3.2 Determinants and Matrix Inverses

In this section, several theorems about determinants are derived. One consequence of these theorems is that a square matrix A is invertible if and only if $\det A \neq 0$. Moreover, determinants are used to give a formula for A^{-1} which, in turn, yields a formula (called Cramer's rule) for the solution of any system of linear equations with an invertible coefficient matrix.

We begin with a remarkable theorem (due to Cauchy in 1812) about the determinant of a product of matrices. The proof is given at the end of this section.

Theorem 3.2.1: Product Theorem

If A and B are $n \times n$ matrices, then $\det(AB) = \det A \det B$.

The complexity of matrix multiplication makes the product theorem quite unexpected. Here is an example where it reveals an important numerical identity.

Example 3.2.1

If $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ and $B = \begin{bmatrix} c & d \\ -d & c \end{bmatrix}$ then $AB = \begin{bmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{bmatrix}$.

Hence $\det A \det B = \det(AB)$ gives the identity

$$(a^2+b^2)(c^2+d^2) = (ac-bd)^2 + (ad+bc)^2$$

Theorem 3.2.1 extends easily to $\det(ABC) = \det A \det B \det C$. In fact, induction gives

$$\det(A_1A_2\cdots A_{k-1}A_k) = \det A_1 \det A_2 \cdots \det A_{k-1} \det A_k$$

for any square matrices $A_1, \dots, A_k$ of the same size. In particular, if each $A_i = A$, we obtain

$$\det(A^k) = (detA)^k, \text{ for any } k \geq 1$$

We can now give the invertibility condition.

Theorem 3.2.2

An $n \times n$ matrix A is invertible if and only if $\det A \neq 0$. *When this is the case,* $\det(A^{-1}) = \frac{1}{\det A}$

Proof. If A is invertible, then $AA^{-1} = I$; so the product theorem gives

$$1 = \det I = \det(AA^{-1}) = \det A \det A^{-1}$$

Hence, $\det A \neq 0$ and also $\det A^{-1} = \frac{1}{\det A}$.

Conversely, if $\det A \neq 0$, we show that A can be carried to I by elementary row operations (and invoke Theorem 2.4.5). Certainly, A can be carried to its reduced row-echelon form R, so $R = E_k \cdots E_2 E_1 A$ where the E_i are elementary matrices (Theorem 2.5.1). Hence the product theorem gives

$$\det R = \det E_k \cdots \det E_2 \det E_1 \det A$$

Since $\det E \neq 0$ for all elementary matrices E, this shows $\det R \neq 0$. In particular, R has no row of zeros, so $R = I$ because R is square and reduced row-echelon. This is what we wanted. □

Example 3.2.2

For which values of c does $A = \left[\begin{array}{rrr} 1 & 0 & -c \\ -1 & 3 & 1 \\ 0 & 2c & -4 \end{array} \right]$ have an inverse?

Solution. Compute $\det A$ by first adding c times column 1 to column 3 and then expanding along row 1.

$$\det A = \det \left[\begin{array}{rrr} 1 & 0 & -c \\ -1 & 3 & 1 \\ 0 & 2c & -4 \end{array} \right] = \det \left[\begin{array}{rrr} 1 & 0 & 0 \\ -1 & 3 & 1-c \\ 0 & 2c & -4 \end{array} \right] = 2(c+2)(c-3)$$

Hence, $\det A = 0$ if $c = -2$ or $c = 3$, and A has an inverse if $c \neq -2$ and $c \neq 3$.

Example 3.2.3

If a product $A_1 A_2 \cdots A_k$ of square matrices is invertible, show that each A_i is invertible.

Solution. We have $\det A_1 \det A_2 \cdots \det A_k = \det (A_1 A_2 \cdots A_k)$ by the product theorem, and $\det (A_1 A_2 \cdots A_k) \neq 0$ by Theorem 3.2.2 because $A_1 A_2 \cdots A_k$ is invertible. Hence

$$\det A_1 \det A_2 \cdots \det A_k \neq 0$$

so $\det A_i \neq 0$ for each i. This shows that each A_i is invertible, again by Theorem 3.2.2.

Theorem 3.2.3

If A is any square matrix, $\det A^T = \det A$.

Proof. Consider first the case of an elementary matrix E. If E is of type I or II, then $E^T = E$; so certainly $\det E^T = \det E$. If E is of type III, then E^T is also of type III; so $\det E^T = 1 = \det E$ by Theorem 3.1.2. Hence, $\det E^T = \det E$ for every elementary matrix E.

Now let A be any square matrix. If A is not invertible, then neither is A^T; so $\det A^T = 0 = \det A$ by Theorem 3.2.2. On the other hand, if A is invertible, then $A = E_k \cdots E_2 E_1$, where the E_i are elementary matrices (Theorem 2.5.2). Hence, $A^T = E_1^T E_2^T \cdots E_k^T$ so the product theorem gives

$$\begin{aligned}\det A^T = \det E_1^T \det E_2^T \cdots \det E_k^T &= \det E_1 \det E_2 \cdots \det E_k \\ &= \det E_k \cdots \det E_2 \det E_1 \\ &= \det A\end{aligned}$$

This completes the proof. □

Example 3.2.4

If $\det A = 2$ and $\det B = 5$, calculate $\det(A^3B^{-1}A^TB^2)$.

