, leaving all other indices in their respective places. Then $\tau(j) < \tau(i)$, i.e. τ inverses the pair of indices $i < j$. Besides, for every index k such that $i < k < j$ we have: $\tau(j) < \tau(k) < \tau(i)$, i.e. both pairs $i < k$ and $k < j$ are inverted. Note that all other pairs of indices are not inverted by τ, and hence $l(\tau) = 2(j - i) + 1$. In particular, *every transposition is odd*: $\varepsilon(\tau) = -1$.

Proposition. ***Composition of two even or two odd permutations is even, and composition of one even and one odd permutation is odd:***

$$\varepsilon(\sigma\sigma') = \varepsilon(\sigma)\varepsilon(\sigma').$$

Proof. We have:

$$\varepsilon(\sigma\sigma')\Delta_n := (\sigma\sigma')\Delta_n = \sigma(\sigma'\Delta_n) = \varepsilon(\sigma')\sigma\Delta_n = \varepsilon(\sigma')\varepsilon(\sigma)\Delta_n.$$

Corollary 1. ***Inverse permutations have the same parity.***

`Corollary 2.` ***Whenever a permutation is written as the product of transpositions, the parity of the number of the transpositions in the product remains the same and coincides with the parity of the permutation: If*** $\sigma = \tau_1 \dots \tau_N$***, then*** $\varepsilon(\sigma) = (-1)^N$***.***

Here are some illustrations of the above properties in connection with the definition of determinants.

`Examples.` (3) The transposition (21) is odd. That is why the term $a_{12}a_{21}$ occurs in 2×2-determinants with the negative sign.

(4) The permutations $\binom{123}{123}$, $\binom{123}{213}$, $\binom{123}{231}$, $\binom{123}{321}$, $\binom{123}{312}$, $\binom{123}{132}$ have lengths $l = 0, 1, 2, 3, 2, 1$ and respectively signs $\varepsilon = 1, -1, 1, -1, 1, -1$ (thus explaining Figure 27). Notice that each next permutation here is obtained from the previous one by an extra flip.

(5) The permutation $\binom{1234}{4321}$ inverses all the 6 pairs of indices and has therefore length $l = 6$. Thus the elementary product $a_{14}a_{23}a_{32}a_{41}$ occurs with the sign $\varepsilon = (-1)^6 = +1$ in the definition of 4×4-determinants.

(6) Since inverse permutations have the same parity, the definition of determinants can be rewritten "by columns:"

$$\det A = \sum_\sigma \varepsilon(\sigma) a_{\sigma(1)1} ... a_{\sigma(n)n}.$$

Indeed, each summand in this formula is equal to the summand in the original definition corresponding to the permutation σ^{-1}, and *vice versa*. Namely, reordering the factors $a_{\sigma(1)1} ... a_{\sigma(n)n}$, so that $\sigma(1), \dots, \sigma(n)$ increase monotonically, yields $a_{1\sigma^{-1}(1)} ... a_{n\sigma^{-1}(n)}$.

EXERCISES

204. List all the 24 permutations of $\{1, 2, 3, 4\}$, find the length and the sign of each of them. ϟ

205. Find the length of the following permutation:

$$\left(\begin{array}{cccccccc} 1 & 2 & \dots & k & k+1 & k+2 & \dots & 2k \\ 1 & 3 & \dots & 2k-1 & 2 & 4 & \dots & 2k \end{array}\right). \qquad \checkmark$$

206. Find the maximal possible length of permutations of $\{1, ..., n\}$. ϟ

207. Find the length of a permutation $\left(\begin{array}{ccc} 1 & \dots & n \\ i_1 & \dots & i_n \end{array}\right)$ given the length l of the permutation $\left(\begin{array}{ccc} 1 & \dots & n \\ i_n & \dots & i_1 \end{array}\right)$. ✓

208. Prove that inverse permutations have the same length. ϟ

209. Compare parities of permutations of the letters *a,g,h,i,l,m,o,r,t* in the words *logarithm* and *algorithm.* ↯

210. Prove that the identity permutations are the only ones of length 0.

211. Find all permutations of length 1. ✓

212.*** Show that every permutation σ can be written as the product of $l(\sigma)$ transpositions of nearby indices. ↯

213.*** Represent the permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 3 & 2 \end{pmatrix}$ as composition of a minimal number of transpositions. ✓

214. Do products $a_{13}a_{24}a_{53}a_{41}a_{35}$ and $a_{21}a_{13}a_{34}a_{55}a_{42}$ occur in the defining formula for determinants of size 5? ✓

215. Find the signs of the elementary products $a_{23}a_{31}a_{42}a_{56}a_{14}a_{65}$ and $a_{32}a_{43}a_{14}a_{51}a_{66}a_{25}$ in the definition of determinants of size 6 by computing the numbers of inverted pairs of indices. ✓

Properties of determinants

(i) ***Transposed matrices have equal determinants:***

$$\det A^t = \det A.$$

This follows from the last Example. Below, we will think of an $n \times n$ matrix as an array $A = [\mathbf{a}_1, \ldots, \mathbf{a}_n]$ of its n columns of size n (vectors from $\mathbb{C}^n$ if you wish) and formulate all further properties of determinants in terms of columns. The same properties hold true for rows, since the transposition of A changes columns into rows without changing the determinant.

