5.3 Determinants and Cramer's Rule

Unique Solution of a 2×2 System

The 2×2 system

$$\begin{array}{rcrcl} ax & + & by & = & e, \\ cx & + & dy & = & f, \end{array} \tag{1}$$

has a unique solution provided $\Delta = ad - bc$ is nonzero, in which case the solution is given by

$$x = \frac{de - bf}{ad - bc}, \quad y = \frac{af - ce}{ad - bc}. \tag{2}$$

This result, called **Cramer's Rule** for 2×2 systems, is usually learned in college algebra as part of determinant theory.

Determinants of Order 2

College algebra introduces matrix notation and determinant notation:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Evaluation of a 2×2 determinant is by **Sarrus' Rule**:

$$\begin{vmatrix} \mathbf{a} & b \\ c & \mathbf{d} \end{vmatrix} = \mathbf{ad} - bc.$$

The boldface product **ad** is the product of the main diagonal entries and the other product bc is from the anti-diagonal.

Cramer's 2×2 rule in determinant notation is

$$x = \frac{\begin{vmatrix} e & b \\ f & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}, \quad y = \frac{\begin{vmatrix} a & e \\ c & f \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}. \tag{3}$$

Unique Solution of an $n \times n$ System

Cramer's rule can be generalized to an $n \times n$ system of equations $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ or

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1, \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2, \\ \vdots & & \vdots & & \cdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \cdots & + & a_{nn}x_n & = & b_n. \end{array} \tag{4}$$

System (4) has a unique solution provided the **determinant of coefficients** $\Delta = \det(A)$ is nonzero, in which case the solution is given by

$$x_1 = \frac{\Delta_1}{\Delta},\; x_2 = \frac{\Delta_2}{\Delta},\; \ldots,\; x_n = \frac{\Delta_n}{\Delta}. \tag{5}$$

The determinant Δ_j equals $\det(B_j)$ where matrix B_j is matrix A with column j replaced by $\vec{\mathbf{b}} = (b_1, \ldots, b_n)$, which is the right side of system (4). The result is called **Cramer's Rule** for $n \times n$ systems. Determinants will be defined shortly; intuition from the 2×2 case and Sarrus' rule should suffice for the moment.

Determinant Notation for Cramer's Rule. The **determinant of coefficients** for system $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is denoted by

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}. \tag{6}$$

The other n determinants in Cramer's rule (5) are given by

$$\Delta_1 = \begin{vmatrix} b_1 & a_{12} & \cdots & a_{1n} \\ b_2 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_n & a_{n2} & \cdots & a_{nn} \end{vmatrix}, \ldots, \Delta_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & b_1 \\ a_{21} & a_{22} & \cdots & b_2 \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & b_n \end{vmatrix}. \tag{7}$$

The literature is filled with conflicting notations for matrices, vectors and determinants. The reader should take care to use vertical bars *only* for determinants and absolute values, e.g., $|A|$ makes sense for a matrix A or a constant A. For clarity, the notation $\det(A)$ is preferred, when A is a matrix. The notation $|A|$ implies that *a determinant is a number*, computed by $|A| = \pm A$ when $n = 1$, and $|A| = a_{11}a_{22} - a_{12}a_{21}$ when $n = 2$. For $n \geq 3$, $|A|$ is computed by similar but increasingly complicated formulas; see Sarrus' rule and the *four properties* below.

Sarrus' Rule for 3×3 Matrices. College algebra supplies the following formula for the determinant of a 3×3 matrix A:

$$\begin{aligned} \det(A) &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ &\quad -a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}. \end{aligned} \tag{8}$$

The number $\det(A)$ can be computed by an algorithm similar to the one for 2×2 matrices, as in Figure 10. We remark that no further generalizations are possible: *there is no Sarrus' rule for* 4×4 *or larger matrices*!

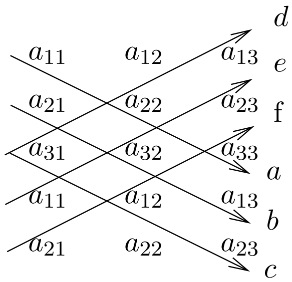

Figure 10. Sarrus' rule for 3×3 matrices, which gives

$$\det(A) = (a + b + c) - (d + e + f).$$

College Algebra Definition of Determinant.

The impractical definition is the formula

$$\det(A) = \sum_{\sigma \in S_n} (-1)^{\text{parity}(\sigma)}\, a_{1\sigma_1} \cdots a_{n\sigma_n}. \tag{9}$$

In the formula, a_{ij} denotes the element in row i and column j of the matrix A. The symbol $\sigma = (\sigma_1, \ldots, \sigma_n)$ stands for a rearrangement of the subscripts 1, 2, …, n and S_n is the set of all possible rearrangements. The nonnegative integer parity(σ) is determined by counting the minimum number of pairwise interchanges required to assemble the list of integers $\sigma_1, \ldots, \sigma_n$ into natural order $1, \ldots, n$.

A consequence of (9) is the relation $\det(A) = \det(A^T)$ where A^T means the transpose of A, obtained by swapping rows and columns. This relation implies that all determinant theory results for rows also apply to columns.

Formula (9) reproduces the definition for 3×3 matrices given in equation (8). We will have no computational use for (9). For computing the value of a determinant, see below *four properties* and *cofactor expansion*.