Solution. We use several of the facts just derived.

$$\begin{aligned}\det(A^3B^{-1}A^TB^2) &= \det(A^3)\det(B^{-1})\det(A^T)\det(B^2) \\ &= (\det A)^3 \frac{1}{\det B} \det A (\det B)^2 \\ &= 2^3 \cdot \tfrac{1}{5} \cdot 2 \cdot 5^2 \\ &= 80\end{aligned}$$

Example 3.2.5

A square matrix is called **orthogonal** if $A^{-1} = A^T$. What are the possible values of $\det A$ if A is orthogonal?

Solution. If A is orthogonal, we have $I = AA^T$. Take determinants to obtain

$$1 = \det I = \det(AA^T) = \det A \det A^T = (\det A)^2$$

Since $\det A$ is a number, this means $\det A = \pm 1$.

Hence Theorems 2.6.4 and 2.6.5 imply that rotation about the origin and reflection about a line through the origin in $\mathbb{R}^2$ have orthogonal matrices with determinants 1 and -1 respectively. In fact they are the *only* such transformations of $\mathbb{R}^2$. We have more to say about this in Section 8.2.

Adjugates

In Section 2.4 we defined the adjugate of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to be $\text{adj}(A) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Then we verified that $A(\text{adj } A) = (\det A)I = (\text{adj } A)A$ and hence that, if $\det A \neq 0$, $A^{-1} = \frac{1}{\det A} \text{adj } A$. We are now able to define the adjugate of an arbitrary square matrix and to show that this formula for the inverse remains valid (when the inverse exists).

Recall that the (i, j)-cofactor $c_{ij}(A)$ of a square matrix A is a number defined for each position (i, j) in the matrix. If A is a square matrix, the **cofactor matrix of** A is defined to be the matrix $\left[c_{ij}(A)\right]$ whose (i, j)-entry is the (i, j)-cofactor of A.

Definition 3.3 Adjugate of a Matrix

The ***adjugate***[4] *of* A, *denoted* $\text{adj}\,(A)$, *is the transpose of this cofactor matrix; in symbols,*

$$\text{adj}\,(A) = \left[c_{ij}(A)\right]^T$$

This agrees with the earlier definition for a 2×2 matrix A as the reader can verify.

Example 3.2.6

Compute the adjugate of $A = \left[\begin{array}{rrr} 1 & 3 & -2 \\ 0 & 1 & 5 \\ -2 & -6 & 7 \end{array}\right]$ and calculate $A(\text{adj}\,A)$ and $(\text{adj}\,A)A$.

Solution. We first find the cofactor matrix.

$$\left[\begin{array}{ccc} c_{11}(A) & c_{12}(A) & c_{13}(A) \\ c_{21}(A) & c_{22}(A) & c_{23}(A) \\ c_{31}(A) & c_{32}(A) & c_{33}(A) \end{array}\right] = \left[\begin{array}{ccc} \left|\begin{array}{rr} 1 & 5 \\ -6 & 7 \end{array}\right| & -\left|\begin{array}{rr} 0 & 5 \\ -2 & 7 \end{array}\right| & \left|\begin{array}{rr} 0 & 1 \\ -2 & -6 \end{array}\right| \\ & & \\ -\left|\begin{array}{rr} 3 & -2 \\ -6 & 7 \end{array}\right| & \left|\begin{array}{rr} 1 & -2 \\ -2 & 7 \end{array}\right| & -\left|\begin{array}{rr} 1 & 3 \\ -2 & -6 \end{array}\right| \\ & & \\ \left|\begin{array}{rr} 3 & -2 \\ 1 & 5 \end{array}\right| & -\left|\begin{array}{rr} 1 & -2 \\ 0 & 5 \end{array}\right| & \left|\begin{array}{rr} 1 & 3 \\ 0 & 1 \end{array}\right| \end{array}\right]$$

$$= \left[\begin{array}{rrr} 37 & -10 & 2 \\ -9 & 3 & 0 \\ 17 & -5 & 1 \end{array}\right]$$

Then the adjugate of A is the transpose of this cofactor matrix.

$$\text{adj}\,A = \left[\begin{array}{rrr} 37 & -10 & 2 \\ -9 & 3 & 0 \\ 17 & -5 & 1 \end{array}\right]^T = \left[\begin{array}{rrr} 37 & -9 & 17 \\ -10 & 3 & -5 \\ 2 & 0 & 1 \end{array}\right]$$

The computation of $A(\text{adj}\,A)$ gives

$$A(\text{adj}\,A) = \left[\begin{array}{rrr} 1 & 3 & -2 \\ 0 & 1 & 5 \\ -2 & -6 & 7 \end{array}\right] \left[\begin{array}{rrr} 37 & -9 & 17 \\ -10 & 3 & -5 \\ 2 & 0 & 1 \end{array}\right] = \left[\begin{array}{rrr} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{array}\right] = 3I$$

and the reader can verify that also $(\text{adj}\,A)A = 3I$. Hence, analogy with the 2×2 case would indicate that $\det A = 3$; this is, in fact, the case.

