Exercise 2.2.16 If a vector $\mathbf{b}$ is a linear combination of the columns of A, show that the system $A\mathbf{x} = \mathbf{b}$ is consistent (that is, it has at least one solution.)

Exercise 2.2.17 If a system $A\mathbf{x} = \mathbf{b}$ is inconsistent (no solution), show that $\mathbf{b}$ is not a linear combination of the columns of A.

Exercise 2.2.18 Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be solutions to the homogeneous system $A\mathbf{x} = \mathbf{0}$.

a. Show that $\mathbf{x}_1 + \mathbf{x}_2$ is a solution to $A\mathbf{x} = \mathbf{0}$.

b. Show that $t\mathbf{x}_1$ is a solution to $A\mathbf{x} = \mathbf{0}$ for any scalar t.

Exercise 2.2.19 Suppose $\mathbf{x}_1$ is a solution to the system $A\mathbf{x} = \mathbf{b}$. If $\mathbf{x}_0$ is any nontrivial solution to the associated homogeneous system $A\mathbf{x} = \mathbf{0}$, show that $\mathbf{x}_1 + t\mathbf{x}_0$, t a scalar, is an infinite one parameter family of solutions to $A\mathbf{x} = \mathbf{b}$. [*Hint*: Example 2.1.7 Section 2.1.]

Exercise 2.2.20 Let A and B be matrices of the same size. If $\mathbf{x}$ is a solution to both the system $A\mathbf{x} = \mathbf{0}$ and the system $B\mathbf{x} = \mathbf{0}$, show that $\mathbf{x}$ is a solution to the system $(A+B)\mathbf{x} = \mathbf{0}$.

Exercise 2.2.21 If A is $m \times n$ and $A\mathbf{x} = \mathbf{0}$ for every $\mathbf{x}$ in $\mathbb{R}^n$, show that $A = 0$ is the zero matrix. [*Hint*: Consider $A\mathbf{e}_j$ where $\mathbf{e}_j$ is the jth column of I_n; that is, $\mathbf{e}_j$ is the vector in $\mathbb{R}^n$ with 1 as entry j and every other entry 0.]

Exercise 2.2.22 Prove part (1) of Theorem 2.2.2.

Exercise 2.2.23 Prove part (2) of Theorem 2.2.2.

2.3 Matrix Multiplication

In Section 2.2 matrix-vector products were introduced. If A is an $m \times n$ matrix, the product $A\mathbf{x}$ was defined for any n-column $\mathbf{x}$ in $\mathbb{R}^n$ as follows: If $A = \left[\begin{array}{cccc} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array} \right]$ where the $\mathbf{a}_j$ are the columns of A, and if $\mathbf{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \right]$, Definition 2.5 reads

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n \tag{2.5}$$

This was motivated as a way of describing systems of linear equations with coefficient matrix A. Indeed every such system has the form $A\mathbf{x} = \mathbf{b}$ where $\mathbf{b}$ is the column of constants.

In this section we extend this matrix-vector multiplication to a way of multiplying matrices in general, and then investigate matrix algebra for its own sake. While it shares several properties of ordinary arithmetic, it will soon become clear that matrix arithmetic is different in a number of ways.

Matrix multiplication is closely related to composition of transformations.

Composition and Matrix Multiplication

Sometimes two transformations "link" together as follows:

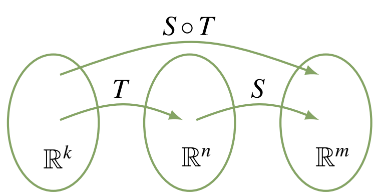
$$\mathbb{R}^k \xrightarrow{T} \mathbb{R}^n \xrightarrow{S} \mathbb{R}^m$$

In this case we can apply T first and then apply S, and the result is a new transformation

$$S \circ T : \mathbb{R}^k \to \mathbb{R}^m$$

called the **composite** of S and T, defined by

$$(S \circ T)(\mathbf{x}) = S[T(\mathbf{x})] \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^k$$



The action of $S \circ T$ can be described as "first T then S" (note the order!)[6]. This new transformation is described in the diagram. The reader will have encountered composition of ordinary functions: For example, consider $\mathbb{R} \xrightarrow{g} \mathbb{R} \xrightarrow{f} \mathbb{R}$ where $f(x) = x^2$ and $g(x) = x + 1$ for all x in $\mathbb{R}$. Then

$$\begin{aligned}(f \circ g)(x) &= f[g(x)] = f(x+1) = (x+1)^2 \\ (g \circ f)(x) &= g[f(x)] = g(x^2) = x^2 + 1\end{aligned}$$

for all x in $\mathbb{R}$.