Four Properties.

The definition of determinant (9) implies the following four properties:

Triangular	The value of $\det(A)$ for either an upper triangular or a lower triangular matrix A is the product of the diagonal elements: $\det(A) = a_{11}a_{22} \cdots a_{nn}$.
Swap	If B results from A by swapping two rows, then $\det(A) = (-1)\det(B)$.
Combination	The value of $\det(A)$ is unchanged by adding a multiple of a row to a different row.
Multiply	If one row of A is multiplied by constant c to create matrix B, then $\det(B) = c\det(A)$.

It is known that these four rules suffice to compute the value of any $n \times n$ determinant. The proof of the four properties is delayed until page 314.

Elementary Matrices and the Four Rules. The rules can be stated in terms of elementary matrices as follows.

Triangular	The value of $\det(A)$ for either an upper triangular or a lower triangular matrix A is the product of the diagonal elements: $\det(A) = a_{11}a_{22}\cdots a_{nn}$. This is a one-arrow Sarrus' rule valid for dimension n.
Swap	If E is an elementary matrix for a swap rule, then $\det(EA) = (-1)\det(A)$.
Combination	If E is an elementary matrix for a combination rule, then $\det(EA) = \det(A)$.
Multiply	If E is an elementary matrix for a multiply rule with multiplier $c \neq 0$, then $\det(EA) = c\det(A)$.

Since $\det(E) = 1$ for a combination rule, $\det(E) = -1$ for a swap rule and $\det(E) = c$ for a multiply rule with multiplier $c \neq 0$, it follows that for any elementary matrix E there is the determinant multiplication rule

$$\det(EA) = \det(E)\det(A).$$

Additional Determinant Rules. The following rules make for efficient evaluation of certain special determinants. The results are stated for rows, but they also hold for columns, because $\det(A) = \det(A^T)$.

Zero row	If one row of A is zero, then $\det(A) = 0$.
Duplicate rows	If two rows of A are identical, then $\det(A) = 0$.
RREF $\neq I$	If $\mathbf{rref}(A) \neq I$, then $\det(A) = 0$.
Common factor	The relation $\det(A) = c\det(B)$ holds, provided A and B differ only in one row, say row j, for which $\mathbf{row}(A,j) = c\,\mathbf{row}(B,j)$.
Row linearity	The relation $\det(A) = \det(B) + \det(C)$ holds, provided A, B and C differ only in one row, say row j, for which $\mathbf{row}(A,j) = \mathbf{row}(B,j) + \mathbf{row}(C,j)$.

The proofs of these properties are delayed until page 314.

Cofactor Expansion

The special subject of cofactor expansions is used to justify Cramer's rule and to provide an alternative method for computation of determinants. There is no claim that cofactor expansion is efficient, only that it is possible, and different than Sarrus' rule or the use of the four properties.

Background from College Algebra. The cofactor expansion theory is most easily understood from the college algebra topic, where the dimension is 3 and row expansion means the following formulas are valid:

$$
\begin{aligned}
|A| &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&= a_{11}(+1)\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}(+1)\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\
&= a_{21}(-1)\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{22}(+1)\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + a_{23}(-1)\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\
&= a_{31}(+1)\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} + a_{32}(-1)\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33}(+1)\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}
\end{aligned}
$$

The formulas expand a 3×3 determinant in terms of 2×2 determinants, along a row of A. The attached signs ± 1 are called the **checkerboard signs**, to be defined shortly. The 2×2 determinants are called **minors** of the 3×3 determinant $|A|$. The checkerboard sign together with a minor is called a **cofactor**.

These formulas are generally used when a row has one or two zeros, making it unnecessary to evaluate one or two of the 2×2 determinants in the expansion. To illustrate, row 1 expansion gives

$$
\begin{vmatrix} 3 & 0 & 0 \\ 2 & 1 & 7 \\ 5 & 4 & 8 \end{vmatrix} = 3(+1)\begin{vmatrix} 1 & 7 \\ 4 & 8 \end{vmatrix} = -60.
$$

A clever time–saving choice is always a row which has the most zeros, although success does not depend upon cleverness. What has been said for rows also applies to columns, due to the transpose formula $|A| = |A^T|$.

Minors and Cofactors. The $(n-1)\times(n-1)$ determinant obtained from $\det(A)$ by striking out row i and column j is called the (i,j)–minor of A and denoted $\textsf{minor}(A,i,j)$ (M_{ij} is common in literature). The (i,j)–cofactor of A is $\textsf{cof}(A,i,j) = (-1)^{i+j}\,\textsf{minor}(A,i,j)$. Multiplicative factor $(-1)^{i+j}$ is called the **checkerboard sign**, because its value can be determined by counting *plus*, *minus*, *plus*, etc., from location $(1,1)$ to location (i,j) in any checkerboard fashion.

Expansion of Determinants by Cofactors. The formulas are

$$
(10) \quad \det(A) = \sum_{j=1}^{n} a_{kj}\,\textsf{cof}(A,k,j), \quad \det(A) = \sum_{i=1}^{n} a_{i\ell}\,\textsf{cof}(A,i,\ell),
$$

where $1 \le k \le n$, $1 \le \ell \le n$. The first expansion in (10) is called a **cofactor row expansion** and the second is called a **cofactor column expansion**. The value $\textbf{cof}(A,i,j)$ is the cofactor of element a_{ij} in $\det(A)$, that is, the checkerboard sign times the minor of a_{ij}. The proof of expansion (10) is delayed until page 314.