The relationship $A(\text{adj}\,A) = (\det A)I$ holds for any square matrix A. To see why this is so, consider

[4]This is also called the classical adjoint of A, but the term "adjoint" has another meaning.

the general 3×3 case. Writing $c_{ij}(A) = c_{ij}$ for short, we have

$$\text{adj } A = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}^T = \begin{bmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{bmatrix}$$

If $A = \left[a_{ij}\right]$ in the usual notation, we are to verify that $A(\text{adj } A) = (\det A)I$. That is,

$$A(\text{adj } A) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} c_{11} & c_{21} & c_{31} \\ c_{12} & c_{22} & c_{32} \\ c_{13} & c_{23} & c_{33} \end{bmatrix} = \begin{bmatrix} \det A & 0 & 0 \\ 0 & \det A & 0 \\ 0 & 0 & \det A \end{bmatrix}$$

Consider the $(1,\ 1)$-entry in the product. It is given by $a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13}$, and this is just the cofactor expansion of $\det A$ along the first row of A. Similarly, the $(2,\ 2)$-entry and the $(3,\ 3)$-entry are the cofactor expansions of $\det A$ along rows 2 and 3, respectively.

So it remains to be seen why the off-diagonal elements in the matrix product $A(\text{adj } A)$ are all zero. Consider the $(1,\ 2)$-entry of the product. It is given by $a_{11}c_{21} + a_{12}c_{22} + a_{13}c_{23}$. This *looks* like the cofactor expansion of the determinant of *some* matrix. To see which, observe that c_{21}, c_{22}, and c_{23} are all computed by *deleting* row 2 of A (and one of the columns), so they remain the same if row 2 of A is changed. In particular, if row 2 of A is replaced by row 1, we obtain

$$a_{11}c_{21} + a_{12}c_{22} + a_{13}c_{23} = \det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = 0$$

where the expansion is along row 2 and where the determinant is zero because two rows are identical. A similar argument shows that the other off-diagonal entries are zero.

This argument works in general and yields the first part of Theorem 3.2.4. The second assertion follows from the first by multiplying through by the scalar $\frac{1}{\det A}$.

Theorem 3.2.4: Adjugate Formula

If A is any square matrix, then

$$A(\text{adj } A) = (\det A)I = (\text{adj } A)A$$

In particular, if $\det A \neq 0$, the inverse of A is given by

$$A^{-1} = \frac{1}{\det A} \text{adj } A$$

It is important to note that this theorem is *not* an efficient way to find the inverse of the matrix A. For example, if A were 10×10, the calculation of $\text{adj } A$ would require computing $10^2 = 100$ determinants of 9×9 matrices! On the other hand, the matrix inversion algorithm would find A^{-1} with about the same effort as finding $\det A$. Clearly, Theorem 3.2.4 is not a *practical* result: its virtue is that it gives a formula for A^{-1} that is useful for *theoretical* purposes.

Example 3.2.7

Find the $(2,\ 3)$-entry of A^{-1} if $A = \left[\begin{array}{rrr} 2 & 1 & 3 \\ 5 & -7 & 1 \\ 3 & 0 & -6 \end{array} \right]$.

Solution. First compute

$$\det A = \left| \begin{array}{rrr} 2 & 1 & 3 \\ 5 & -7 & 1 \\ 3 & 0 & -6 \end{array} \right| = \left| \begin{array}{rrr} 2 & 1 & 7 \\ 5 & -7 & 11 \\ 3 & 0 & 0 \end{array} \right| = 3 \left| \begin{array}{rr} 1 & 7 \\ -7 & 11 \end{array} \right| = 180$$

Since $A^{-1} = \frac{1}{\det A} \operatorname{adj} A = \frac{1}{180} \left[c_{ij}(A) \right]^T$, the $(2,\ 3)$-entry of A^{-1} is the $(3,\ 2)$-entry of the matrix $\frac{1}{180} \left[c_{ij}(A) \right]$; that is, it equals $\frac{1}{180} c_{32}(A) = \frac{1}{180} \left(- \left| \begin{array}{rr} 2 & 3 \\ 5 & 1 \end{array} \right| \right) = \frac{13}{180}$.

Example 3.2.8

If A is $n \times n$, $n \geq 2$, show that $\det (\operatorname{adj} A) = (\det A)^{n-1}$.

Solution. Write $d = \det A$; we must show that $\det (\operatorname{adj} A) = d^{n-1}$. We have $A(\operatorname{adj} A) = dI$ by Theorem 3.2.4, so taking determinants gives $d \det (\operatorname{adj} A) = d^n$. Hence we are done if $d \neq 0$. Assume $d = 0$; we must show that $\det (\operatorname{adj} A) = 0$, that is, $\operatorname{adj} A$ is not invertible. If $A \neq 0$, this follows from $A(\operatorname{adj} A) = dI = 0$; if $A = 0$, it follows because then $\operatorname{adj} A = 0$.

