11

More Determinants

Historically, determinants were defined to solve linear systems. In this chapter, we explain the connection between determinants and linear systems, and relate determinants to permutations.

11.0 The Cramer Rule

HOW CAN WE USE DETERMINANTS TO SOLVE LINEAR SYSTEMS? Whenever a linear system has a unique solution, there is an explicit formula for this solution involving determinants.

11.0.0 Proposition (Laplace expansion)**.** *Let n and i be integers such that $1 \leqslant i \leqslant n$. The determinant of an $(n \times n)$-matrix $\mathbf{A}$ is*

$$\det(\mathbf{A}) = \sum_{k=1}^{n} (-1)^{i+k} \, a_{i,k} \, \det\big(\mathbf{A}(\hat{\imath}, \hat{k})\big) = \sum_{j=1}^{n} (-1)^{j+i} \, a_{j,i} \, \det\big(\mathbf{A}(\hat{\jmath}, \hat{\imath})\big) \,,$$

where $a_{j,k}$ denotes the (j,k)-entry in the matrix $\mathbf{A}$, for all $1 \leqslant j \leqslant n$ and all $1 \leqslant k \leqslant n$, and $\mathbf{A}(\hat{\jmath}, \hat{k})$ is the submatrix of $\mathbf{A}$ obtained by deleting j-th row and k-th column.

When $i = 1$, we recover the definition of the determinant [10.0.1].

Proof. The first equation can be obtain from the definition of the determinant [10.0.1] by using $i-1$ adjacent row swaps to move the i-th row into the 1-st row. Since an adjacent row swap changes the sign of the determinant, it follows that the sign of the determinant changes by $(-1)^{i-1}$. Since the determinant is invariant under the transpose [10.2.2], the second equation follows from the first. □

$$\begin{bmatrix} + & - & + & - & + & - & \cdots \\ - & + & - & + & - & + & \cdots \\ + & - & + & - & + & - & \cdots \\ - & + & - & + & - & + & \cdots \\ + & - & + & - & + & - & \cdots \\ - & + & - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Figure 11.0: Sign pattern in the Laplace expansion and the adjugate

11.0.1 Definition. The ***adjugate*** of a square matrix $\mathbf{A}$, denoted $\operatorname{adj}(\mathbf{A})$, is the matrix whose (j,k)-entry is $(-1)^{j+k} \det\big(\mathbf{A}(\hat{k}, \hat{\jmath})\big)$ where $\mathbf{A}(\hat{k}, \hat{\jmath})$ is the submatrix obtained by deleting the k-th row and j-th column in the matrix $\mathbf{A}$.

The row and columns indices are interchanged.

11.0.2 Problem. Compute the adjugate of $\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$.

Solution. We have

$$\operatorname{adj}\left(\begin{bmatrix} 1 & 1 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix}\right) = \begin{bmatrix} \det\left(\left[\begin{smallmatrix} 2 & 1 \\ 0 & 2 \end{smallmatrix}\right]\right) & (-1)\det\left(\left[\begin{smallmatrix} 0 & 1 \\ 1 & 2 \end{smallmatrix}\right]\right) & \det\left(\left[\begin{smallmatrix} 0 & 2 \\ 1 & 0 \end{smallmatrix}\right]\right) \\ (-1)\det\left(\left[\begin{smallmatrix} 1 & 2 \\ 0 & 2 \end{smallmatrix}\right]\right) & \det\left(\left[\begin{smallmatrix} 1 & 2 \\ 1 & 2 \end{smallmatrix}\right]\right) & (-1)\det\left(\left[\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right]\right) \\ \det\left(\left[\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right]\right) & (-1)\det\left(\left[\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right]\right) & \det\left(\left[\begin{smallmatrix} 1 & 1 \\ 0 & 2 \end{smallmatrix}\right]\right) \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 4 & 1 & -2 \\ -2 & 0 & 1 \\ -3 & -1 & 2 \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} 4 & -2 & -3 \\ 1 & 0 & -1 \\ -2 & 1 & 2 \end{bmatrix} ,$$

and $\operatorname{adj}\left(\begin{bmatrix} a & c \\ b & d \end{bmatrix}\right) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix}$. □