The Adjugate Matrix. The **adjugate** $\textbf{adj}(A)$ of an $n \times n$ matrix A is the transpose of the matrix of cofactors,

$$\textbf{adj}(A) = \begin{pmatrix} \textbf{cof}(A,1,1) & \textbf{cof}(A,1,2) & \cdots & \textbf{cof}(A,1,n) \\ \textbf{cof}(A,2,1) & \textbf{cof}(A,2,2) & \cdots & \textbf{cof}(A,2,n) \\ \vdots & \vdots & \cdots & \vdots \\ \textbf{cof}(A,n,1) & \textbf{cof}(A,n,2) & \cdots & \textbf{cof}(A,n,n) \end{pmatrix}^T.$$

A cofactor $\textbf{cof}(A,i,j)$ is the checkerboard sign $(-1)^{i+j}$ times the corresponding minor determinant $\textbf{minor}(A,i,j)$. In the 2×2 case,

$$\textbf{adj}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

In words: *swap the diagonal elements and change the sign of the off–diagonal elements.*

The Inverse Matrix. The adjugate appears in the formula for the inverse matrix A^{-1}:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

This formula is verified by direct matrix multiplication:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} = (a_{11}a_{22} - a_{12}a_{21}) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For an $n \times n$ matrix, $A \cdot \textbf{adj}(A) = \det(A)\, I$, which gives the formula

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} \textbf{cof}(A,1,1) & \textbf{cof}(A,1,2) & \cdots & \textbf{cof}(A,1,n) \\ \textbf{cof}(A,2,1) & \textbf{cof}(A,2,2) & \cdots & \textbf{cof}(A,2,n) \\ \vdots & \vdots & \cdots & \vdots \\ \textbf{cof}(A,n,1) & \textbf{cof}(A,n,2) & \cdots & \textbf{cof}(A,n,n) \end{pmatrix}^T$$

The proof of $A \cdot \textbf{adj}(A) = \det(A)\, I$ is delayed to page 316.

Elementary Matrices. An elementary matrix E is the result of applying a combination, multiply or swap rule to the identity matrix. This definition implies that an elementary matrix is the identity matrix with a minor change applied, to wit:

Combination	Change an off-diagonal zero of I to c.
Multiply	Change a diagonal one of I to multiplier $m \neq 0$.
Swap	Swap two rows of I.

Theorem 9 (Determinants and Elementary Matrices)
Let E be an $n \times n$ elementary matrix. Then

Combination	$\det(E) = 1$
Multiply	$\det(E) = m$ for multiplier m.
Swap	$\det(E) = -1$
Product	$\det(EX) = \det(E)\det(X)$ for all $n \times n$ matrices X.

Theorem 10 (Determinants and Invertible Matrices)
Let A be a given invertible matrix. Then

$$\det(A) = \frac{(-1)^s}{m_1 m_2 \cdots m_r}$$

where s is the number of swap rules applied and m_1, m_2, ..., m_r are the nonzero multipliers used in multiply rules when A is reduced to $\mathbf{rref}(A)$.

Determinant Product Rule. The determinant rules of combination, multiply and swap imply that $\det(EX) = \det(E)\det(X)$ for elementary matrices E and square matrices X. We show that a more general relationship holds.

Theorem 11 (Determinant Product Rule)
Let A and B be given $n \times n$ matrices. Then

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

Proof:

Used in the proof is the equivalence of invertibility of a square matrix C with $\det(C) \neq 0$ and $\mathbf{rref}(C) = I$.

Assume one of A or B has zero determinant. Then $\det(A)\det(B) = 0$. If $\det(B) = 0$, then $B\mathbf{x} = \mathbf{0}$ has infinitely many solutions, in particular a nonzero solution $\mathbf{x}$. Multiply $B\mathbf{x} = \mathbf{0}$ by A, then $AB\mathbf{x} = \mathbf{0}$ which implies AB is not invertible. Then the identity $\det(AB) = \det(A)\det(B)$ holds, because both sides are zero. If $\det(B) \neq 0$ but $\det(A) = 0$, then there is a nonzero $\mathbf{y}$ with $A\mathbf{y} = \mathbf{0}$. Define $\mathbf{x} = AB^{-1}\mathbf{y}$. Then $AB\mathbf{x} = A\mathbf{y} = \mathbf{0}$, with $\mathbf{x} \neq \mathbf{0}$, which implies the identity holds.. This completes the proof when one of A or B is not invertible.