(ii) ***Interchanging any two columns changes the sign of the determinant:***

$$\det[..., \mathbf{a}_j, ..., \mathbf{a}_i, ...] = -\det[..., \mathbf{a}_i, ..., \mathbf{a}_j, ...].$$

Indeed, the operation replaces each permutation in the definition of determinants by its composition with the transposition of the indices i and j. Thus changes the parity of the permutation, and thus reverses the sign of each summand.

Rephrasing this property, one says that the determinant, considered as a function of n vectors $\mathbf{a}_1, \ldots, \mathbf{a}_n$ is **totally anti-symmetric**, i.e. changes the sign under every odd permutation of the vectors, and stays invariant under even. It implies that *a matrix with two equal columns has zero determinant.* It also allows one to formulate further column properties of determinants referring to the 1st column only, since the properties of all columns are alike.

(iii) ***Multiplication of a column by a number multiplies the determinant by this number:***

$$\det[\lambda \mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n] = \lambda \det[\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n].$$

Indeed, this operation simply multiplies each of the $n!$ elementary products by the factor of λ.

This property shows that *a matrix with a zero column has zero determinant.*

(iv) ***The determinant function is* additive *with respect to each column:***

$$\det[\mathbf{a}_1' + \mathbf{a}_1'', \mathbf{a}_2, ..., \mathbf{a}_n] = \det[\mathbf{a}_1', \mathbf{a}_2, ..., \mathbf{a}_n] + \det[\mathbf{a}_1'', \mathbf{a}_2, ..., \mathbf{a}_n].$$

Indeed, each elementary product contains exactly one factor picked from the 1-st column and thus splits into the sum of two elementary products $a'_{\sigma(1)1} a_{\sigma(2)2} ... a_{\sigma(n)n}$ and $a''_{\sigma(1)1} a_{\sigma(2)2} ... a_{\sigma(n)n}$. Summing up over all permutations yields the sum of two determinants on the right hand side of the formula.

The properties (*iv*) and (*iii*) together mean that *the determinant function is linear with respect to each column* separately. Together with the property (*ii*), they show that ***adding a multiple of one column to another one does not change the determinant of the matrix***. Indeed,

$$|\mathbf{a}_1 + \lambda \mathbf{a}_2, \mathbf{a}_2, ...| = |\mathbf{a}_1, \mathbf{a}_2, ...| + \lambda \, |\mathbf{a}_2, \mathbf{a}_2, ...| = |\mathbf{a}_1, \mathbf{a}_2, ...| \, ,$$

since the second summand has two equal columns.

The determinant function shares all the above properties with the identically zero function. The following property shows that these functions do not coincide.

(v) $\det I = 1$.

Indeed, since all off-diagonal entries of the identity matrix are zeroes, the only elementary product in the definition of $\det A$ that survives is $a_{11} ... a_{nn} = 1$.

The same argument shows that *the determinant of any diagonal matrix equals the product of the diagonal entries.* It is not hard to generalize the argument in order to see that the determinant of any upper or lower triangular matrix is equal to the product of the diagonal entries. One can also deduce this from the following factorization property valid for block triangular matrices.

Consider an $n \times n$-matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ subdivided into four **blocks** A, B, C, D of sizes $m \times m$, $m \times l$, $l \times m$ and $l \times l$ respectively (where of course $m + l = n$). We will call such a matrix **block triangular** if C or B is the zero matrix 0. We claim that

$$\det \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \det A \ \det D.$$

Indeed, consider a permutation σ of $\{1, ..., n\}$ which sends at least one of the indices $\{1, ..., m\}$ to the other part of the set, $\{m+1, ..., m+l\}$. Then σ must send at least one of $\{m+1, ..., m+l\}$ back to $\{1, ..., m\}$. This means that every elementary product in our $n \times n$-determinant which contains a factor from B must also contain a factor from C, and hence vanish, if $C = 0$. Thus only the permutations σ which permute $\{1, ..., m\}$ separately from $\{m+1, ..., m+l\}$ contribute to the determinant in question. Elementary products corresponding to such permutations factor into elementary products from $\det A$ and $\det D$ and eventually add up to the product $\det A \det D$.

Of course, the same holds true if $B = 0$ instead of $C = 0$.

We will use the factorization formula in the 1st proof of the following fundamental property of determinants.