11.0.3 Theorem (Adjugate equation)**.** *For any square matrix* $\mathbf{A}$*, we have*

$$\operatorname{adj}(\mathbf{A})\,\mathbf{A} = \det(\mathbf{A})\,\mathbf{I}\,.$$

Proof. For all $1 \leqslant j \leqslant n$ and all $1 \leqslant k \leqslant n$, let $a_{j,k} \in \mathbb{K}$ denote the (j,k)-entry in the matrix $\mathbf{A}$. Combining the definitions of the matrix multiplication and the adjugate, we see that the (i,k)-entry in matrix product $\operatorname{adj}(\mathbf{A})\,\mathbf{A}$ equals

$$\sum_{j=1}^{n} (-1)^{j+i} \det\big(\mathbf{A}(\hat{\jmath},\hat{\imath})\big)\, a_{j,k} = (-1)^{i+1} \det\big(\mathbf{A}(\hat{1},\hat{\imath})\big)\, a_{1,k} + \det\big(\mathbf{A}(\hat{2},\hat{\imath})\big)\, a_{2,k} + \cdots + \det\big(\mathbf{A}(\hat{n},\hat{\imath})\big)\, a_{n,k}\,.$$

When $i = k$, the (i,k)-entry is equal to the Laplace expansion for $\det(\mathbf{A})$ along the k-th column. On the other hand, suppose that $i \neq k$. If $\mathbf{C}$ is the matrix obtained from $\mathbf{A}$ by replacing the j-th column of $\mathbf{A}$ with the k-th column of $\mathbf{A}$, then (i,k)-entry is equal to the Laplace expansion of $\det(\mathbf{C})$ along the k-th column. Since two columns of $\mathbf{C}$ are equal, we conclude that $\det(\mathbf{C}) = 0$. □

11.0.4 Corollary (Relation between adjugates and inverses)**.** *When the matrix* $\mathbf{A}$ *is invertible, we have* $\mathbf{A}^{-1} = \big(\det(\mathbf{A})\big)^{-1} \operatorname{adj}(\mathbf{A})$.

Proof. Since the matrix $\mathbf{A}$ is invertible, the characterization of the determinant [10.2.1] establishes that $\det(\mathbf{A}) \neq 0$. The adjugate equation implies that $\big(\det(\mathbf{A})\big)^{-1} \operatorname{adj}(\mathbf{A})\,\mathbf{A} = \mathbf{I}$, so the characterizations of invertible matrices [8.0.3] shows that $\mathbf{A}^{-1} = \big(\det(\mathbf{A})\big)^{-1} \operatorname{adj}(\mathbf{A})$. □

For any (3×3)-matrix, this corollary leads to an effective method for computing the inverse by hand.

11.0.5 Algorithm (The Bayer method)**.**

This algorithm is named after D.A. Bayer.

input: a (3×3)-matrix invertible $\mathbf{A} := [a_{j,k}]$.
output: the inverse $\mathbf{A}^{-1}$.

Write out $\mathbf{A}^{\mathsf{T}}$ leaving blank lines between the rows and columns.
Duplicate the first two columns on the right.
Duplicate the first two rows on the bottom.
Cross-out the first row and column in the (5×5)-array.
For the 9 adjacent (2×2)-submatrices, compute the determinant.
Multiply the (3×3)-matrix $\mathbf{B}$ formed by the determinants by $\mathbf{A}$.
If the product equals $d\,\mathbf{I}$ for some nonzero scalar d, then return $d^{-1}\,\mathbf{B}$.
If the product is zero, then the matrix $\mathbf{A}$ is not invertible.

