11 • *Two-by-Two Matrices Related to the Fibonacci Numbers*

A two-by-two *matrix* is represented by a symbol such as

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

where a, b, c, and d represent any real numbers. These numbers are called *elements*. In this section, all matrices mentioned are to be considered to be two-by-two matrices.

Two matrices,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} e & f \\ g & h \end{pmatrix},$$

are *equal* if and only if their corresponding elements are equal. That is,

$$A = B$$

if and only if

$$a = e, \quad b = f, \quad c = g, \quad \text{and} \quad d = h.$$

The matrix which is the *sum* of two matrices A and B is defined to be

$$A + B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a + e & b + f \\ c + g & d + h \end{pmatrix}.$$

The zero matrix Z is defined as

$$Z = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

It can be shown (Exercise 1, page 68) that

$$A + Z = Z + A = A,$$

and so Z is the *additive identity element* for the system of two-by-two matrices.

Also, it can be shown (Exercise 2, page 68) that, in general,

$$A + B = B + A.$$

Thus, *matrix addition is commutative.* It can also be shown (Exercise 3, page 68) that *matrix addition is associative.*

We define the *negative*, or *additive inverse, of A* to be

$$-A = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$$

and the *difference* of A and B as

$$A - B = A + (-B).$$

You can verify (Exercise 4, page 68) that

$$A + (-A) = Z.$$

When we are dealing with matrices, we refer to the real numbers as *scalars.* The product of a scalar s and a matrix A is defined to be

$$sA = s\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} sa & sb \\ sc & sd \end{pmatrix}.$$

It can be shown (Exercise 5, page 68) that

$$(sr)A = s(rA).$$

The matrix which is the *product* of two matrices A and B is defined to be

$$AB = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}.$$

The identity matrix I is defined as

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

It can be shown (Exercise 7, page 68) that

$$AI = IA = A,$$

and so I is the *multiplicative identity element* for this system. It can also be shown that *matrix multiplication is associative.*

Let $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}$; then:

$$AB = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 3 + 1 \cdot 0 & 1 \cdot 1 + 1 \cdot 1 \\ 1 \cdot 3 + 0 \cdot 0 & 1 \cdot 1 + 0 \cdot 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 3 & 1 \end{pmatrix}$$

$$BA = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 \cdot 1 + 1 \cdot 1 & 3 \cdot 1 + 1 \cdot 0 \\ 0 \cdot 1 + 1 \cdot 1 & 0 \cdot 1 + 1 \cdot 0 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 1 & 0 \end{pmatrix}$$

Therefore, in this case $AB \neq BA$. This demonstrates that *matrix multiplication is not always commutative.*

If $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, then we have

$$AB = \begin{pmatrix} 1 \cdot 1 + (-1)1 & 1 \cdot 1 + (-1)1 \\ (-1)1 + 1 \cdot 1 & (-1)1 + 1 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = Z.$$

Here we note that the product AB is the zero matrix although neither A nor B is the zero matrix. In elementary mathematics, if the product of two numbers is zero, then at least one of the numbers has to be zero. For matrices (which are not numbers) this rule does not apply.

If A, B, and C are matrices, it can be shown (Exercise 8, page 68) that *matrix multiplication is distributive over matrix addition*. That is,

$$A(B + C) = AB + AC \qquad \text{and} \qquad (B + C)A = BA + CA.$$

Since matrix multiplication is not always commutative, these two distributive laws do not necessarily say the same thing.

Associated with each matrix A is a number, called the *determinant* of matrix A, and denoted by det A, which is defined as

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

The matrix A is said to be *nonsingular* if det $A \neq 0$; otherwise, matrix A is *singular*.

We now prove a simple theorem.

THEOREM I

The determinant, det AB, of the product of two matrices, A and B, is the product of the determinants, det A and det B. That is,

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

Proof. Using the matrices A and B shown at the beginning of this section, we have

$$\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc, \qquad \det B = \begin{vmatrix} e & f \\ g & h \end{vmatrix} = eh - fg,$$

$$\begin{aligned} \det AB &= \begin{vmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{vmatrix} \\ &= (ae + bg)(cf + dh) - (af + bh)(ce + dg) \\ &= acef + adeh + bcfg + bdgh - (acef + adfg + bceh + bdgh) \\ &= adeh + bcfg - adfg - bceh \\ &= (ad - bc)(eh - fg) \\ &= (\det A)(\det B). \end{aligned}$$

Before looking at powers of a two-by-two matrix, let us recall how powers were defined for real numbers. If a is a nonzero real number, then one can define all integral powers of a as follows:

$$a^0 = 1, \qquad a^1 = a, \qquad a^{n+1} = a^n a^1,$$
$$\text{and} \quad a^{-n} = \frac{1}{a^n} \quad \text{for positive integral } n$$

The definition $a^{n+1} = a^n a^1$ is called an *inductive definition.* We shall also use an inductive definition for powers of matrices. For any nonzero matrix A,

$$A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I, \qquad A^1 = A, \qquad \text{and} \qquad A^{n+1} = A^n A^1$$

for positive integral n.

We shall now consider a particular matrix

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{with } \det Q = -1,$$

and we shall prove the following theorem.