Cramer's Rule

Theorem 3.2.4 has a nice application to linear equations. Suppose

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

is a system of n equations in n variables $x_1,\ x_2,\ \dots,\ x_n$. Here A is the $n \times n$ coefficient matrix, and $\mathbf{x}$ and $\mathbf{b}$ are the columns

$$\mathbf{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \right] \text{ and } \mathbf{b} = \left[\begin{array}{c} b_1 \\ b_2 \\ \vdots \\ b_n \end{array} \right]$$

of variables and constants, respectively. If $\det A \neq 0$, we left multiply by A^{-1} to obtain the solution $\mathbf{x} = A^{-1}\mathbf{b}$. When we use the adjugate formula, this becomes

$$\left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \right] = \frac{1}{\det A} (\operatorname{adj} A)\mathbf{b}$$

$$= \frac{1}{\det A} \begin{bmatrix} c_{11}(A) & c_{21}(A) & \cdots & c_{n1}(A) \\ c_{12}(A) & c_{22}(A) & \cdots & c_{n2}(A) \\ \vdots & \vdots & & \vdots \\ c_{1n}(A) & c_{2n}(A) & \cdots & c_{nn}(A) \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Hence, the variables $x_1, x_2, \dots, x_n$ are given by

$$\begin{aligned} x_1 &= \tfrac{1}{\det A}\left[b_1c_{11}(A) + b_2c_{21}(A) + \cdots + b_nc_{n1}(A)\right] \\ x_2 &= \tfrac{1}{\det A}\left[b_1c_{12}(A) + b_2c_{22}(A) + \cdots + b_nc_{n2}(A)\right] \\ &\vdots \qquad\qquad \vdots \\ x_n &= \tfrac{1}{\det A}\left[b_1c_{1n}(A) + b_2c_{2n}(A) + \cdots + b_nc_{nn}(A)\right] \end{aligned}$$

Now the quantity $b_1c_{11}(A) + b_2c_{21}(A) + \cdots + b_nc_{n1}(A)$ occurring in the formula for x_1 looks like the cofactor expansion of the determinant of a matrix. The cofactors involved are $c_{11}(A), c_{21}(A), \dots, c_{n1}(A)$, corresponding to the first column of A. If A_1 is obtained from A by replacing the first column of A by $\mathbf{b}$, then $c_{i1}(A_1) = c_{i1}(A)$ for each i because column 1 is deleted when computing them. Hence, expanding $\det(A_1)$ by the first column gives

$$\begin{aligned} \det A_1 &= b_1c_{11}(A_1) + b_2c_{21}(A_1) + \cdots + b_nc_{n1}(A_1) \\ &= b_1c_{11}(A) + b_2c_{21}(A) + \cdots + b_nc_{n1}(A) \\ &= (\det A)x_1 \end{aligned}$$

Hence, $x_1 = \frac{\det A_1}{\det A}$ and similar results hold for the other variables.

Theorem 3.2.5: Cramer's Rule[5]

If A is an invertible $n \times n$ matrix, the solution to the system

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

of n equations in the variables $x_1, x_2, \dots, x_n$ is given by

$$x_1 = \frac{\det A_1}{\det A}, \quad x_2 = \frac{\det A_2}{\det A}, \quad \cdots, \quad x_n = \frac{\det A_n}{\det A}$$

where, for each k, A_k is the matrix obtained from A by replacing column k by $\mathbf{b}$.

Example 3.2.9

Find x_1, given the following system of equations.

$$\begin{aligned} 5x_1 + x_2 - x_3 &= 4 \\ 9x_1 + x_2 - x_3 &= 1 \\ x_1 - x_2 + 5x_3 &= 2 \end{aligned}$$

[5] Gabriel Cramer (1704–1752) was a Swiss mathematician who wrote an introductory work on algebraic curves. He popularized the rule that bears his name, but the idea was known earlier.

Solution. Compute the determinants of the coefficient matrix A and the matrix A_1 obtained from it by replacing the first column by the column of constants.

$$\det A = \det \left[\begin{array}{rrr} 5 & 1 & -1 \\ 9 & 1 & -1 \\ 1 & -1 & 5 \end{array} \right] = -16$$

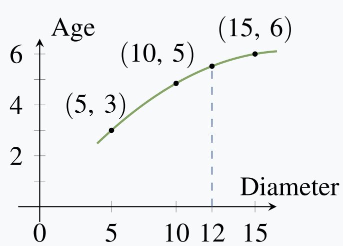
$$\det A_1 = \det \left[\begin{array}{rrr} 4 & 1 & -1 \\ 1 & 1 & -1 \\ 2 & -1 & 5 \end{array} \right] = 12$$

Hence, $x_1 = \frac{\det A_1}{\det A} = -\frac{3}{4}$ by Cramer's rule.

Cramer's rule is *not* an efficient way to solve linear systems or invert matrices. True, it enabled us to calculate x_1 here without computing x_2 or x_3. Although this might seem an advantage, the truth of the matter is that, for large systems of equations, the number of computations needed to find *all* the variables by the gaussian algorithm is comparable to the number required to find *one* of the determinants involved in Cramer's rule. Furthermore, the algorithm works when the matrix of the system is not invertible and even when the coefficient matrix is not square. Like the adjugate formula, then, Cramer's rule is *not* a practical numerical technique; its virtue is theoretical.

Polynomial Interpolation

Example 3.2.10



A forester wants to estimate the age (in years) of a tree by measuring the diameter of the trunk (in cm). She obtains the following data:

	Tree 1	Tree 2	Tree 3
Trunk Diameter	5	10	15
Age	3	5	6

Estimate the age of a tree with a trunk diameter of 12 cm.