EXERCISES

216. Compute the determinants

$$\begin{vmatrix} 13247 & 13347 \\ 28469 & 28569 \end{vmatrix}, \quad \begin{vmatrix} 246 & 427 & 327 \\ 1014 & 543 & 443 \\ -342 & 721 & 621 \end{vmatrix}. \qquad \checkmark$$

217. The numbers 195, 247, and 403 are divisible by 13. Prove that the following determinant is also divisible by 13: $\begin{vmatrix} 1 & 9 & 5 \\ 2 & 4 & 7 \\ 4 & 0 & 3 \end{vmatrix}$. ϟ

218. Professor Dumbel writes his office and home phone numbers as a 7×1-matrix O and 1×7-matrix H respectively. Help him compute $\det(OH)$. ✓

219. How does a determinant change if all its n columns are rewritten in the opposite order? ✓

220.* Solve the equation $\left| \begin{array}{ccccc} 1 & x & x^2 & ... & x^n \\ 1 & a_1 & a_1^2 & ... & a_1^n \\ 1 & a_2 & a_2^2 & ... & a_2^n \\ & & & ... & \\ 1 & a_n & a_n^2 & ... & a_n^n \end{array} \right| = 0$, where all $a_1, ..., a_n$ are given distinct numbers. ✓

221. Prove that an anti-symmetric matrix of size n has zero determinant if n is odd. ϟ

Multiplicativity

Theorem. ***The determinant is multiplicative with respect to matrix products: for arbitrary*** $n \times n$***-matrices*** A ***and*** B***,***

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

We give two proofs: one *ad hoc*, the other more conceptual.

Proof I. Consider the auxiliary $2n \times 2n$ matrix $\left[\begin{array}{cc} A & 0 \\ -I & B \end{array} \right]$ with the determinant equal to the product $(\det A)(\det B)$ according to the factorization formula. We begin to change the matrix by adding to the last n columns linear combinations of the first n columns with such coefficients that the submatrix B is eventually replaced by zero submatrix. Thus, in order to kill the entry b_{kj} we must add the b_{kj}-multiple of the k-th column to the $n+j$-th column. According to the properties of determinants (see (iv)) these operations do not change the determinant but transform the matrix to the form $\left[\begin{array}{cc} A & C \\ -I & 0 \end{array} \right]$. We ask the reader to check that the entry c_{ij} of the submatrix C in the upper right corner equals $a_{i1}b_{1j} + ... + a_{in}b_{nj}$ so that $C = AB$ is the matrix product! Now, interchanging the i-th and $n+i$-th columns, $i = 1, ..., n$, we change the determinant by the factor of $(-1)^n$ and transform the matrix to the form $\left[\begin{array}{cc} C & A \\ 0 & -I \end{array} \right]$. The factorization formula applies again and yields $\det C \det(-I)$. We conclude that $\det C = \det A \det B$ since $\det(-I) = (-1)^n$ compensates for the previous factor $(-1)^n$.□

Proof II. We will first show that the properties (i – v) completely characterize $\det[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ as a function of n columns $\mathbf{v}_i$ of size n.

Indeed, consider a function f, which to n columns $\mathbf{v}_1, \ldots, \mathbf{v}_n$, associates a number $f(\mathbf{v}_1, \ldots, \mathbf{v}_n)$. Suppose that f is *linear* with respect to each column. Let $\mathbf{e}_i$ denote the ith column of the identity matrix. Since $\mathbf{v}_1 = \sum_{i=1}^n v_{i1}\mathbf{e}_i$, we have:

$$f(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n) = \sum_{i=1}^n v_{i1} f(\mathbf{e}_i, \mathbf{v}_2, \ldots, \mathbf{v}_n).$$

Using linearity with respect to the 2nd column $\mathbf{v}_2 = \sum_{j=1}^n v_{j2}\mathbf{e}_j$, we similarly obtain:

$$f(\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n) = \sum_{i=1}^n \sum_{j=1}^n v_{i1} v_{j2} f(\mathbf{e}_i, \mathbf{e}_j, \mathbf{v}_3, \ldots, \mathbf{v}_n).$$

Proceeding the same way with all columns, we get:

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_n) = \sum_{i_1, \ldots, i_n} v_{i_1 1} \cdots v_{i_n n} f(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}).$$

Thus, f is determined by its values $f(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n})$ on strings of n basis vectors.

Let us assume now that f is *totally anti-symmetric*. Then, if any two of the indices $i_1, \ldots, i_n$ coincide, we have: $f(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) = 0$. All other coefficients correspond to *permutations* $\sigma = \begin{pmatrix} 1 & \ldots & n \\ i_1 & \ldots & i_n \end{pmatrix}$ of the indices $(1, \ldots, n)$, and hence satisfy:

$$f(\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_n}) = \varepsilon(\sigma) f(\mathbf{e}_1, \ldots, \mathbf{e}_n).$$

Therefore, we find:

$$\begin{aligned} f(\mathbf{v}_1, \ldots, \mathbf{v}_n) &= \sum_\sigma v_{\sigma(1)1} \ldots v_{\sigma(n)n} \varepsilon(\sigma) f(\mathbf{e}_1, \ldots, \mathbf{e}_n), \\ &= f(\mathbf{e}_1, \ldots, \mathbf{e}_n) \det[\mathbf{v}_1, \ldots, \mathbf{v}_n]. \end{aligned}$$