Our concern here is with matrix transformations. Suppose that A is an $m \times n$ matrix and B is an $n \times k$ matrix, and let $\mathbb{R}^k \xrightarrow{T_B} \mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^m$ be the matrix transformations induced by B and A respectively, that is:

$$T_B(\mathbf{x}) = B\mathbf{x} \text{ for all } \mathbf{x} \text{ in } \mathbb{R}^k \quad \text{and} \quad T_A(\mathbf{y}) = A\mathbf{y} \text{ for all } \mathbf{y} \text{ in } \mathbb{R}^n$$

Write $B = \left[\begin{array}{cccc} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{array}\right]$ where $\mathbf{b}_j$ denotes column j of B for each j. Hence each $\mathbf{b}_j$ is an n-vector (B is $n \times k$) so we can form the matrix-vector product $A\mathbf{b}_j$. In particular, we obtain an $m \times k$ matrix

$$\left[\begin{array}{cccc} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_k \end{array}\right]$$

with columns $A\mathbf{b}_1, A\mathbf{b}_2, \cdots, A\mathbf{b}_k$. Now compute $(T_A \circ T_B)(\mathbf{x})$ for any $\mathbf{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_k \end{array}\right]$ in $\mathbb{R}^k$:

$$\begin{array}{rcll}
(T_A \circ T_B)(\mathbf{x}) & = & T_A[T_B(\mathbf{x})] & \text{Definition of } T_A \circ T_B \\
 & = & A(B\mathbf{x}) & A \text{ and } B \text{ induce } T_A \text{ and } T_B \\
 & = & A(x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_k\mathbf{b}_k) & \text{Equation 2.5 above} \\
 & = & A(x_1\mathbf{b}_1) + A(x_2\mathbf{b}_2) + \cdots + A(x_k\mathbf{b}_k) & \text{Theorem 2.2.2} \\
 & = & x_1(A\mathbf{b}_1) + x_2(A\mathbf{b}_2) + \cdots + x_k(A\mathbf{b}_k) & \text{Theorem 2.2.2} \\
 & = & \left[\begin{array}{cccc} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_k \end{array}\right]\mathbf{x} & \text{Equation 2.5 above}
\end{array}$$

Because $\mathbf{x}$ was an arbitrary vector in $\mathbb{R}^n$, this shows that $T_A \circ T_B$ is the matrix transformation induced by the matrix $\left[\begin{array}{cccc} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_n \end{array}\right]$. This motivates the following definition.

[6]When reading the notation $S \circ T$, we read S first and then T even though the action is "first T then S". This annoying state of affairs results because we write $T(\mathbf{x})$ for the effect of the transformation T on $\mathbf{x}$, with T on the left. If we wrote this instead as $(\mathbf{x})T$, the confusion would not occur. However the notation $T(\mathbf{x})$ is well established.

Definition 2.9 Matrix Multiplication

Let A be an $m \times n$ matrix, let B be an $n \times k$ matrix, and write $B = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{bmatrix}$ where $\mathbf{b}_j$ is column j of B for each j. The product matrix AB is the $m \times k$ matrix defined as follows:

$$AB = A \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{bmatrix} = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_k \end{bmatrix}$$

Thus the product matrix AB is given in terms of its columns $A\mathbf{b}_1, A\mathbf{b}_2, \dots, A\mathbf{b}_n$: Column j of AB is the matrix-vector product $A\mathbf{b}_j$ of A and the corresponding column $\mathbf{b}_j$ of B. Note that each such product $A\mathbf{b}_j$ makes sense by Definition 2.5 because A is $m \times n$ and each $\mathbf{b}_j$ is in $\mathbb{R}^n$ (since B has n rows). Note also that if B is a column matrix, this definition reduces to Definition 2.5 for matrix-vector multiplication.