Solution.
The forester decides to "fit" a quadratic polynomial

$$p(x) = r_0 + r_1x + r_2x^2$$

to the data, that is choose the coefficients r_0, r_1, and r_2 so that $p(5) = 3$, $p(10) = 5$, and $p(15) = 6$, and then use $p(12)$ as the estimate. These conditions give three linear equations:

$$\begin{aligned} r_0 + 5r_1 + 25r_2 &= 3 \\ r_0 + 10r_1 + 100r_2 &= 5 \\ r_0 + 15r_1 + 225r_2 &= 6 \end{aligned}$$

The (unique) solution is $r_0 = 0$, $r_1 = \frac{7}{10}$, and $r_2 = -\frac{1}{50}$, so

$$p(x) = \frac{7}{10}x - \frac{1}{50}x^2 = \frac{1}{50}x(35-x)$$

Hence the estimate is $p(12) = 5.52$.

As in Example 3.2.10, it often happens that two variables x and y are related but the actual functional form $y = f(x)$ of the relationship is unknown. Suppose that for certain values $x_1, x_2, \dots, x_n$ of x the corresponding values $y_1, y_2, \dots, y_n$ are known (say from experimental measurements). One way to estimate the value of y corresponding to some other value a of x is to find a polynomial[6]

$$p(x) = r_0 + r_1x + r_2x^2 + \cdots + r_{n-1}x^{n-1}$$

that "fits" the data, that is $p(x_i) = y_i$ holds for each $i = 1, 2, \dots, n$. Then the estimate for y is $p(a)$. As we will see, such a polynomial always exists if the x_i are distinct.

The conditions that $p(x_i) = y_i$ are

$$\begin{array}{ccccccccc}
r_0 & + & r_1x_1 & + & r_2x_1^2 & + \cdots + & r_{n-1}x_1^{n-1} & = & y_1 \\
r_0 & + & r_1x_2 & + & r_2x_2^2 & + \cdots + & r_{n-1}x_2^{n-1} & = & y_2 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
r_0 & + & r_1x_n & + & r_2x_n^2 & + \cdots + & r_{n-1}x_n^{n-1} & = & y_n
\end{array}$$

In matrix form, this is

$$\left[\begin{array}{ccccc}
1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\
1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\
\vdots & \vdots & \vdots & & \vdots \\
1 & x_n & x_n^2 & \cdots & x_n^{n-1}
\end{array}\right]
\left[\begin{array}{c} r_0 \\ r_1 \\ \vdots \\ r_{n-1} \end{array}\right]
=
\left[\begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_n \end{array}\right] \tag{3.3}$$

It can be shown (see Theorem 3.2.7) that the determinant of the coefficient matrix equals the product of all terms $(x_i - x_j)$ with $i > j$ and so is nonzero (because the x_i are distinct). Hence the equations have a unique solution $r_0, r_1, \dots, r_{n-1}$. This proves

Theorem 3.2.6

Let n data pairs $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ be given, and assume that the x_i are distinct. Then there exists a unique polynomial

$$p(x) = r_0 + r_1x + r_2x^2 + \cdots + r_{n-1}x^{n-1}$$

such that $p(x_i) = y_i$ for each $i = 1, 2, \dots, n$.

The polynomial in Theorem 3.2.6 is called the **interpolating polynomial** for the data.

---

[6]A **polynomial** is an expression of the form $a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ where the a_i are numbers and x is a variable. If $a_n \neq 0$, the integer n is called the degree of the polynomial, and a_n is called the leading coefficient. See Appendix D.

We conclude by evaluating the determinant of the coefficient matrix in Equation 3.3. If $a_1, a_2, \dots, a_n$ are numbers, the determinant

$$\det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix}$$

is called a **Vandermonde determinant**.[7] There is a simple formula for this determinant. If $n = 2$, it equals $(a_2 - a_1)$; if $n = 3$, it is $(a_3 - a_2)(a_3 - a_1)(a_2 - a_1)$ by Example 3.1.8. The general result is the product

$$\prod_{1 \leq j < i \leq n} (a_i - a_j)$$

of all factors $(a_i - a_j)$ where $1 \leq j < i \leq n$. For example, if $n = 4$, it is

$$(a_4 - a_3)(a_4 - a_2)(a_4 - a_1)(a_3 - a_2)(a_3 - a_1)(a_2 - a_1)$$

Theorem 3.2.7

Let $a_1, a_2, \dots, a_n$ be numbers where $n \geq 2$. Then the corresponding Vandermonde determinant is given by

$$\det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{bmatrix} = \prod_{1 \leq j < i \leq n} (a_i - a_j)$$

Proof. We may assume that the a_i are distinct; otherwise both sides are zero. We proceed by induction on $n \geq 2$; we have it for $n = 2, 3$. So assume it holds for $n - 1$. The trick is to replace a_n by a variable x, and consider the determinant

$$p(x) = \det \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^{n-1} \\ 1 & x & x^2 & \cdots & x^{n-1} \end{bmatrix}$$