Thus, we have established:

`Proposition 1.` ***Every totally anti-symmetric function of*** n ***coordinate vectors of size*** n ***which is linear in each of them is proportional to the determinant function.***

Next, given an $n \times n$ matrix C, put

$$f(\mathbf{v}_1, \ldots, \mathbf{v}_n) := \det[C\mathbf{v}_1, \ldots, C\mathbf{v}_n].$$

Obviously, the function f is totally anti-symmetric in all $\mathbf{v}_i$ (since det is). Multiplication by C is linear:

$$C(\lambda\mathbf{u} + \mu\mathbf{v}) = \lambda C\mathbf{u} + \mu C\mathbf{v} \quad \text{for all } \mathbf{u}, \mathbf{v} \text{ and } \lambda, \mu.$$

Therefore, f is linear with respect to each $\mathbf{v}_i$ (as composition of two linear operations). By the previous result, f is proportional to det. Since $C\mathbf{e}_i$ are columns of C, we conclude that the coefficient of proportionality $f(\mathbf{e}_1, \dots, \mathbf{e}_n) = \det C$. Thus, we have found the following interpretation of $\det C$.

`Proposition 2.` $\det C$ ***is the factor by which the determinant function of*** n ***vectors*** $\mathbf{v}_i$ ***is multiplied when the vectors are replaced with*** $C\mathbf{v}_i$***.***

Now our theorem follows from the fact that when $C = AB$, the substitution $\mathbf{v} \mapsto C\mathbf{v}$ is the composition $\mathbf{v} \mapsto A\mathbf{v} \mapsto AB\mathbf{v}$ of consecutive substitutions defined by A and B. Under the action of A, the function det is multiplied by the factor $\det A$, then under the action of B by another factor $\det B$. But the resulting factor $(\det A)(\det B)$ must be equal to $\det C$. □

`Corollary.` ***If*** A ***is invertible, then*** $\det A$ ***is invertible.***

Indeed, $(\det A)(\det A^{-1}) = \det I = 1$, and hence $\det A^{-1}$ is reciprocal to $\det A$. The converse statement: that matrices with invertible determinants are invertible, is also true due to the explicit formula for the inverse matrix, described in the next section.

`Remark.` Of course, a real or complex number $\det A$ is invertible whenever $\det A \neq 0$. Yet over the integers $\mathbb{Z}$ this is not the case: the only invertible integers are ± 1. The above formulation, and several similar formulations that follow, which refer to invertibility of determinants, are preferable as they are more general.

EXERCISES

222. How do similarity transformations of a given matrix affect its determinant? ✓

223. Prove that the sign of the determinant of the coefficient marix of a real quadratic form does not depend on the coordinate system. ϟ

The Cofactor Theorem

In the determinant formula for an $n \times n$-matrix A each elementary product $\pm a_{1\sigma(1)} \dots$ begins with one of the entries $a_{11}, \dots, a_{1n}$ of the first row. The sum of all terms containing a_{11} in the 1-st place is the product of a_{11} with the determinant of the $(n-1) \times (n-1)$-matrix obtained from A by crossing out the 1-st row and the 1-st column. Similarly, the sum of all terms containing a_{12} in the 1-st place looks like the product of a_{12} with the determinant obtained by crossing out the 1-st row and the 2-nd column of A. In fact it differs by the factor of -1 from this product, since switching the columns 1 and 2 changes signs of all terms in the determinant formula and interchanges the roles of a_{11} and a_{12}. Proceeding in this way with $a_{13}, ..., a_{1n}$ we arrive at the **cofactor expansion** formula for $\det A$ which can be stated as follows.

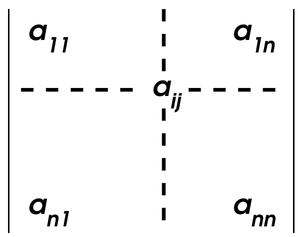

Figure 29

i \ j	1	2	3	4	5
1	+	−	+	−	+
2	−	+	−	+	−
3	+	−	+	−	+
4	−	+	−	+	−
5	+	−	+	−	+

Figure 30

The determinant of the $(n-1)\times(n-1)$-matrix obtained from A by crossing out the row i and column j is called the (ij)**-minor** of A (Figure 28). Denote it by M_{ij}. The (ij)**-cofactor** A_{ij} of the matrix A is the number that differs from the minor M_{ij} by a factor ± 1:

$$A_{ij} = (-1)^{i+j} M_{ij}.$$

The chess-board of the signs $(-1)^{i+j}$ is shown on Figure 29. With these notations, the cofactor expansion formula reads:

$$\det A = a_{11}A_{11} + a_{12}A_{12} + ... + a_{1n}A_{1n}.$$

`Example.`

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Using the properties (i) and (ii) of determinants we can adjust the cofactor expansion to the i-th row or j-th column:

$$\det A = a_{i1}A_{i1} + ... + a_{in}A_{in} = a_{1j}A_{1j} + ... + a_{nj}A_{nj},\ \ i, j = 1, ..., n.$$

These formulas reduce evaluation of $n \times n$-determinants to that of $(n-1) \times (n-1)$-determinants and can be useful in recursive computations.