Given matrices A and B, Definition 2.9 and the above computation give

$$A(B\mathbf{x}) = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_n \end{bmatrix} \mathbf{x} = (AB)\mathbf{x}$$

for all $\mathbf{x}$ in $\mathbb{R}^k$. We record this for reference.

Theorem 2.3.1

Let A be an $m \times n$ matrix and let B be an $n \times k$ matrix. Then the product matrix AB is $m \times k$ and satisfies

$$A(B\mathbf{x}) = (AB)\mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^k$$

Here is an example of how to compute the product AB of two matrices using Definition 2.9.

Example 2.3.1

Compute AB if $A = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 7 \\ 0 & 1 & 8 \end{bmatrix}$ and $B = \begin{bmatrix} 8 & 9 \\ 7 & 2 \\ 6 & 1 \end{bmatrix}$.

Solution. The columns of B are $\mathbf{b}_1 = \begin{bmatrix} 8 \\ 7 \\ 6 \end{bmatrix}$ and $\mathbf{b}_2 = \begin{bmatrix} 9 \\ 2 \\ 1 \end{bmatrix}$, so Definition 2.5 gives

$$A\mathbf{b}_1 = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 7 \\ 0 & 1 & 8 \end{bmatrix} \begin{bmatrix} 8 \\ 7 \\ 6 \end{bmatrix} = \begin{bmatrix} 67 \\ 78 \\ 55 \end{bmatrix} \text{ and } A\mathbf{b}_2 = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 7 \\ 0 & 1 & 8 \end{bmatrix} \begin{bmatrix} 9 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 29 \\ 24 \\ 10 \end{bmatrix}$$

Hence Definition 2.9 above gives $AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 \end{bmatrix} = \begin{bmatrix} 67 & 29 \\ 78 & 24 \\ 55 & 10 \end{bmatrix}$.

Example 2.3.2

If A is $m \times n$ and B is $n \times k$, Theorem 2.3.1 gives a simple formula for the composite of the matrix transformations T_A and T_B:

$$T_A \circ T_B = T_{AB}$$

Solution. Given any $\mathbf{x}$ in $\mathbb{R}^k$,

$$\begin{aligned}(T_A \circ T_B)(\mathbf{x}) &= T_A[T_B(\mathbf{x})] \\ &= A[B\mathbf{x}] \\ &= (AB)\mathbf{x} \\ &= T_{AB}(\mathbf{x})\end{aligned}$$

While Definition 2.9 is important, there is another way to compute the matrix product AB that gives a way to calculate each individual entry. In Section 2.2 we defined the dot product of two n-tuples to be the sum of the products of corresponding entries. We went on to show (Theorem 2.2.5) that if A is an $m \times n$ matrix and $\mathbf{x}$ is an n-vector, then entry j of the product $A\mathbf{x}$ is the dot product of row j of A with $\mathbf{x}$. This observation was called the "dot product rule" for matrix-vector multiplication, and the next theorem shows that it extends to matrix multiplication in general.

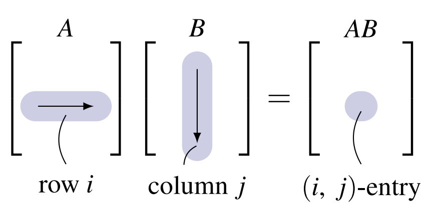
Theorem 2.3.2: Dot Product Rule

Let A and B be matrices of sizes $m \times n$ and $n \times k$, respectively. Then the $(i,\ j)$-entry of AB is the dot product of row i of A with column j of B.

Proof. Write $B = \left[\begin{array}{cccc} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_n \end{array}\right]$ in terms of its columns. Then $A\mathbf{b}_j$ is column j of AB for each j. Hence the $(i,\ j)$-entry of AB is entry i of $A\mathbf{b}_j$, which is the dot product of row i of A with $\mathbf{b}_j$. This proves the theorem. □

Thus to compute the $(i,\ j)$-entry of AB, proceed as follows (see the diagram):

Go *across* row i of A, and *down* column j of B, multiply corresponding entries, and add the results.



Note that this requires that the rows of A must be the same length as the columns of B. The following rule is useful for remembering this and for deciding the size of the product matrix AB.