Correctness of algorithm. The (3×3)-matrix produced by the first five steps is the adjugate of the matrix $\mathbf{A}$. Hence, the conditions in the final step follow immediately from the adjugate equation. □

11.0.6 Problem. Use the Bayer method to compute the inverse of

$$\mathbf{A} := \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}.$$

Solution. As the last steps in the Bayer method, we calculate

$$\begin{bmatrix} 2 & -1 & -1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} \Rightarrow \mathbf{A}^{-1} = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ 1 & 1 & -2 \\ 1 & 1 & 1 \end{bmatrix}. \square$$

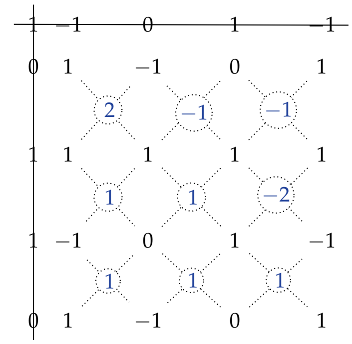

Figure 11.1: The Bayer array

The adjugate equation also produces an explicit formula for the solutions of many linear systems.

11.0.7 Corollary (The Cramer Rule)**.** *For any invertible matrix* $\mathbf{A}$*, the unique solution to the non-homogeneous linear system* $\mathbf{A}\,\vec{x} = \vec{b}$ *satisfies*

$$x_k = \frac{1}{\det(\mathbf{A})} \left(\sum_{k=1}^{n} (-1)^{j+k}\, b_k \det\big(\mathbf{A}(\hat{k}, \hat{\jmath})\big) \right)$$
$$= \tfrac{1}{\det(\mathbf{A})} \Big((-1)^{j+1} b_1 \det\big(\mathbf{A}(\hat{1}, \hat{\jmath})\big) + (-1)^{j+2} b_2 \det\big(\mathbf{A}(\hat{2}, \hat{\jmath})\big) + \cdots + (-1)^{j+n} b_n \det\big(\mathbf{A}(\hat{n}, \hat{\jmath})\big) \Big).$$

The numerator is the Laplace expansion along the j-th column of the matrix obtain from $\mathbf{A}$ *by replacing the j-th column with the vector* $\vec{b}$*.*

This rule is named after G. Cramer who published it in 1750.

Proof. Since the matrix $\mathbf{A}$ is invertible, the relation between adjugates and inverses implies that $\mathbf{A}^{-1} = \big(\det(\mathbf{A})\big)^{-1} \operatorname{adj}(\mathbf{A})$. The unique solution to linear system $\mathbf{A}\,\vec{x} = \vec{b}$ is $\vec{x} = \mathbf{A}^{-1}\,\vec{b} = \big(\det(\mathbf{A})\big)^{-1} \operatorname{adj}(\mathbf{A})\,\vec{b}$. Combining the definition of the matrix product and the definition of the adjugate yields the required formula. □

Computing a single determinant takes about as much work as solving $\mathbf{A}\,\vec{x} = \vec{b}$ via row reduction. Thus, for large matrices, the Cramer rule is hopelessly inefficient.

11.0.8 Problem. Use the Cramer rule to solve $\begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$.

Solution. The Cramer rule gives

$$x = \frac{\det\left[\begin{smallmatrix} 6 & -2 \\ 8 & 4 \end{smallmatrix}\right]}{\det\left[\begin{smallmatrix} 3 & -2 \\ -5 & 4 \end{smallmatrix}\right]} = \frac{(6)(4) - (-2)(8)}{(3)(4) - (-2)(-5)} = \frac{40}{2} = 20,$$
$$y = \frac{\det\left[\begin{smallmatrix} 3 & 6 \\ -5 & 8 \end{smallmatrix}\right]}{\det\left[\begin{smallmatrix} 3 & -2 \\ -5 & 4 \end{smallmatrix}\right]} = \frac{(3)(8) - (6)(-5)}{(3)(4) - (-2)(-5)} = \frac{54}{2} = 27. \qquad \square$$

11.0.9 Problem. Determine the values of t for which the non-homogeneous linear system has a unique solution and use the Cramer rule to describe this solution:

$$\begin{bmatrix} 3t & -2 \\ -6 & t \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}.$$