Assume A, B are invertible and then $C = AB$ is invertible. In particular, $\mathbf{rref}(A^{-1}) = \mathbf{rref}(B^{-1}) = I$. Write $I = \mathbf{rref}(A^{-1}) = E_1 E_2 \cdots E_k A^{-1}$ and $I = \mathbf{rref}(B^{-1}) = F_1 F_2 \cdots F_m B^{-1}$ for elementary matrices E_i, F_j. Then

$$AB = E_1 E_2 \cdots E_k F_1 F_2 \cdots F_m. \tag{11}$$

The theorem follows from repeated application of the basic identity $\det(EX) = \det(E)\det(X)$ to relation (11), because

$$\det(A) = \det(E_1)\cdots\det(E_k), \quad \det(B) = \det(F_1)\cdots\det(F_m).$$

The Cayley-Hamilton Theorem

Presented here is an adjoint formula $F^{-1} = \mathbf{adj}(F)/\det(F)$ derivation for the celebrated Cayley-Hamilton formula

$$(-A)^n + p_{n-1}(-A)^{n-1} + \cdots + p_0 I = 0. \tag{12}$$

The $n\times n$ matrix A is given and I is the identity matrix. The coefficients p_k in (12) are determined by the **characteristic polynomial** of matrix A, which is defined by the determinant expansion formula

$$\det(A - \lambda I) = (-\lambda)^n + p_{n-1}(-\lambda)^{n-1} + \cdots + p_0. \tag{13}$$

The Cayley-Hamilton Theorem is summarized as follows:

A square matrix A satisfies its own characteristic equation.

Proof of (12): Define $x = -\lambda$, $F = A + xI$ and $G = \mathbf{adj}(F)$. A cofactor of $\det(F)$ is a polynomial in x of degree at most $n-1$. Therefore, there are $n \times n$ constant matrices $C_0, \ldots, C_{n-1}$ such that

$$\mathbf{adj}(F) = x^{n-1}C_{n-1} + \cdots + xC_1 + C_0.$$

The adjoint formula for F^{-1} gives $\det(F)I = \mathbf{adj}(F)\,F$. Relation (13) implies $\det(F) = x^n + p_{n-1}x^{n-1} + \cdots + p_0$. Expand the matrix product $\mathbf{adj}(F)F$ in powers of x as follows:

$$\begin{aligned} \mathbf{adj}(F)F &= \left(\sum_{j=0}^{n-1} x^j C_j\right)(A + xI) \\ &= C_0 A + \sum_{i=1}^{n-1} x^i (C_i A + C_{i-1}) + x^n C_{n-1}. \end{aligned}$$

Match coefficients on each side of $\det(F)I = \mathbf{adj}(F)F$ to give the relations

$$\left\{ \begin{aligned} p_0 I &= C_0 A, \\ p_1 I &= C_1 A + C_0, \\ p_2 I &= C_2 A + C_1, \\ &\vdots \\ I &= C_{n-1}. \end{aligned} \right. \tag{14}$$

To complete the proof of the Cayley-Hamilton identity (12), multiply the equations in (14) by I, $(-A)$, $(-A)^2$, $(-A)^3$, …, $(-A)^n$, respectively. Then add all the equations. The left side matches (12). The right side is a telescoping sum which adds to the zero matrix. The proof is complete.

2 Example (Four Properties) Apply the four properties of a determinant to justify the formula

$$\det\begin{pmatrix} 12 & 6 & 0 \\ 11 & 5 & 1 \\ 10 & 2 & 2 \end{pmatrix} = 24.$$

Solution: Let D denote the value of the determinant. Then

$D = \det\begin{pmatrix} 12 & 6 & 0 \\ 11 & 5 & 1 \\ 10 & 2 & 2 \end{pmatrix}$ Given.

$= \det\begin{pmatrix} 12 & 6 & 0 \\ -1 & -1 & 1 \\ -2 & -4 & 2 \end{pmatrix}$ Combination rule: row 1 subtracted from the others.

$= 6\det\begin{pmatrix} 2 & 1 & 0 \\ -1 & -1 & 1 \\ -2 & -4 & 2 \end{pmatrix}$ Multiply rule.

$= 6\det\begin{pmatrix} 0 & -1 & 2 \\ -1 & -1 & 1 \\ 0 & -3 & 2 \end{pmatrix}$ Combination rule: add row 1 to row 3, then add twice row 2 to row 1.

$= -6\det\begin{pmatrix} -1 & -1 & 1 \\ 0 & -1 & 2 \\ 0 & -3 & 2 \end{pmatrix}$ Swap rule.

$= 6\det\begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 2 \\ 0 & -3 & 2 \end{pmatrix}$ Multiply rule.

$= 6\det\begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 2 \\ 0 & 0 & -4 \end{pmatrix}$ Combination rule.

$= 6(1)(-1)(-4)$ Triangular rule.

$= 24$ Formula verified.

3 Example (Hybrid Method) Justify by cofactor expansion and the four properties the identity

$$\det\begin{pmatrix} 10 & 5 & 0 \\ 11 & 5 & a \\ 10 & 2 & b \end{pmatrix} = 5(6a - b).$$

Solution: Let D denote the value of the determinant. Then

$D = \det\begin{pmatrix} 10 & 5 & 0 \\ 11 & 5 & a \\ 10 & 2 & b \end{pmatrix}$ Given.

$= \det\begin{pmatrix} 10 & 5 & 0 \\ 1 & 0 & a \\ 0 & -3 & b \end{pmatrix}$ Combination: subtract row 1 from the other rows.

$$= \det \begin{pmatrix} 0 & 5 & -10a \\ 1 & 0 & a \\ 0 & -3 & b \end{pmatrix}$$ Combination: add -10 times row 2 to row 1.

$$= (1)(-1) \det \begin{pmatrix} 5 & -10a \\ -3 & b \end{pmatrix}$$ Cofactor expansion on column 1.

$$= (1)(-1)(5b - 30a)$$ Sarrus' rule for $n = 2$.

$$= 5(6a - b).$$ Formula verified.