Compatibility Rule

$$\begin{array}{cc} A & B \\ m \times n & n' \times k \end{array}$$

Let A and B denote matrices. If A is $m \times n$ and B is $n' \times k$, the product AB can be formed if and only if $n = n'$. In this case the size of the product matrix AB is $m \times k$, and we say that AB is **defined**, or that A and B are **compatible** for multiplication.

The diagram provides a useful mnemonic for remembering this. We adopt the following convention:

Convention

Whenever a product of matrices is written, it is tacitly assumed that the sizes of the factors are such that the product is defined.

To illustrate the dot product rule, we recompute the matrix product in Example 2.3.1.

Example 2.3.3

Compute AB if $A = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 7 \\ 0 & 1 & 8 \end{bmatrix}$ and $B = \begin{bmatrix} 8 & 9 \\ 7 & 2 \\ 6 & 1 \end{bmatrix}$.

Solution. Here A is 3×3 and B is 3×2, so the product matrix AB is defined and will be of size 3×2. Theorem 2.3.2 gives each entry of AB as the dot product of the corresponding row of A with the corresponding column of B_j that is,

$$AB = \begin{bmatrix} 2 & 3 & 5 \\ 1 & 4 & 7 \\ 0 & 1 & 8 \end{bmatrix} \begin{bmatrix} 8 & 9 \\ 7 & 2 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} 2\cdot 8+3\cdot 7+5\cdot 6 & 2\cdot 9+3\cdot 2+5\cdot 1 \\ 1\cdot 8+4\cdot 7+7\cdot 6 & 1\cdot 9+4\cdot 2+7\cdot 1 \\ 0\cdot 8+1\cdot 7+8\cdot 6 & 0\cdot 9+1\cdot 2+8\cdot 1 \end{bmatrix} = \begin{bmatrix} 67 & 29 \\ 78 & 24 \\ 55 & 10 \end{bmatrix}$$

Of course, this agrees with Example 2.3.1.

Example 2.3.4

Compute the $(1,\ 3)$- and $(2,\ 4)$-entries of AB where

$$A = \begin{bmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 1 & 6 & 0 \\ 0 & 2 & 3 & 4 \\ -1 & 0 & 5 & 8 \end{bmatrix}.$$

Then compute AB.

Solution. The $(1,\ 3)$-entry of AB is the dot product of row 1 of A and column 3 of B (highlighted in the following display), computed by multiplying corresponding entries and adding the results.

$$\begin{bmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 & 6 & 0 \\ 0 & 2 & 3 & 4 \\ -1 & 0 & 5 & 8 \end{bmatrix} \quad (1,\ 3)\text{-entry} = 3\cdot 6 + (-1)\cdot 3 + 2\cdot 5 = 25$$

Similarly, the $(2,\ 4)$-entry of AB involves row 2 of A and column 4 of B.

$$\left[\begin{array}{rrr} 3 & -1 & 2 \\ 0 & 1 & 4 \end{array}\right]\left[\begin{array}{rrrr} 2 & 1 & 6 & 0 \\ 0 & 2 & 3 & 4 \\ -1 & 0 & 5 & 8 \end{array}\right] \quad (2,\ 4)\text{-entry} = 0\cdot 0 + 1\cdot 4 + 4\cdot 8 = 36$$

Since A is 2×3 and B is 3×4, the product is 2×4.

$$AB = \left[\begin{array}{rrr} 3 & -1 & 2 \\ 0 & 1 & 4 \end{array}\right]\left[\begin{array}{rrrr} 2 & 1 & 6 & 0 \\ 0 & 2 & 3 & 4 \\ -1 & 0 & 5 & 8 \end{array}\right] = \left[\begin{array}{rrrr} 4 & 1 & 25 & 12 \\ -4 & 2 & 23 & 36 \end{array}\right]$$

Example 2.3.5

If $A = \left[\begin{array}{rrr} 1 & 3 & 2 \end{array}\right]$ and $B = \left[\begin{array}{r} 5 \\ 6 \\ 4 \end{array}\right]$, compute A^2, AB, BA, and B^2 when they are defined.[7]