Furthermore, we claim that applying the cofactor formula to the entries of the i-th row but picking the cofactors of another row we get the zero sum:

$$a_{i1}A_{j1} + ... + a_{in}A_{jn} = 0 \text{ if } i \neq j.$$

Indeed, construct a new matrix $\tilde{A}$ replacing the j-th row by a copy of the i-th row. This forgery does not change the cofactors $A_{j1}, ..., A_{jn}$ (since the j-th row is crossed out anyway) and yields the cofactor expansion $a_{i1}A_{j1} + ... + a_{in}A_{jn}$ for $\det \tilde{A}$. But $\tilde{A}$ has two identical rows and hence $\det \tilde{A} = 0$. The same arguments applied to the columns yield the dual statement:

$$a_{1i}A_{1j} + ... + a_{ni}A_{nj} = 0 \text{ if } i \neq j.$$

All the above formulas can be summarized in a single matrix identity. Introduce the $n \times n$-matrix $\mathrm{adj}(A)$, called **adjugate** to A, by placing the cofactor A_{ij} on the intersection of j-th row and i-th column. In other words, each a_{ij} is replaced with the corresponding cofactor A_{ij}, and then the resulting matrix is transposed:

$$\mathrm{adj} \left[\begin{array}{ccc} a_{11} & \dots & a_{1n} \\ \dots & a_{ij} & \dots \\ a_{n1} & \dots & a_{nn} \end{array} \right] = \left[\begin{array}{ccc} A_{11} & \dots & A_{n1} \\ \dots & A_{ji} & \dots \\ A_{1n} & \dots & A_{nn} \end{array} \right].$$

`Theorem.` $A \ \mathrm{adj}(A) = (\det A)\ I = \mathrm{adj}(A)\ A.$

`Corollary.` ***If*** $\det A$ ***is invertible then*** A ***is invertible, and***

$$A^{-1} = \frac{1}{\det A} \mathrm{adj}(A).$$

`Example.` If $ad - bc \neq 0$, then $\left[\begin{array}{cc} a & b \\ c & d \end{array} \right]^{-1} = \frac{1}{ad-bc} \left[\begin{array}{cc} d & -b \\ -c & a \end{array} \right].$

EXERCISES

224. Prove that the adjugate matrix of an upper (lower) triangular matrix is upper (lower) triangular.

225. Which triangular matrices are invertible?

226. Compute the determinants: ($*$ is a wild card):

$$(a) \left| \begin{array}{cccc} * & * & * & a_n \\ * & * & \dots & 0 \\ * & a_2 & 0 & \dots \\ a_1 & 0 & \dots & 0 \end{array} \right|, \quad (b) \left| \begin{array}{cccc} * & * & a & b \\ * & * & c & d \\ e & f & 0 & 0 \\ g & h & 0 & 0 \end{array} \right|. \qquad \checkmark$$

227. Compute determinants using cofactor expansions:

$$(a)\ \begin{vmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 0 & 2 \\ 2 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 \end{vmatrix}, \quad (b)\ \begin{vmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{vmatrix}. \qquad \checkmark$$

228. Compute inverses of matrices using the Cofactor Theorem:

$$(a)\ \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix}, \quad (b)\ \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \qquad \checkmark$$

229. Solve the systems of linear equations $A\mathbf{x} = \mathbf{b}$ where A is one of the matrices of the previous exercise, and $\mathbf{b} = [1, 0, 1]^t$. ✓

230. Compute

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1}.$$

231. Express $\det(\operatorname{adj}(A))$ of the adjugate matrix via $\det A$. ✓

232. Which integer matrices have integer inverses? ✓

Cramer's Rule

This is an application of the Cofactor Theorem to systems of linear equations. Consider a system

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\cdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= b_n \end{aligned}$$

of n linear equations with n unknowns $(x_1, \ldots, x_n)$. It can be written in the matrix form

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

where A is the $n \times n$-matrix of the coefficients a_{ij}, $\mathbf{b} = [b_1, \ldots, b_n]^t$ is the column of the right hand sides, and $\mathbf{x}$ is the column of unknowns. In the following Corollary, $\mathbf{a}_i$ denote columns of A.