Solution. The Cramer rule gives

$$x = \frac{\det\left[\begin{smallmatrix} 4 & -2 \\ 1 & t \end{smallmatrix}\right]}{\det\left[\begin{smallmatrix} 3t & -2 \\ -6 & t \end{smallmatrix}\right]} = \frac{4t + 2}{3t^2 - 12} = \frac{2(2t+1)}{3(t+2)(t-2)},$$
$$y = \frac{\det\left[\begin{smallmatrix} 3t & 4 \\ -6 & 1 \end{smallmatrix}\right]}{\det\left[\begin{smallmatrix} 3t & -2 \\ -6 & t \end{smallmatrix}\right]} = \frac{3t + 24}{3(t+2)(t-2)} = \frac{t+8}{(t+2)(t-2)},$$

so this linear system has a unique solution if and only if $t \neq \pm 2$. □

The Cramer rule is important because it expresses $\mathbf{A}^{-1}$ and solutions to $\mathbf{A}\,\vec{x} = \vec{b}$ as quotients of polynomials in the entries of $\mathbf{A}$ and $\vec{b}$ with integer coefficients. Hence, if the entries of $\mathbf{A}$ and $\vec{b}$ are all continuous functions, then so are the solutions x_j.

Exercises

11.0.10 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. Every matrix has an adjugate.
ii. The adjugate of a matrix has the same size as the original matrix.
iii. The product of a matrix and its adjugate equals the determinant.
iv. The Bayer method provides a useful mnemonic for remembering all adjugates.
v. The Cramer rule always provides the best method for solving a linear system.

11.1 Other Determinantal Formula

HOW CAN WE EXPRESS THE DETERMINANT VIA PERMUTATIONS? To relate determinants and permutations, we first need another numerical invariant of a permutation.

11.1.0 Definition. The ***sign*** of a permutation σ of the set $\{1, 2, \dots, n\}$, denoted by $\operatorname{sgn}(\sigma)$, equals the determinant of the corresponding permutation matrix.

11.1.1 Problem. Compute the sign of the permutation $3\ \ 4\ \ 5\ \ 2\ \ 1$.

Solution. We calculate the determinant of its permutation matrix:

$$\det\left(\begin{bmatrix} 0&0&0&0&1\\ 0&0&0&1&0\\ 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&1&0&0 \end{bmatrix}\right) \xrightarrow[=]{\substack{\vec{r}_3 \mapsto \vec{r}_1\\ \vec{r}_1 \mapsto \vec{r}_3\\ \vec{r}_4 \mapsto \vec{r}_2\\ \vec{r}_2 \mapsto \vec{r}_4}} \det\left(\begin{bmatrix} 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&0&0&1\\ 0&0&0&1&0\\ 0&0&1&0&0 \end{bmatrix}\right) \xrightarrow[=]{\substack{\vec{r}_3 \mapsto \vec{r}_5\\ \vec{r}_5 \mapsto \vec{r}_3}} (-1)\det\left(\begin{bmatrix} 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&1&0\\ 0&0&0&0&1 \end{bmatrix}\right) = -1\,.$$

Therefore, we see that $\operatorname{sgn}(3\ \ 4\ \ 5\ \ 2\ \ 1) = -1$. □

Alternatively, the sign of a permutation can be derived from its decomposition into the product of transpositions.

11.1.2 Lemma. *When the permutation σ of the set $\{1, 2, \dots, n\}$ is a product of ℓ transpositions, we have $\operatorname{sgn}(\sigma) = (-1)^{\ell}$.*

Proof. Let $\mathbf{P}$ be the $(n \times n)$-matrix associated to the permutation σ. When the permutation σ is a product of ℓ transpositions, there exists a sequence $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_\ell$ of elementary matrices corresponding row swaps such that $\mathbf{P} = \mathbf{R}_1\,\mathbf{R}_2\,\cdots\,\mathbf{R}_\ell$. The compatibility of determinants with multiplication and the determinants for elementary matrices [10.1.8] give

$$\operatorname{sgn}(\sigma) = \det(\mathbf{P}) = \det(\mathbf{R}_1\,\mathbf{R}_2\,\cdots\,\mathbf{R}_\ell) = \prod_{j=1}^{\ell} \det(\mathbf{R}_j) = (-1)^{\ell}\,. \quad \square$$

Our next formula expresses the determinant of a square matrix in terms of permutations of the matrix entries.