Then $p(x)$ is a polynomial of degree at most $n - 1$ (expand along the last row), and $p(a_i) = 0$ for each $i = 1, 2, \dots, n - 1$ because in each case there are two identical rows in the determinant. In particular, $p(a_1) = 0$, so we have $p(x) = (x - a_1)p_1(x)$ by the factor theorem (see Appendix D). Since $a_2 \neq a_1$, we obtain $p_1(a_2) = 0$, and so $p_1(x) = (x - a_2)p_2(x)$. Thus $p(x) = (x - a_1)(x - a_2)p_2(x)$. As the a_i are distinct, this process continues to obtain

$$p(x) = (x - a_1)(x - a_2) \cdots (x - a_{n-1})d \tag{3.4}$$

[7] Alexandre Théophile Vandermonde (1735–1796) was a French mathematician who made contributions to the theory of equations.

where d is the coefficient of x^{n-1} in $p(x)$. By the cofactor expansion of $p(x)$ along the last row we get

$$d = (-1)^{n+n} \det \left[\begin{array}{ccccc} 1 & a_1 & a_1^2 & \cdots & a_1^{n-2} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & a_{n-1} & a_{n-1}^2 & \cdots & a_{n-1}^{n-2} \end{array} \right]$$

Because $(-1)^{n+n} = 1$ the induction hypothesis shows that d is the product of all factors $(a_i - a_j)$ where $1 \leq j < i \leq n-1$. The result now follows from Equation 3.4 by substituting a_n for x in $p(x)$. □

Proof of Theorem 3.2.1. If A and B are $n \times n$ matrices we must show that

$$\det(AB) = \det A \det B \tag{3.5}$$

Recall that if E is an elementary matrix obtained by doing one row operation to I_n, then doing that operation to a matrix C (Lemma 2.5.1) results in EC. By looking at the three types of elementary matrices separately, Theorem 3.1.2 shows that

$$\det(EC) = \det E \det C \quad \text{ for any matrix } C \tag{3.6}$$

Thus if $E_1, E_2, \dots, E_k$ are all elementary matrices, it follows by induction that

$$\det(E_k \cdots E_2 E_1 C) = \det E_k \cdots \det E_2 \det E_1 \det C \text{ for any matrix } C \tag{3.7}$$

Lemma. If A has no inverse, then $\det A = 0$.

Proof. Let $A \to R$ where R is reduced row-echelon, say $E_n \cdots E_2 E_1 A = R$. Then R has a row of zeros by Part (4) of Theorem 2.4.5, and hence $\det R = 0$. But then Equation 3.7 gives $\det A = 0$ because $\det E \neq 0$ for any elementary matrix E. This proves the Lemma.

Now we can prove Equation 3.5 by considering two cases.

Case 1. A has no inverse. Then AB also has no inverse (otherwise $A[B(AB)^{-1}] = I$) so A is invertible by Corollary 2.4.2 to Theorem 2.4.5. Hence the above Lemma (twice) gives

$$\det(AB) = 0 = 0 \det B = \det A \det B$$

proving Equation 3.5 in this case.

Case 2. A has an inverse. Then A is a product of elementary matrices by Theorem 2.5.2, say $A = E_1 E_2 \cdots E_k$. Then Equation 3.7 with $C = I$ gives

$$\det A = \det(E_1 E_2 \cdots E_k) = \det E_1 \det E_2 \cdots \det E_k$$

But then Equation 3.7 with $C = B$ gives

$$\det(AB) = \det[(E_1 E_2 \cdots E_k)B] = \det E_1 \det E_2 \cdots \det E_k \det B = \det A \det B$$

and Equation 3.5 holds in this case too. □

Exercises for 3.2

Exercise 3.2.1 Find the adjugate of each of the following matrices.

a. $\left[\begin{array}{rrr} 5 & 1 & 3 \\ -1 & 2 & 3 \\ 1 & 4 & 8 \end{array}\right]$

b. $\left[\begin{array}{rrr} 1 & -1 & 2 \\ 3 & 1 & 0 \\ 0 & -1 & 1 \end{array}\right]$

c. $\left[\begin{array}{rrr} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{array}\right]$

d. $\frac{1}{3}\left[\begin{array}{rrr} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{array}\right]$

Exercise 3.2.2 Use determinants to find which real values of c make each of the following matrices invertible.

a. $\left[\begin{array}{rrr} 1 & 0 & 3 \\ 3 & -4 & c \\ 2 & 5 & 8 \end{array}\right]$

b. $\left[\begin{array}{rrr} 0 & c & -c \\ -1 & 2 & 1 \\ c & -c & c \end{array}\right]$

c. $\left[\begin{array}{rrr} c & 1 & 0 \\ 0 & 2 & c \\ -1 & c & 5 \end{array}\right]$

d. $\left[\begin{array}{rrr} 4 & c & 3 \\ c & 2 & c \\ 5 & c & 4 \end{array}\right]$

e. $\left[\begin{array}{rrr} 1 & 2 & -1 \\ 0 & -1 & c \\ 2 & c & 1 \end{array}\right]$

f. $\left[\begin{array}{rrr} 1 & c & -1 \\ c & 1 & 1 \\ 0 & 1 & c \end{array}\right]$

Exercise 3.2.3 Let A, B, and C denote $n \times n$ matrices and assume that $\det A = -1$, $\det B = 2$, and $\det C = 3$. Evaluate:

a. $\det (A^3BC^TB^{-1})$

b. $\det (B^2C^{-1}AB^{-1}C^T)$

Exercise 3.2.4 Let A and B be invertible $n \times n$ matrices. Evaluate:

a. $\det (B^{-1}AB)$

b. $\det (A^{-1}B^{-1}AB)$

Exercise 3.2.5 If A is 3×3 and $\det (2A^{-1}) = -4$ and $\det (A^3(B^{-1})^T) = -4$, find $\det A$ and $\det B$.