`Corollary.` ***If*** $\det A$ ***is invertible then the system of linear equations*** $A\mathbf{x} = \mathbf{b}$ ***has a unique solution given by the***

formulas:

$$x_1 = \frac{\det[\mathbf{b}, \mathbf{a}_2, ..., \mathbf{a}_n]}{\det[\mathbf{a}_1, ..., \mathbf{a}_n]}, \ \ldots \ , x_n = \frac{\det[\mathbf{a}_1, ..., \mathbf{a}_{n-1}, \mathbf{b}]}{\det[\mathbf{a}_1, ..., \mathbf{a}_n]}.$$

Indeed, when $\det A \neq 0$, the matrix A is invertible. Multiplying the matrix equation $A\mathbf{x} = \mathbf{b}$ by A^{-1} on the left, we find: $\mathbf{x} = A^{-1}\mathbf{b}$. Thus the solution is unique, and $x_i = (\det A)^{-1}(A_{1i}b_1 + ... + A_{ni}b_n)$ according to the cofactor formula for the inverse matrix. But the sum $b_1A_{1i} + ... + b_nA_{ni}$ is the cofactor expansion for $\det[\mathbf{a}_1, ..., \mathbf{a}_{i-1}, \mathbf{b}, \mathbf{a}_{i+1}, ..., \mathbf{a}_n]$ with respect to the i-th column.

`Example.` Suppose that $a_{11}a_{22} \neq a_{12}a_{21}$. Then the system

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_2 + a_{22}x_2 &= b_2 \end{aligned}$$

has a unique solution

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

EXERCISES

233. Solve systems of equations using Cramer's rule:

$$(a) \begin{array}{rcr} 2x_1 - x_2 - x_3 &=& 4 \\ 3x_1 + 4x_2 - 2x_3 &=& 11 \\ 3x_1 - 2x_2 + 4x_3 &=& 11 \end{array}, \quad (b) \begin{array}{rcr} x_1 + 2x_2 + 4x_3 &=& 31 \\ 5x_1 + x_2 + 2x_3 &=& 29 \\ 3x_1 - x_2 + x_3 &=& 10 \end{array}. \qquad \checkmark$$

Three Cool Formulas

We collect here some useful generalizations of previous results.

A. We don't know of any reasonable generalization of determinants to the situation when matrix entries do *not* commute. However the following generalization of the formula $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$ is instrumental in some non-commutative applications.[10]

[10]Notably in the definition of *Berezinian* in super-mathematics [7].

In the block matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$***, assume that*** D^{-1} ***exists. Then*** $\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det(A - BD^{-1}C) \det D$***.***

`Proof:` $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix} = \begin{bmatrix} A - BD^{-1}C & B \\ 0 & D \end{bmatrix}$.

B. Laplace's formula[11] below generalizes cofactor expansions.

By a **multi-index** I of length $|I| = k$ we mean an increasing sequence $i_1 < \cdots < i_k$ of k indices from the set $\{1, \ldots, n\}$. Given and $n \times n$-matrix A and two multi-indices I, J of the same length k, we define the (IJ)**-minor** of A as the determinant of the $k \times k$-matrix formed by the entries $a_{\mathbf{i}_\alpha j_\beta}$ of A located at the intersections of the rows $i_1, \ldots, i_k$ with columns $j_1, \ldots, j_k$ (see Figure 30). Also, denote by $\bar{I}$ the multi-index **complementary** to I, i.e. formed by those $n - k$ indices from $\{1, \ldots, n\}$ which are *not* contained in I.

For each multi-index $I = (i_1, \ldots, i_k)$***, the following*** **cofactor expansion with respect to rows** $i_1, \ldots, i_k$ **holds true:**

$$\det A = \sum_{J:|J|=k} (-1)^{i_1+\cdots+i_k+j_1+\cdots+j_k} M_{IJ} M_{\bar{I}\bar{J}},$$

where the sum is taken over all multi-indices $J = (j_1, \ldots, j_k)$ ***of length*** k***.***

Similarly, one can similarly write Laplace's cofactor expansion formula with respect to given k columns.

`Example.` Let $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ and $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4$ be 8 vectors on the plane. Then $\begin{vmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 \\ \mathbf{b}_1 & \mathbf{b}_2 & \mathbf{b}_3 & \mathbf{b}_4 \end{vmatrix} = |\mathbf{a}_1\ \mathbf{a}_2||\mathbf{b}_3\ \mathbf{b}_4| - |\mathbf{a}_1\ \mathbf{a}_3||\mathbf{b}_2\ \mathbf{b}_4|$ $+ |\mathbf{a}_1\ \mathbf{a}_4||\mathbf{b}_2\ \mathbf{b}_3| + |\mathbf{a}_2\ \mathbf{a}_3||\mathbf{b}_1\ \mathbf{b}_4| - |\mathbf{a}_2\ \mathbf{a}_4||\mathbf{b}_1\ \mathbf{b}_3| + |\mathbf{a}_3\ \mathbf{a}_4||\mathbf{b}_1\ \mathbf{b}_2|$.

[11] After Pierre-Simon **Laplace** (1749–1827).

In the proof of Laplace's formula, it suffices to assume that it is written with respect to the *first* k rows, i.e. that $I = (1, \dots, k)$. Indeed, interchanging them with the rows $i_1 < \dots < i_k$ takes $(i_1 - 1) + (i_2 - 2) + \dots + (i_k - k)$ transpositions, which is accounted for by the sign $(-1)^{i_1+\dots+i_k}$ in the formula.