The Leibniz formula has too many terms to be useful for computation. It is important because it establishes that determinants are polynomials in the n^2 entries of the matrix with coefficents ± 1. For example, if each matrix entry is a differentiable function of a single variable, then $\det(\mathbf{A})$ is also a differentiable function.

11.1.3 Theorem (Leibniz formula). *For all $1 \leqslant j \leqslant n$ and all $1 \leqslant k \leqslant n$, let $a_{j,k} \in \mathbb{K}$ denote the (j,k)-entry of the matrix $\mathbf{A}$. We have*

$$\det(\mathbf{A}) = \sum_{\sigma} \left(\operatorname{sgn}(\sigma) \prod_{k=1}^{n} a_{\sigma(k),k} \right),$$

where the sum is over the $n!$ permutations σ of the set $\{1, 2, \dots, n\}$.

Proof. Given the characterization of the determinant [10.2.1], it suffices to show that the function $f \colon \mathbb{K}^{n \times n} \to \mathbb{K}^n$ defined by

$$f(\mathbf{A}) := \sum_{\sigma} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} a_{\sigma(k),k}$$

satisfies the three key properties.

(identity) The (j,k)-entry in the identity matrix $\mathbf{I}$ is the Kronecker delta $\delta_{j,k}$. Hence, the product $\prod_{k=1}^{n} \delta_{\sigma(k),k}$ is zero unless $\sigma(k) = k$ for all $1 \leqslant k \leqslant n$ or $\sigma = 1\ 2\ \cdots\ n$. We conclude that

$$f(\mathbf{I}) = \sum_{\sigma} \operatorname{sgn}(\sigma) \prod_{k=1}^{n} \delta_{\sigma(k),k} = 1\,.$$

(linearity) Let $\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2, \dots, \vec{\boldsymbol{a}}_{i-1}, \vec{\boldsymbol{v}}, \vec{\boldsymbol{w}}, \vec{\boldsymbol{a}}_{i+1}, \vec{\boldsymbol{a}}_{i+2}, \dots, \vec{\boldsymbol{a}}_n \in \mathbb{K}^n$ be column vectors and let $c, d \in \mathbb{K}$ be scalars. It follows that

$$\begin{aligned}
& f\big(\big[\vec{\boldsymbol{a}}_1\ \ \vec{\boldsymbol{a}}_2\ \ \cdots\ \ \vec{\boldsymbol{a}}_{i-1}\ \ c\vec{\boldsymbol{v}} + d\,\vec{\boldsymbol{w}}\ \ \vec{\boldsymbol{a}}_{i+1}\ \ \cdots\ \ \vec{\boldsymbol{a}}_n\big]\big) \\
&= \sum_{\sigma} \left(\operatorname{sgn}(\sigma)\ (c\,v_{\sigma(i)} + d\,w_{\sigma(i)}) \prod_{k \neq i} a_{\sigma(k),k} \right) \\
&= c \sum_{\sigma} \left(\operatorname{sgn}(\sigma)\ v_{\sigma(i)} \prod_{k \neq i} a_{\sigma(k),k} \right) + d \sum_{\sigma} \left(\operatorname{sgn}(\sigma)\ w_{\sigma(i)} \prod_{k \neq i} a_{\sigma(k),k} \right) \\
&= c\, f\big(\big[\vec{\boldsymbol{a}}_1\ \ \vec{\boldsymbol{a}}_2\ \ \cdots\ \ \vec{\boldsymbol{a}}_{i-1}\ \ \vec{\boldsymbol{v}}\ \ \vec{\boldsymbol{a}}_{i+1}\ \ \cdots\ \ \vec{\boldsymbol{a}}_n\big]\big) + d\, f\big(\big[\vec{\boldsymbol{a}}_1\ \ \vec{\boldsymbol{a}}_2\ \ \cdots\ \ \vec{\boldsymbol{a}}_{i-1}\ \ \vec{\boldsymbol{w}}\ \ \vec{\boldsymbol{a}}_{i+1}\ \ \cdots\ \ \vec{\boldsymbol{a}}_n\big]\big)
\end{aligned}$$