THEOREM II

For $n \geq 1$, the nth power of Q is given by

$$Q^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}.$$

Proof. We shall use mathematical induction. Clearly, for $n = 1$,

$$Q^1 = \begin{pmatrix} F_2 & F_1 \\ F_1 & F_0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Assume that $Q^k = \begin{pmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{pmatrix}$; then

$$\begin{aligned} Q^{k+1} = Q^k Q^1 &= \begin{pmatrix} F_{k+1} & F_k \\ F_k & F_{k-1} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} F_{k+1} + F_k & F_{k+1} \\ F_k + F_{k-1} & F_k \end{pmatrix} \\ &= \begin{pmatrix} F_{k+2} & F_{k+1} \\ F_{k+1} & F_k \end{pmatrix}. \end{aligned}$$

The proof is complete by mathematical induction.

THEOREM III

$$\det Q^n = (-1)^n, \qquad n \geq 1.$$

Proof. The proof is left to you (Exercise 9, page 68).

From Theorem II and Theorem III we have

$$\det Q^n = \begin{vmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{vmatrix} = F_{n-1}F_{n+1} - F_n^2 = (-1)^n, \quad n \geq 1.$$

Thus, we have still another derivation of (I_{13}) of Section 10.

Of much technical interest in the algebra of matrices is the characteristic polynomial. For a matrix A, the *characteristic polynomial* is defined to be

$$\begin{aligned} P(x) = \det(A - xI) &= \left|\begin{pmatrix} a & b \\ c & d \end{pmatrix} - x\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right| \\ &= \left|\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} -x & 0 \\ 0 & -x \end{pmatrix}\right| \\ &= \begin{vmatrix} a - x & b \\ c & d - x \end{vmatrix} \\ &= x^2 - (a + d)x + (ad - bc). \end{aligned}$$

The equation $P(x) = 0$, or

$$x^2 - (a + d)x + (ad - bc) = 0,$$

is called the *characteristic equation* of matrix A. Note that the constant term is the determinant of matrix A and the linear term has coefficient $-(a + d)$, which is the negative of $(a + d)$. The sum $(a + d)$ of the elements on the diagonal of matrix A is called the *trace* of matrix A. The roots of the characteristic equation are called the *characteristic roots* of matrix A.

For Q, the characteristic polynomial is

$$\begin{aligned} \det(Q - xI) &= \left|\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}\right| \\ &= \begin{vmatrix} 1 - x & 1 \\ 1 & -x \end{vmatrix} = x^2 - x - 1, \end{aligned}$$

and the characteristic equation is

$$x^2 - x - 1 = 0,$$

which we have called the Fibonacci quadratic equation (page 11). The characteristic roots of Q are then, of course,

$$\alpha = \frac{1 + \sqrt{5}}{2} \qquad \text{and} \qquad \beta = \frac{1 - \sqrt{5}}{2}.$$

Now what are the characteristic equation and the characteristic roots of Q^n?

$$\det(Q^n - xI) = \begin{vmatrix} F_{n+1} - x & F_n \\ F_n & F_{n-1} - x \end{vmatrix}$$
$$= x^2 - (F_{n+1} + F_{n-1})x + (F_{n+1}F_{n-1} - F_n^2)$$

But $F_{n+1} + F_{n-1} = L_n$ [from (I_8) of Section 10] and $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$ [(I_{13})]. Thus,

$$\det(Q^n - xI) = x^2 - L_n x + (-1)^n,$$

and the characteristic equation is

$$x^2 - L_n x + (-1)^n = 0.$$

The characteristic roots are given by (using the quadratic formula)

$$x = \frac{L_n \pm \sqrt{L_n^2 - 4(-1)^n}}{2}.$$

We now recall from (I_{12}) in Section 10 that $L_n^2 - 4(-1)^n = 5F_n^2$; thus, the roots are

$$\frac{L_n + \sqrt{5}\,F_n}{2} \quad \text{and} \quad \frac{L_n - \sqrt{5}\,F_n}{2}.$$

However, we recall that these are the expressions that we found for α^n and β^n on page 27, and so we have the following theorem.

THEOREM IV

The characteristic roots of matrix Q^n are the nth powers of the characteristic roots of Q, and the trace of Q^n is L_n.

You can easily verify (Exercise 10, page 68) that

$$Q^2 = Q + I,$$

or

$$Q^2 - Q - I = Z.$$

Thus, Q may be said to satisfy its characteristic equation

$$x^2 - x - 1 = 0.$$

This is an instance of the following theorem, which we state without proof.

THEOREM V (Hamilton-Cayley* Theorem)

Every square matrix satisfies its own characteristic equation.

* Sir W. R. Hamilton (1805–1865), Irish mathematician; Arthur Cayley (1821–1895), English mathematician.

EXERCISES

1. Using

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad Z = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

and the definition of addition, show that

$$A + Z = Z + A = A.$$

2. Using

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

and the definition of addition, show that

$$A + B = B + A.$$

3. Using

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}, \quad \text{and} \quad C = \begin{pmatrix} i & j \\ k & l \end{pmatrix}$$

and the definition of addition, show that

$$(A + B) + C = A + (B + C).$$

4. Show that $A + (-A) = Z$.

5. Using the definition of scalar multiplication, show that $(sr)A = s(rA)$.

6. Show that $-A = (-1)(A)$.

7. Using

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and the definition of multiplication, show that

$$AI = IA = A.$$

8. Using A, B, and C as given in Exercise 3, show that

$$A(B + C) = AB + AC \quad \text{and} \quad (B + C)A = BA + CA.$$

9. Prove that $\det(Q^n) = (-1)^n$ for $Q = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, $n \geq 1$.

10. Show that $Q^2 = Q + I$, or $Q^2 - Q - I = Z$.