Next, multiplying out $M_{IJ}M_{\bar{I}\bar{J}}$, we find $k!(n-k)!$ elementary products of the form:

$$\pm a_{1,j_{\alpha_1}} \cdots a_{k,j_{\alpha_k}} \, a_{k+1,\bar{j}_{\beta_1}} \cdots a_{n,\bar{j}_{\beta_{n-k}}},$$

where $\alpha = \begin{pmatrix} 1 & \dots & k \\ \alpha_1 & \dots & \alpha_k \end{pmatrix}$ and $\beta = \begin{pmatrix} 1 & \dots & n-k \\ \beta_1 & \dots & \beta_{n-k} \end{pmatrix}$ are permutations, and $j_{\alpha_\mu} \in J$, $\bar{j}_{\beta_\nu} \in \bar{J}$. It is clear that the total sum over multi-indices I contains each elementary product from $\det A$, and does it exactly once. Thus, to finish the proof, we need to compare the signs.

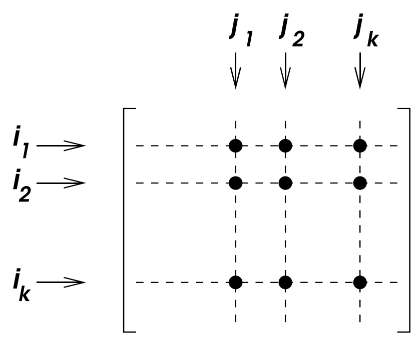

Figure 31

The sign $\pm$ in the above formula is equal to $\varepsilon(\alpha)\varepsilon(\beta)$, the product of the signs of the permutations α and β. The sign of this elementary product in the definition of $\det A$ is equal to the sign of the permutation $\begin{pmatrix} 1 & \dots & k & k+1 & \dots & n \\ j_{\alpha_1} & \dots & j_{\alpha_k} & \bar{j}_{\beta_1} & \dots & \bar{j}_{\beta_{n-k}} \end{pmatrix}$ on the set $J \cup \bar{J} = \{1, \dots, n\}$. Reordering separately the first k and last $n-k$ indices in the increasing order changes the sign of the permutation by $\varepsilon(\alpha)\varepsilon(\beta)$. Therefore the signs of all summands of $\det A$ which occur in $M_{IJ}M_{\bar{I}\bar{J}}$ are *coherent*. It remains to find the total sign with which $M_{IJ}M_{\bar{I}\bar{J}}$ occurs in $\det A$, by computing the sign of the permutation $\sigma := \begin{pmatrix} 1 & \dots & k & k+1 & \dots & n \\ j_1 & \dots & j_k & \bar{j}_1 & \dots & \bar{j}_{n-k} \end{pmatrix}$, where $j_1 < \cdots j_k$ and $\bar{j}_1 < \dots < \bar{j}_{n-k}$.

Starting with the identity permutation $(1, 2 \dots, j_1, \dots, j_2, \dots, n)$, it takes $j_1 - 1$ transpositions of nearby indices to move j_1 to the 1st place. Then it takes $j_2 - 2$ such transpositions to move j_2 to the 2nd place. Continuing this way, we find that

$$\varepsilon(\sigma) = (-1)^{(j_1-1)+\cdots+(j_k-k)} = (-1)^{1+\cdots+k+j_1+\cdots+\mathbf{j}_k}.$$

This agrees with Laplace's formula, since $I = \{1, \dots, k\}$. $\square$.

C. Let A and B be $k \times n$ and $n \times k$ matrices (think of $k < n$). For each multi-index $I = (i_1, \dots, i_k)$, denote by A_I and B_I the $k \times k$-matrices formed by respectively: columns of A and rows of B with the indices $i_1, \dots, i_k$.

The determinant of the $k \times k$***-matrix*** AB ***is given by the following*** **Binet–Cauchy formula**:[12]

$$\det AB = \sum_I (\det A_I)(\det B_I).$$

Note that when $k = n$, this turns into the multiplicative property of determinants: $\det(AB) = (\det A)(\det B)$. Our second proof of it can be generalized to establish the formula of Binet–Cauchy. Namely, let $\mathbf{a}_1, \dots, \mathbf{a}_n$ denote columns of A. Then the jth column of $C = AB$ is the linear combination: $\mathbf{c}_j = \mathbf{a}_1 b_{1j} + \cdots + \mathbf{a}_n b_{nj}$. Using linearity in each $\mathbf{c}_j$, we find:

$$\det[\mathbf{c}_1, \dots, \mathbf{c}_k] = \sum_{1 \le i_1, \dots, i_k \le n} \det[\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_k}] b_{i_1 1} \cdots b_{i_k k}.$$

If any two of the indices i_α coincide, $\det[\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_k}] = 0$. Thus the sum is effectively taken over all *permutations* $\begin{pmatrix} 1 & \dots & k \\ i_1 & \dots & i_k \end{pmatrix}$ on the *set*[13] $\{i_1, \dots, i_k\}$. Reordering the columns $\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_k}$ in the increasing order of the indices (and paying the "fees" ± 1 according to parities of permutations) we obtain the sum over all multi-indices of length k:

$$\sum_{i'_1 < \cdots < i'_k} \det[\mathbf{a}_{i'_1}, \dots, \mathbf{a}_{i'_k}] \sum_\sigma \varepsilon(\sigma) b_{i_1 1} \cdots b_{i_k k}.$$

The sum on the right is taken over permutations $\sigma = \begin{pmatrix} i'_1 & \dots & i'_k \\ i_1 & \dots & i_k \end{pmatrix}$. It is equal to $\det B_I$, where $I = (i'_1, \dots, i'_k)$. $\square$

`Corollary 1.` *If* $k > n$, $\det AB = 0$.