4 Example (Cramer's Rule) Solve by Cramer's rule the system of equations

$$\begin{array}{rcrcrcrcr} 2x_1 & + & 3x_2 & + & x_3 & - & x_4 & = & 1, \\ x_1 & + & x_2 & - & & & x_4 & = & -1, \\ & & 3x_2 & + & x_3 & + & x_4 & = & 3, \\ x_1 & + & & & x_3 & - & x_4 & = & 0, \end{array}$$

verifying $x_1 = 1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$.

Solution: Form the four determinants $\Delta_1, \ldots, \Delta_4$ from the base determinant Δ as follows:

$$\Delta = \det \begin{pmatrix} 2 & 3 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ 0 & 3 & 1 & 1 \\ 1 & 0 & 1 & -1 \end{pmatrix},$$

$$\Delta_1 = \det \begin{pmatrix} \mathbf{1} & 3 & 1 & -1 \\ -\mathbf{1} & 1 & 0 & -1 \\ \mathbf{3} & 3 & 1 & 1 \\ \mathbf{0} & 0 & 1 & -1 \end{pmatrix}, \quad \Delta_2 = \det \begin{pmatrix} 2 & \mathbf{1} & 1 & -1 \\ 1 & -\mathbf{1} & 0 & -1 \\ 0 & \mathbf{3} & 1 & 1 \\ 1 & \mathbf{0} & 1 & -1 \end{pmatrix},$$

$$\Delta_3 = \det \begin{pmatrix} 2 & 3 & \mathbf{1} & -1 \\ 1 & 1 & -\mathbf{1} & -1 \\ 0 & 3 & \mathbf{3} & 1 \\ 1 & 0 & \mathbf{0} & -1 \end{pmatrix}, \quad \Delta_4 = \det \begin{pmatrix} 2 & 3 & 1 & \mathbf{1} \\ 1 & 1 & 0 & -\mathbf{1} \\ 0 & 3 & 1 & \mathbf{3} \\ 1 & 0 & 1 & \mathbf{0} \end{pmatrix}.$$

Five repetitions of the methods used in the previous examples give the answers $\Delta = -2$, $\Delta_1 = -2$, $\Delta_2 = 0$, $\Delta_3 = -2$, $\Delta_4 = -4$, therefore Cramer's rule implies the solution $x_i = \Delta_i/\Delta$, $1 \le i \le 4$. Then $x_1 = 1$, $x_2 = 0$, $x_3 = 1$, $x_4 = 2$.

Maple code. The details of the computation above can be checked in computer algebra system `maple` as follows.

```
with(linalg):
A:=matrix([
  [2,  3,  1, -1], [1,  1,  0, -1],
  [0,  3,  1,  1], [1,  0,  1, -1]]);
Delta:= det(A);
B1:=matrix([
  [ 1,  3,  1, -1], [-1,  1,  0, -1],
  [ 3,  3,  1,  1], [ 0,  0,  1, -1]]);
Delta1:=det(B1);
x[1]:=Delta1/Delta;
```

An Applied Definition of Determinant

To be developed here is another way to look at formula (9), which emphasizes the column and row structure of a determinant. The definition, which agrees with (9), leads to a short proof of the four properties, which are used to find the value of any determinant.

Permutation Matrices. A matrix P obtained from the identity matrix I by swapping rows is called a **permutation matrix**. There are $n!$ permutation matrices. To illustrate, the 3×3 permutation matrices are

$$\begin{pmatrix} 1&0&0\\0&1&0\\0&0&1 \end{pmatrix}, \begin{pmatrix} 1&0&0\\0&0&1\\0&1&0 \end{pmatrix}, \begin{pmatrix} 0&1&0\\1&0&0\\0&0&1 \end{pmatrix}, \begin{pmatrix} 0&1&0\\0&0&1\\1&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&1\\1&0&0\\0&1&0 \end{pmatrix}, \begin{pmatrix} 0&0&1\\0&1&0\\1&0&0 \end{pmatrix}.$$

Define for a permutation matrix P the determinant by

$$\det(P) = (-1)^k$$

where k is the least number of row swaps required to convert P to the identity. The number k satisfies $r = k + 2m$, where r is *any* count of row swaps that changes P to the identity, and m is some integer. Therefore, $\det(P) = (-1)^k = (-1)^r$. In the illustration, the corresponding determinants are $1, -1, -1, 1, 1, -1$, as computed from $\det(P) = (-1)^r$, where r row swaps change P into I.

It can be verified that $\det(P)$ agrees with the value reported by formula (9). Each σ in (9) corresponds to a permutation matrix P with rows arranged in the order specified by σ. The summation in (9) for $A = P$ has exactly one nonzero term.

Sampled–product. Let be given an $n \times n$ matrix A and an $n \times n$ permutation matrix P. The matrix P has ones (1) in exactly n locations. The sampled–product $A.P$ uses these locations to select entries from the matrix A, whereupon $A.P$ is the product of those entries. More precisely, let $\vec{A}_1, \ldots, \vec{A}_n$ be the rows of A and let $\vec{P}_1, \ldots, \vec{P}_n$ be the rows of P. Define via the normal dot product $(\cdot)$ the **sampled–product**

$$A.P = (A_1 \cdot P_1)(A_2 \cdot P_2) \cdots (A_n \cdot P_n) \\ = a_{1\sigma_1} \cdots a_{n\sigma_n}, \tag{15}$$

where the rows of P are rows $\sigma_1, \ldots, \sigma_n$ of I. The formula implies that $A.P$ is a linear function of the rows of A. A similar construction shows $A.P$ is a linear function of the columns of A.