Solution. Here, A is a 1×3 matrix and B is a 3×1 matrix, so A^2 and B^2 are not defined. However, the compatibility rule reads

$$\begin{array}{cc} A & B \\ 1\times 3 & 3\times 1 \end{array} \quad \text{and} \quad \begin{array}{cc} B & A \\ 3\times 1 & 1\times 3 \end{array}$$

so both AB and BA can be formed and these are 1×1 and 3×3 matrices, respectively.

$$AB = \left[\begin{array}{rrr} 1 & 3 & 2 \end{array}\right]\left[\begin{array}{r} 5 \\ 6 \\ 4 \end{array}\right] = \left[\begin{array}{c} 1\cdot 5 + 3\cdot 6 + 2\cdot 4 \end{array}\right] = \left[31\right]$$

$$BA = \left[\begin{array}{r} 5 \\ 6 \\ 4 \end{array}\right]\left[\begin{array}{rrr} 1 & 3 & 2 \end{array}\right] = \left[\begin{array}{ccc} 5\cdot 1 & 5\cdot 3 & 5\cdot 2 \\ 6\cdot 1 & 6\cdot 3 & 6\cdot 2 \\ 4\cdot 1 & 4\cdot 3 & 4\cdot 2 \end{array}\right] = \left[\begin{array}{rrr} 5 & 15 & 10 \\ 6 & 18 & 12 \\ 4 & 12 & 8 \end{array}\right]$$

Unlike numerical multiplication, matrix products AB and BA *need not be equal*. In fact they need not even be the same size, as Example 2.3.5 shows. It turns out to be rare that $AB = BA$ (although it is by no means impossible), and A and B are said to **commute** when this happens.

Example 2.3.6

Let $A = \left[\begin{array}{rr} 6 & 9 \\ -4 & -6 \end{array}\right]$ and $B = \left[\begin{array}{rr} 1 & 2 \\ -1 & 0 \end{array}\right]$. Compute A^2, AB, BA.

[7] As for numbers, we write $A^2 = A\cdot A$, $A^3 = A\cdot A\cdot A$, etc. Note that A^2 is defined if and only if A is of size $n\times n$ for some n.

Solution. $A^2 = \begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix} \begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, so $A^2 = 0$ can occur even if $A \neq 0$. Next,

$$AB = \begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -3 & 12 \\ 2 & -8 \end{bmatrix}$$
$$BA = \begin{bmatrix} 1 & 2 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 6 & 9 \\ -4 & -6 \end{bmatrix} = \begin{bmatrix} -2 & -3 \\ -6 & -9 \end{bmatrix}$$

Hence $AB \neq BA$, even though AB and BA are the same size.

Example 2.3.7

If A is any matrix, then $IA = A$ and $AI = A$, and where I denotes an identity matrix of a size so that the multiplications are defined.

Solution. These both follow from the dot product rule as the reader should verify. For a more formal proof, write $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix}$ where $\mathbf{a}_j$ is column j of A. Then Definition 2.9 and Example 2.2.11 give

$$IA = \begin{bmatrix} I\mathbf{a}_1 & I\mathbf{a}_2 & \cdots & I\mathbf{a}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = A$$

If $\mathbf{e}_j$ denotes column j of I, then $A\mathbf{e}_j = \mathbf{a}_j$ for each j by Example 2.2.12. Hence Definition 2.9 gives:

$$AI = A\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix} = \begin{bmatrix} A\mathbf{e}_1 & A\mathbf{e}_2 & \cdots & A\mathbf{e}_n \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} = A$$

The following theorem collects several results about matrix multiplication that are used everywhere in linear algebra.

Theorem 2.3.3

Assume that a is any scalar, and that A, B, and C are matrices of sizes such that the indicated matrix products are defined. Then:

1. *$IA = A$ and $AI = A$ where I denotes an identity matrix.*
2. *$A(BC) = (AB)C$.*
3. *$A(B+C) = AB + AC$.*
4. *$(B+C)A = BA + CA$.*
5. *$a(AB) = (aA)B = A(aB)$.*
6. *$(AB)^T = B^T A^T$.*

Proof. Condition (1) is Example 2.3.7; we prove (2), (4), and (6) and leave (3) and (5) as exercises.