which shows that f is linear in the i-th column.

(alternating) Let $\vec{\boldsymbol{a}}_1, \vec{\boldsymbol{a}}_2, \dots, \vec{\boldsymbol{a}}_{i-1}, \vec{\boldsymbol{v}}, \vec{\boldsymbol{a}}_{i+2}, \vec{\boldsymbol{a}}_{i+3}, \dots, \vec{\boldsymbol{a}}_n \in \mathbb{K}^n$ be column vectors. If τ is the transposition that interchanges i and $i+1$, then the set of permutations of $\{1, 2, \dots, n\}$ is a disjoint union subsets of equal size: $\{\sigma \mid \sigma(i) < \sigma(i+1)\} \sqcup \{\sigma \circ \tau \mid \sigma(i) < \sigma(i+1)\}$. It follows that

$$\begin{aligned}
& f\big(\big[\vec{\boldsymbol{a}}_1\ \ \vec{\boldsymbol{a}}_2\ \ \cdots\ \ \vec{\boldsymbol{a}}_{i-1}\ \ \vec{\boldsymbol{v}}\ \ \vec{\boldsymbol{v}}\ \ \vec{\boldsymbol{a}}_{i+2}\ \ \cdots\ \ \vec{\boldsymbol{a}}_n\big]\big) \\
&= \sum_{\sigma(i) < \sigma(i+1)} \left(\operatorname{sgn}(\sigma)\ v_{\sigma(i)}\ v_{\sigma(i+1)} \prod_{k \neq i, i+1} a_{\sigma(k),k} + \operatorname{sgn}(\sigma \circ \tau)\ v_{\sigma\circ\tau(i)}\ v_{\sigma\circ\tau(i+1)} \prod_{k \neq i, i+1} a_{\sigma\circ\tau(k),k} \right) \\
&= \sum_{\sigma(i) < \sigma(i+1)} \left(\operatorname{sgn}(\sigma)\ v_{\sigma(i)}\ v_{\sigma(i+1)} \prod_{k \neq i, i+1} a_{\sigma(k),k} - \operatorname{sgn}(\sigma)\ v_{\sigma(i+1)}\ v_{\sigma(i)} \prod_{k \neq i, i+1} a_{\sigma(k),k} \right) \\
&= \sum_{\sigma(i) < \sigma(i+1)} \left(\operatorname{sgn}(\sigma)\ \big(v_{\sigma(i)}\ v_{\sigma(i+1)} - v_{\sigma(i+1)}\ v_{\sigma(i)}\big) \prod_{k \neq i, i+1} a_{\sigma(k),k} \right) = 0\,,
\end{aligned}$$

which establishes that f vanishes when the i-th and $(i+1)$-st columns are equal. □

11.1.4 Problem (Vandermonde determinant)**.** For any nonnegative integer n, prove that

$$\det(\mathbf{V}_n) = \prod_{j=1}^{n-1} \left(\prod_{k=j+1}^{n} (x_k - x_j) \right) = \prod_{1 \leqslant j < k \leqslant n} (x_k - x_j)$$

where

$$\mathbf{V}_n := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{bmatrix} .$$

Although A.-T. Vandemonde made major contributions to the theory of determinants, this specific determinant does not appear in his four mathematical papers on the subject. It seems likely that this attribution arose from someone's misreading Vandermonde's notation.