Applied determinant formula. The determinant is defined by

$$\det(A) = \textstyle\sum_P \det(P)\, A.P, \tag{16}$$

where the summation extends over all possible permutation matrices P. The definition emphasizes the explicit linear dependence of the determinant upon the rows of A (or the columns of A). A tedious but otherwise routine justification shows that (9) and (16) give the same result.

Verification of the Four Properties:

Triangular. If A is $n \times n$ triangular, then in (16) appears only one nonzero term, due to zero factors in the product $A.P$. The term that appears corresponds to P=identity, therefore $A.P$ is the product of the diagonal elements of A. Since $\det(P) = \det(I) = 1$, the result follows. A similar proof can be constructed from determinant definition (9).

Swap. Let Q be obtained from I by swapping rows i and j. Let $B = QA$, so that B is A with rows i and j swapped. We must show $\det(A) = -\det(B)$. Observe that $B.P = A.QP$ and $\det(QP) = -\det(P)$. The matrices QP over all possible P simply relist all permutation matrices, hence definition (16) implies the result.

Combination. Let matrix B be obtained from matrix A by adding to row j the vector k times row i ($i \neq j$). Then $\mathbf{row}(B, j) = \mathbf{row}(A, j) + k\,\mathbf{row}(A, i)$ and $B.P = (B_1 \cdot P) \cdots (B_n \cdot P) = A.P + k\,C.P$, where C is the matrix obtained from A by *replacing* $\mathbf{row}(A, j)$ with $\mathbf{row}(A, i)$. Then C has equal rows $\mathbf{row}(C, i) = \mathbf{row}(C, j) = \mathbf{row}(A, i)$. By the swap rule applied to rows i and j, $\det(C) = -\det(C)$, or $\det(C) = 0$. Add on P across $B.P = A.P + k\,C.P$ to obtain $\det(B) = \det(A) + k\det(C)$. This implies $\det(B) = \det(A)$.

Multiply. Let matrices A and B have the same rows, except for some index i, $\mathbf{row}(B, i) = c\,\mathbf{row}(A, i)$. Then $B.P = c\,A.P$. Add on P across this equation to obtain $\det(B) = c\det(A)$.

Verification of the Additional Rules:

Duplicate rows. The swap rule applies to the two duplicate rows to give $\det(A) = -\det(A)$, hence $\det(A) = 0$.

Zero row. Apply the common factor rule with $c = 2$, possible since the row has all zero entries. Then $\det(A) = 2\det(A)$, giving $\det(A) = 0$.

Common factor and row linearity. The sampled–product $A.P$ is a linear function of each row, therefore the same is true of $\det(A)$.

Derivations: Cofactors and Cramer's Rule

Derivation of cofactor expansion (10): The column expansion formula is derived from the row expansion formula applied to the transpose. We consider only the derivation of the row expansion formula (10) for $k = 1$, because the case for general k is the same except for notation. The plan is to establish equality of the two sides of (10) for $k = 1$, which in terms of $\mathbf{minor}(A, 1, j) = (-1)^{1+j}\,\mathbf{cof}(A, 1, j)$ is the equality

$$\det(A) = \sum_{j=1}^{n} a_{1j} (-1)^{1+j}\,\mathbf{minor}(A, 1, j). \tag{17}$$

The details require expansion of $\textbf{minor}(A,1,j)$ in (17) via the definition of determinant $\det(A) = \sum_\sigma (-1)^{\text{parity}(\sigma)} a_{1\sigma_1} \cdots a_{n\sigma_n}$. A typical term on the right in (17) after expansion looks like

$$a_{1j}\,(-1)^{1+j}(-1)^{\text{parity}(\alpha)} a_{2\alpha_2} \cdots a_{n\alpha_n}.$$

Here, α is a rearrangement of the set of $n-1$ elements consisting of $1, \ldots, j-1, j+1, \ldots, n$. Define $\sigma = (j, \alpha_2, \ldots, \alpha_n)$, which is a rearrangement of symbols $1, \ldots, n$. After parity(α) interchanges, α is changed into $(1, \ldots, j-1, j+1, \ldots, n)$ and therefore these same interchanges transform σ into $(j, 1, \ldots, j-1, j+1, \ldots, n)$. An additional $j-1$ interchanges will transform σ into natural order $(1, \ldots, n)$. This establishes, because of $(-1)^{j-1} = (-1)^{j+1}$, the identity

$$\begin{aligned} (-1)^{\text{parity}(\sigma)} &= (-1)^{j-1+\text{parity}(\alpha)} \\ &= (-1)^{j+1+\text{parity}(\alpha)}. \end{aligned}$$

Collecting formulas gives

$$(-1)^{\text{parity}(\sigma)} a_{1\sigma_1} \cdots a_{n\sigma_n} = a_{1j}\,(-1)^{1+j}(-1)^{\text{parity}(\alpha)} a_{2\alpha_2} \cdots a_{n\alpha_n}.$$

Adding across this formula over all α and j gives a sum on the right which matches the right side of (17). Some additional thought reveals that the terms on the left add exactly to $\det(A)$, hence (17) is proved.