Exercise 3.2.6 Let $A = \left[\begin{array}{ccc} a & b & c \\ p & q & r \\ u & v & w \end{array}\right]$ and assume that $\det A = 3$. Compute:

a. $\det (2B^{-1})$ where $B = \left[\begin{array}{ccc} 4u & 2a & -p \\ 4v & 2b & -q \\ 4w & 2c & -r \end{array}\right]$

b. $\det (2C^{-1})$ where $C = \left[\begin{array}{ccc} 2p & -a+u & 3u \\ 2q & -b+v & 3v \\ 2r & -c+w & 3w \end{array}\right]$

Exercise 3.2.7 If $\det \left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = -2$ calculate:

a. $\det \left[\begin{array}{ccc} 2 & -2 & 0 \\ c+1 & -1 & 2a \\ d-2 & 2 & 2b \end{array}\right]$

b. $\det \left[\begin{array}{ccc} 2b & 0 & 4d \\ 1 & 2 & -2 \\ a+1 & 2 & 2(c-1) \end{array}\right]$

c. $\det (3A^{-1})$ where $A = \left[\begin{array}{cc} 3c & a+c \\ 3d & b+d \end{array}\right]$

Exercise 3.2.8 Solve each of the following by Cramer's rule:

a. $\begin{array}{rcr} 2x + y &=& 1 \\ 3x + 7y &=& -2 \end{array}$

b. $\begin{array}{rcr} 3x + 4y &=& 9 \\ 2x - y &=& -1 \end{array}$

c. $\begin{array}{rcr} 5x + y - z &=& -7 \\ 2x - y - 2z &=& 6 \\ 3x \phantom{{}+y} + 2z &=& -7 \end{array}$

d. $\begin{array}{rcr} 4x - y + 3z &=& 1 \\ 6x + 2y - z &=& 0 \\ 3x + 3y + 2z &=& -1 \end{array}$

Exercise 3.2.9 Use Theorem 3.2.4 to find the $(2, 3)$-entry of A^{-1} if:

a. $A = \left[\begin{array}{rrr} 3 & 2 & 1 \\ 1 & 1 & 2 \\ -1 & 2 & 1 \end{array}\right]$

b. $A = \left[\begin{array}{rrr} 1 & 2 & -1 \\ 3 & 1 & 1 \\ 0 & 4 & 7 \end{array}\right]$

Exercise 3.2.10 Explain what can be said about $\det A$ if:

a. $A^2 = A$

b. $A^2 = I$

c. $A^3 = A$

d. $PA = P$ and P is invertible

e. $A^2 = uA$ and A is $n \times n$

f. $A = -A^T$ and A is $n \times n$

g. $A^2 + I = 0$ and A is $n \times n$

Exercise 3.2.11 Let A be $n \times n$. Show that $uA = (uI)A$, and use this with Theorem 3.2.1 to deduce the result in Theorem 3.1.3: $\det(uA) = u^n \det A$.

Exercise 3.2.12 If A and B are $n \times n$ matrices, if $AB = -BA$, and if n is odd, show that either A or B has no inverse.

Exercise 3.2.13 Show that $\det AB = \det BA$ holds for any two $n \times n$ matrices A and B.

Exercise 3.2.14 If $A^k = 0$ for some $k \geq 1$, show that A is not invertible.

Exercise 3.2.15 If $A^{-1} = A^T$, describe the cofactor matrix of A in terms of A.

Exercise 3.2.16 Show that no 3×3 matrix A exists such that $A^2 + I = 0$. Find a 2×2 matrix A with this property.

Exercise 3.2.17 Show that $\det(A + B^T) = \det(A^T + B)$ for any $n \times n$ matrices A and B.

Exercise 3.2.18 Let A and B be invertible $n \times n$ matrices. Show that $\det A = \det B$ if and only if $A = UB$ where U is a matrix with $\det U = 1$.

Exercise 3.2.19 For each of the matrices in Exercise 2, find the inverse for those values of c for which it exists.

Exercise 3.2.20 In each case either prove the statement or give an example showing that it is false:

a. If $\operatorname{adj} A$ exists, then A is invertible.

b. If A is invertible and $\operatorname{adj} A = A^{-1}$, then $\det A = 1$.

c. $\det(AB) = \det(B^T A)$.

d. If $\det A \neq 0$ and $AB = AC$, then $B = C$.

e. If $A^T = -A$, then $\det A = -1$.

f. If $\operatorname{adj} A = 0$, then $A = 0$.

g. If A is invertible, then $\operatorname{adj} A$ is invertible.

h. If A has a row of zeros, so also does $\operatorname{adj} A$.

i. $\det(A^T A) > 0$ for all square matrices A.

j. $\det(I + A) = 1 + \det A$.

k. If AB is invertible, then A and B are invertible.

l. If $\det A = 1$, then $\operatorname{adj} A = A$.

m. If A is invertible and $\det A = d$, then $\operatorname{adj} A = dA^{-1}$.