[12] After Jacques **Binet** (1786–1856) and Augustin Louis **Cauchy** (1789–1857).
[13] Remember that in a set, elements are unordered!

This is because no multi-indices of length $k > n$ can be formed from $\{1, \dots, n\}$. In the oppositely extreme case $k = 1$, Binet–Cauchy's formula turns into the expression $\mathbf{u}^t\mathbf{v} = \sum u_i v_i$ for the dot product of coordinate vectors. A "Pythagorean" interpretation of the following identity will come to light in the next chapter, in connection with volumes of parallelepipeds.

`Corollary 2.` $\det AA^t = \sum_I (\det A_I)^2$.

EXERCISES

234.* Compute determinants:

$$(a)\ \begin{vmatrix} 0 & x_1 & x_2 & \dots & x_n \\ x_1 & 1 & 0 & \dots & 0 \\ x_2 & 0 & 1 & \dots & 0 \\ . & . & . & . & . \\ x_n & 0 & \dots & 0 & 1 \end{vmatrix}, \quad (b)\ \begin{vmatrix} a & 0 & 0 & 0 & 0 & b \\ 0 & a & 0 & 0 & b & 0 \\ 0 & 0 & a & b & 0 & 0 \\ 0 & 0 & c & d & 0 & 0 \\ 0 & c & 0 & 0 & d & 0 \\ c & 0 & 0 & 0 & 0 & d \end{vmatrix} \qquad ϟ\ ✓.$$

235.* Let P_{ij}, $1 \leq i < j \leq 4$, denote the 2×2-minor of a 2×4-matrix formed by the columns i and j. Prove the following **Plücker identity**[14]

$$P_{12}P_{34} - P_{13}P_{24} + P_{14}P_{23} = 0. \qquad ✓$$

236. The **cross product** of two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ is defined by

$$\mathbf{x} \times \mathbf{y} := \left(\begin{vmatrix} x_2 & x_3 \\ y_2 & y_3 \end{vmatrix}, \begin{vmatrix} x_3 & x_1 \\ y_3 & y_1 \end{vmatrix}, \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} \right).$$

Prove that the length $|\mathbf{x} \times \mathbf{y}| = \sqrt{|\mathbf{x}|^2|\mathbf{y}|^2 - \langle \mathbf{x}, \mathbf{y} \rangle^2}$. ϟ

237.* Prove that $a_n + \cfrac{1}{a_{n-1} + \cfrac{1}{\dots + \cfrac{1}{a_1 + \cfrac{1}{a_0}}}} = \dfrac{\Delta_n}{\Delta_{n-1}}$,

where $\Delta_n = \begin{vmatrix} a_0 & 1 & 0 & \dots & 0 \\ -1 & a_1 & 1 & \dots & 0 \\ . & . & . & . & . \\ 0 & \dots & -1 & a_{n-1} & 1 \\ 0 & \dots & 0 & -1 & a_n \end{vmatrix}$. ϟ

238.* Compute: $\begin{vmatrix} \lambda & -1 & 0 & \dots & 0 \\ 0 & \lambda & -1 & \dots & 0 \\ . & . & . & . & . \\ 0 & \dots & 0 & \lambda & -1 \\ a_n & a_{n-1} & \dots & a_2 & \lambda + a_1 \end{vmatrix}$. ✓

[14] After Julius **Plücker** (1801–1868).

239.* Compute: $\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \binom{2}{1} & \binom{3}{1} & \dots & \binom{n}{1} \\ 1 & \binom{3}{2} & \binom{4}{2} & \dots & \binom{n+1}{2} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & \binom{n}{n-1} & \binom{n+1}{n-1} & \dots & \binom{2n-2}{n-1} \end{vmatrix}$. ϟ ✓

240.* Prove **Vandermonde's identity**[15]

$$\begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{vmatrix} = \prod_{1 \le i < j \le n} (x_j - x_i).$$ ϟ

241.* Compute: $\begin{vmatrix} 1 & 2 & 3 & \dots & n \\ 1 & 2^3 & 3^3 & \dots & n^3 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 2^{2n-1} & 3^{2n-1} & \dots & n^{2n-1} \end{vmatrix}$. ϟ ✓

[15] After Alexandre-Theóphile **Vandermonde** (1735–1796).