Derivation of Cramer's Rule: The cofactor column expansion theory implies that the Cramer's rule solution of $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ is given by

$$x_j = \frac{\Delta_j}{\Delta} = \frac{1}{\Delta}\sum_{k=1}^{n} b_k\, \textbf{cof}(A,k,j). \tag{18}$$

We will verify that $A\vec{\mathbf{x}} = \vec{\mathbf{b}}$. Let $E_1, \ldots, E_n$ be the rows of the identity matrix. The question reduces to showing that $E_pA\vec{\mathbf{x}} = b_p$. The details will use the fact

$$\sum_{j=1}^{n} a_{pj}\, \textbf{cof}(A,k,j) = \begin{cases} \det(A) & \text{for } k = p, \\ 0 & \text{for } k \neq p, \end{cases} \tag{19}$$

Equation (19) follows by cofactor row expansion, because the sum on the left is $\det(B)$ where B is matrix A with row k replaced by row p. If B has two equal rows, then $\det(B) = 0$; otherwise, $B = A$ and $\det(B) = \det(A)$.

$$\begin{aligned} E_pA\vec{\mathbf{x}} &= \sum_{j=1}^{n} a_{pj} x_j \\ &= \frac{1}{\Delta}\sum_{j=1}^{n} a_{pj} \sum_{k=1}^{n} b_k\, \textbf{cof}(A,k,j) && \text{Apply formula (18).} \\ &= \frac{1}{\Delta}\sum_{k=1}^{n} b_k \left(\sum_{j=1}^{n} a_{pj}\, \textbf{cof}(A,k,j)\right) && \text{Switch order of summation.} \\ &= b_p && \text{Apply (19).} \end{aligned}$$

Derivation of $A \cdot \mathbf{adj}(A) = \det(A)I$: The proof uses formula (19). Consider column k of $\mathbf{adj}(A)$, denoted $\vec{X}$, multiplied against matrix A, which gives

$$A\vec{X} = \begin{pmatrix} \sum_{j=1}^n a_{1j}\,\mathbf{cof}(A,k,j) \\ \sum_{j=1}^n a_{2j}\,\mathbf{cof}(A,k,j) \\ \vdots \\ \sum_{j=1}^n a_{nj}\,\mathbf{cof}(A,k,j) \end{pmatrix}.$$

By formula (19),

$$\sum_{j=1}^n a_{ij}\,\mathbf{cof}(A,k,j) = \begin{cases} \det(A) & i = k, \\ 0 & i \neq k. \end{cases}$$

Therefore, $A\vec{X}$ is $\det(A)$ times column k of the identity I. This completes the proof.

Three Properties that Define a Determinant

Write the determinant $\det(A)$ in terms of the rows $A_1, \ldots, A_n$ of the matrix A as follows:

$$D_1(A_1, \ldots, A_n) = \sum_P \det(P)\, A.P.$$

Already known is that $D_1(A_1, \ldots, A_n)$ is a function D that satisfies the following three properties:

Linearity	D is linear in each argument $A_1, \ldots, A_n$.
Swap	D changes sign if two arguments are swapped. Equivalently, $D = 0$ if two arguments are equal.
Identity	$D = 1$ when $A = I$.

The equivalence reported in **swap** is obtained by expansion, e.g., for $n = 2$, $A_1 = A_2$ implies $D(A_1, A_2) = -D(A_1, A_2)$ and hence $D = 0$. Similarly, $D(A_1 + A_2, A_1 + A_2) = 0$ implies by linearity that $D(A_1, A_2) = -D(A_2, A_1)$, which is the swap property for $n = 2$.

It is less obvious that *the three properties uniquely define the determinant*, that is:

Theorem 12 (Uniqueness)
If $D(A_1, \ldots, A_n)$ satisfies the properties of **linearity**, **swap** and **identity**, then $D(A_1, \ldots, A_n) = \det(A)$.

Proof: The rows of the identity matrix I are denoted $E_1, \ldots, E_n$, so that for $1 \leq j \leq n$ we may write the expansion

$$A_j = a_{j1}E_1 + a_{j2}E_2 + \cdots + a_{jn}E_n. \tag{20}$$

We illustrate the proof for the case $n = 2$:

$$
\begin{aligned}
D(A_1, A_2) &= D(a_{11}E_1 + a_{12}E_2, A_2) && \text{By (20).}\\
&= a_{11}D(E_1, A_2) + a_{12}D(E_2, A_2) && \text{By linearity.}\\
&= a_{11}a_{22}D(E_1, E_2) + a_{11}a_{21}D(E_1, E_1) && \text{Repeat for } A_2.\\
&\quad + a_{12}a_{21}D(E_2, E_1) + a_{12}a_{22}D(E_2, E_2)
\end{aligned}
$$

The swap and identity properties give $D(E_1, E_1) = D(E_2, E_2) = 0$ and $1 = D(E_1, E_2) = -D(E_2, E_1)$. Therefore, $D(A_1, A_2) = a_{11}a_{22} - a_{12}a_{21}$ and this implies that $D(A_1, A_2) = \det(A)$.