Exercise 3.2.21 If A is 2×2 and $\det A = 0$, show that one column of A is a scalar multiple of the other. [*Hint*: Definition 2.5 and Part (2) of Theorem 2.4.5.]

Exercise 3.2.22 Find a polynomial $p(x)$ of degree 2 such that:

a. $p(0) = 2$, $p(1) = 3$, $p(3) = 8$

b. $p(0) = 5$, $p(1) = 3$, $p(2) = 5$

Exercise 3.2.23 Find a polynomial $p(x)$ of degree 3 such that:

a. $p(0) = p(1) = 1$, $p(-1) = 4$, $p(2) = -5$

b. $p(0) = p(1) = 1$, $p(-1) = 2$, $p(-2) = -3$

Exercise 3.2.24 Given the following data pairs, find the interpolating polynomial of degree 3 and estimate the value of y corresponding to $x = 1.5$.

a. (0, 1), (1, 2), (2, 5), (3, 10)

b. (0, 1), (1, 1.49), (2, −0.42), (3, −11.33)

c. (0, 2), (1, 2.03), (2, −0.40), (−1, 0.89)

Exercise 3.2.25 If $A = \begin{bmatrix} 1 & a & b \\ -a & 1 & c \\ -b & -c & 1 \end{bmatrix}$ show that $\det A = 1 + a^2 + b^2 + c^2$. Hence, find A^{-1} for any a, b, and c.

Exercise 3.2.26

a. Show that $A = \begin{bmatrix} a & p & q \\ 0 & b & r \\ 0 & 0 & c \end{bmatrix}$ has an inverse if and only if $abc \neq 0$, and find A^{-1} in that case.

b. Show that if an upper triangular matrix is invertible, the inverse is also upper triangular.

Exercise 3.2.27 Let A be a matrix each of whose entries are integers. Show that each of the following conditions implies the other.

1. A is invertible and A^{-1} has integer entries.

2. $\det A = 1$ or -1.

Exercise 3.2.28 If $A^{-1} = \left[\begin{array}{rrr} 3 & 0 & 1 \\ 0 & 2 & 3 \\ 3 & 1 & -1 \end{array} \right]$ find $\text{adj } A$.

Exercise 3.2.29 If A is 3×3 and $\det A = 2$, find $\det (A^{-1} + 4 \text{ adj } A)$.

Exercise 3.2.30 Show that $\det \left[\begin{array}{cc} 0 & A \\ B & X \end{array} \right] = \det A \det B$ when A and B are 2×2. What if A and B are 3×3?

[*Hint*: Block multiply by $\left[\begin{array}{cc} 0 & I \\ I & 0 \end{array} \right]$.]

Exercise 3.2.31 Let A be $n \times n$, $n \geq 2$, and assume one column of A consists of zeros. Find the possible values of $\text{rank}(\text{adj } A)$.

Exercise 3.2.32 If A is 3×3 and invertible, compute $\det (-A^2(\text{adj } A)^{-1})$.

Exercise 3.2.33 Show that $\text{adj}(uA) = u^{n-1} \text{adj } A$ for all $n \times n$ matrices A.

Exercise 3.2.34 Let A and B denote invertible $n \times n$ matrices. Show that:

a. $\text{adj}(\text{adj } A) = (\det A)^{n-2} A$ (here $n \geq 2$) [*Hint*: See Example 3.2.8.]

b. $\text{adj}(A^{-1}) = (\text{adj } A)^{-1}$

c. $\text{adj}(A^T) = (\text{adj } A)^T$

d. $\text{adj}(AB) = (\text{adj } B)(\text{adj } A)$ [*Hint*: Show that $AB \text{ adj}(AB) = AB \text{ adj } B \text{ adj } A$.]

3.3 Diagonalization and Eigenvalues

The world is filled with examples of systems that evolve in time—the weather in a region, the economy of a nation, the diversity of an ecosystem, etc. Describing such systems is difficult in general and various methods have been developed in special cases. In this section we describe one such method, called *diagonalization,* which is one of the most important techniques in linear algebra. A very fertile example of this procedure is in modelling the growth of the population of an animal species. This has attracted more attention in recent years with the ever increasing awareness that many species are endangered. To motivate the technique, we begin by setting up a simple model of a bird population in which we make assumptions about survival and reproduction rates.

Example 3.3.1

Consider the evolution of the population of a species of birds. Because the number of males and females are nearly equal, we count only females. We assume that each female remains a juvenile for one year and then becomes an adult, and that only adults have offspring. We make three assumptions about reproduction and survival rates:

1. The number of juvenile females hatched in any year is twice the number of adult females alive the year before (we say the **reproduction rate** is 2).

2. Half of the adult females in any year survive to the next year (the **adult survival rate** is $\frac{1}{2}$).

3. One quarter of the juvenile females in any year survive into adulthood (the **juvenile survival rate** is $\frac{1}{4}$).

If there were 100 adult females and 40 juvenile females alive initially, compute the population of females k years later.