Polynomial solution. Since the (j,k)-entry of the Vandermonde matrix $\mathbf{V}_n$ is x_k^{j-1}, the Leibniz formula shows that the determinant equals $\sum_\sigma \operatorname{sgn}(\sigma) \prod_{k=1}^n x_k^{\sigma(k)-1}$. Regarding this expression as a polynomial in the variables $x_1, x_2, \dots, x_n$, we see that each monomial $\prod_{k=1}^n x_k^{\sigma(k)-1}$ has degree $\sum_{j=1}^n j - 1 = \sum_{j=0}^{n-1} j = n(n-1)/2$. Furthermore, the monomial $\prod_{j=1}^n x_j^{j-1} = x_2 x_3^2 x_4^3 \cdots x_n^{n-1}$, which corresponds to the identity permutation, has coefficient 1. This polynomial is divisible by $x_k - x_j$ for all $1 \leqslant j < k \leqslant n$, because the determinant vanishes when $x_j = x_k$. On the other hand, the homogeneous polynomial

$$\prod_{1 \leqslant j < k \leqslant n} (x_k - x_j) = \prod_{k=2}^{n} \left(\prod_{j=1}^{k-1} (x_k - x_j) \right)$$

has degree $n(n-1)/2$ and the coefficient of $x_2 x_3^2 x_4^3 \cdots x_n^{n-1}$ is 1, so we conclude that

$$\sum_\sigma \left(\operatorname{sgn}(\sigma) \prod_{k=1}^{n} x_k^{\sigma(k)-1} \right) = \prod_{k=2}^{n} \left(\prod_{j=1}^{k-1} (x_k - x_j) \right) = \prod_{1 \leqslant j < k \leqslant n} (x_k - x_j) . \quad \square$$

Inductive solution. The base case $n = 0$ is vacuously true. For each $2 \leqslant j \leqslant n$, add $-x_n$ times $(j-1)$-th row of V_n to the j-th row V_n. Since determinants are invariant under elementary row operations, we have

$$\det(\mathbf{V}_n) = \det\left(\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ x_n - x_1 & x_n - x_2 & \cdots & x_n - x_{n-1} & 0 \\ x_1(x_n - x_1) & x_2(x_2 - x_1) & \cdots & x_{n-1}(x_n - x_{n-1}) & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^{n-2}(x_n - x_1) & x_2^{n-2}(x_2 - x_1) & \cdots & x_{n-1}^{n-2}(x_n - x_{n-1}) & 0 \end{bmatrix} \right) .$$

The induction hypothesis states that $\det(\mathbf{V}_{n-1}) = \prod_{k=2}^{n-1} \prod_{j=1}^{k-1} (x_k - x_j)$. Hence, the Laplace expansion along the last column, together with linearity in each column, yields

$$\begin{aligned}
\det(\mathbf{V}_n) &= \left(\prod_{k=1}^{n-1} (x_n - x_k) \right) \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_{n-1} \\ x_1^2 & x_2^2 & \cdots & x_{n-1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-2} & x_2^{n-2} & \cdots & x_{n-1}^{n-2} \end{bmatrix} \\
&= \left(\prod_{j=1}^{n-1} (x_n - x_j) \right) \det(\mathbf{V}_{n-1}) \\
&= \left(\prod_{k=n}^{n} \prod_{j=1}^{k-1} (x_k - x_j) \right) \left(\prod_{k=2}^{n-1} \prod_{j=1}^{k-1} (x_k - x_j) \right) = \prod_{k=2}^{n} \prod_{j=1}^{k-1} (x_k - x_j) \,. \square
\end{aligned}$$

Exercises

11.1.5 Problem. Determine which of the following statements are true. If a statement is false, then provide a counterexample.

i. The sign of a permutation is either $+1$ or -1.
ii. The complete expansion for a determinant of an $(n \times n)$-matrix has $n!$ terms.
iii. The columns in a Vandermonde matrix are terms in a geometric sequence.
iv. For an $(n \times n)$-matrix, the Vandermonde determinant is the product of $\frac{n(n-1)}{2}$ linear factors.