The proof for general n depends upon the identity

$$
\begin{aligned}
D(E_{\sigma_1}, \ldots, E_{\sigma_n}) &= (-1)^{\text{parity}(\sigma)} D(E_1, \ldots, E_n)\\
&= (-1)^{\text{parity}(\sigma)}
\end{aligned}
$$

where $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a rearrangement of the integers $1, \ldots, n$. This identity is implied by the **swap** and **identity** properties. Then, as in the case $n = 2$, **linearity** implies that

$$
\begin{aligned}
D(A_1, \ldots, A_n) &= \textstyle\sum_\sigma a_{1\sigma_1} \cdots a_{n\sigma_n} D(E_{\sigma_1}, \ldots, E_{\sigma_n})\\
&= \textstyle\sum_\sigma (-1)^{\text{parity}(\sigma)} a_{1\sigma_1} \cdots a_{n\sigma_n}\\
&= \det(A).
\end{aligned}
$$

Exercises 5.3

Determinant Notation. Write formulae for x and y as quotients of 2×2 determinants. Do not evaluate the determinants!

1. $\begin{pmatrix} 1 & -1 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -10 \\ 3 \end{pmatrix}$

Unique Solution of a 2×2 System. Solve $A\mathbf{X} = \mathbf{b}$ for $\mathbf{X}$ using Cramer's rule for 2×2 systems.

2. $A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

Sarrus' 2×2 rule. Evaluate $\det(A)$.

3. $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

Sarrus' rule 3×3. Evaluate $\det(A)$.

4. $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$

Definition of Determinant.

5. Let A be 3×3 with first row all zeros. Use the college algebra definition of determinant to show that $\det(A) = 0$.

Four Properties. Evaluate $\det(A)$ using the four properties for determinants, page **??**.

6. $A = \begin{pmatrix} -1 & 5 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & -3 \end{pmatrix}$

Elementary Matrices and the Four Rules. Find $\det(A)$.

7. A is 3×3 and obtained from the identity matrix I by three row swaps.

8. A is 7×7, obtained from I by swapping rows 1 and 2, then rows 4 and 1, then rows 1 and 3.

More Determinant Rules. Cite the determinant rule that verifies $\det(A) = 0$. **Never** expand $\det(A)$!

9. $A = \begin{pmatrix} -1 & 5 & 1 \\ 2 & -4 & -4 \\ 1 & 1 & -3 \end{pmatrix}$

Cofactor Expansion and College Algebra. Evaluate $\det(A)$ using the most efficient cofactor expansion.

10. $A = \begin{pmatrix} 2 & 5 & 1 \\ 2 & 0 & -4 \\ 1 & 0 & 0 \end{pmatrix}$

Minors and Cofactors. Write out and then evaluate the minor and cofactor of each element cited for the matrix $A = \begin{pmatrix} 2 & 5 & y \\ x & -1 & -4 \\ 1 & 2 & z \end{pmatrix}$

11. Row 1 and column 3.

Cofactor Expansion. Use cofactors to evaluate $\det(A)$.

12. $A = \begin{pmatrix} 2 & 7 & 1 \\ -1 & 0 & -4 \\ 1 & 0 & 3 \end{pmatrix}$

Adjugate and Inverse Matrix. Find the adjugate of A and the inverse of A. Check the answers via the formula $A\,\mathbf{adj}(A) = \det(A)I$.

13. $A = \begin{pmatrix} 2 & 7 \\ -1 & 0 \end{pmatrix}$

14. $A = \begin{pmatrix} 5 & 1 & 1 \\ 0 & 0 & 2 \\ 1 & 0 & 3 \end{pmatrix}$

Elementary Matrices. Find the determinant of the product A of elementary matrices.

15. Let $A = E_1E_2$ be 9×9, where E_1 multiplies row 3 of the identity by -7 and E_2 swaps rows 3 and 5 of the identity.

Determinants and Invertible Matrices. Test for invertibility of A. If invertible, then find A^{-1} by the best method: **rref** method or the adjugate formula.

16. $A = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 0 & 2 \\ 1 & 0 & 4 \end{pmatrix}$

Determinant Product Rule. Apply the product rule $\det(AB) = \det(A)\det(B)$.

17. Let $\det(A) = 5$ and $\det(B) = -2$. Find $\det(A^2B^3)$.

18. Let $\det(A) = 4$ and $A(B - 2A) = 0$. Find $\det(B)$.

19. Let $A = E_1E_2E_3$ where E_1, E_2 are elementary swap matrices and E_3 is an elementary combination matrix. Find $\det(A)$.

20. Assume $\det(AB + A) = 0$ and $\det(A) \neq 0$. Show that $\det(B + I) = 0$.

Cayley-Hamilton Theorem.

21. Let $\lambda^2 - 2\lambda + 1 = 0$ be the characteristic equation of a matrix A. Find a formula for A^2 in terms of A and I.

22. Let A be an $n \times n$ triangular matrix with all diagonal entries zero. Prove that $A^n = 0$.

Applied Definition of Determinant.

23. Given A, find the sampled product for the permutation matrix P.

$$A = \begin{pmatrix} 5 & 3 & 1 \\ 0 & 5 & 7 \\ 1 & 9 & 4 \end{pmatrix},$$

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

24. Determine the permutation matrices P required to evaluate $\det(A)$ when A is 4×4.

Three Properties.

25. Assume $n = 3$. Prove that the three properties imply $D = 0$ when two rows are identical.

26. Assume $n = 3$. Prove that the three properties imply $D = 0$ when a row is zero.