1. If $C = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_k \end{bmatrix}$ in terms of its columns, then $BC = \begin{bmatrix} B\mathbf{c}_1 & B\mathbf{c}_2 & \cdots & B\mathbf{c}_k \end{bmatrix}$ by Defini-

tion 2.9, so

$$\begin{aligned} A(BC) &= \left[\begin{array}{cccc} A(B\mathbf{c}_1) & A(B\mathbf{c}_2) & \cdots & A(B\mathbf{c}_k) \end{array}\right] && \text{Definition 2.9} \\ &= \left[\begin{array}{cccc} (AB)\mathbf{c}_1 & (AB)\mathbf{c}_2 & \cdots & (AB)\mathbf{c}_k \end{array}\right] && \text{Theorem 2.3.1} \\ &= (AB)C && \text{Definition 2.9} \end{aligned}$$

4. We know (Theorem 2.2.2) that $(B+C)\mathbf{x} = B\mathbf{x} + C\mathbf{x}$ holds for every column $\mathbf{x}$. If we write $A = \left[\begin{array}{cccc} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array}\right]$ in terms of its columns, we get

$$\begin{aligned} (B+C)A &= \left[\begin{array}{cccc} (B+C)\mathbf{a}_1 & (B+C)\mathbf{a}_2 & \cdots & (B+C)\mathbf{a}_n \end{array}\right] && \text{Definition 2.9} \\ &= \left[\begin{array}{cccc} B\mathbf{a}_1 + C\mathbf{a}_1 & B\mathbf{a}_2 + C\mathbf{a}_2 & \cdots & B\mathbf{a}_n + C\mathbf{a}_n \end{array}\right] && \text{Theorem 2.2.2} \\ &= \left[\begin{array}{cccc} B\mathbf{a}_1 & B\mathbf{a}_2 & \cdots & B\mathbf{a}_n \end{array}\right] + \left[\begin{array}{cccc} C\mathbf{a}_1 & C\mathbf{a}_2 & \cdots & C\mathbf{a}_n \end{array}\right] && \text{Adding Columns} \\ &= BA + CA && \text{Definition 2.9} \end{aligned}$$

6. As in Section 2.1, write $A = [a_{ij}]$ and $B = [b_{ij}]$, so that $A^T = [a'_{ij}]$ and $B^T = [b'_{ij}]$ where $a'_{ij} = a_{ji}$ and $b'_{ji} = b_{ij}$ for all i and j. If c_{ij} denotes the $(i,\ j)$-entry of $B^T A^T$, then c_{ij} is the dot product of row i of B^T with column j of A^T. Hence

$$\begin{aligned} c_{ij} = b'_{i1}a'_{1j} + b'_{i2}a'_{2j} + \cdots + b'_{im}a'_{mj} &= b_{1i}a_{j1} + b_{2i}a_{j2} + \cdots + b_{mi}a_{jm} \\ &= a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jm}b_{mi} \end{aligned}$$

But this is the dot product of row j of A with column i of B; that is, the $(j,\ i)$-entry of AB; that is, the $(i,\ j)$-entry of $(AB)^T$. This proves (6). □

Property 2 in Theorem 2.3.3 is called the **associative law** of matrix multiplication. It asserts that the equation $A(BC) = (AB)C$ holds for all matrices (if the products are defined). Hence this product is the same no matter how it is formed, and so is written simply as ABC. This extends: The product $ABCD$ of four matrices can be formed several ways—for example, $(AB)(CD)$, $[A(BC)]D$, and $A[B(CD)]$—but the associative law implies that they are all equal and so are written as $ABCD$. A similar remark applies in general: Matrix products can be written unambiguously with no parentheses.

However, a note of caution about matrix multiplication must be taken: The fact that AB and BA need *not* be equal means that the *order* of the factors is important in a product of matrices. For example $ABCD$ and $ADCB$ may *not* be equal.

Warning

If the order of the factors in a product of matrices is changed, the product matrix may change (or may not be defined). Ignoring this warning is a source of many errors by students of linear algebra!

Properties 3 and 4 in Theorem 2.3.3 are called **distributive laws**. They assert that $A(B+C) = AB + AC$ and $(B+C)A = BA + CA$ hold whenever the sums and products are defined. These rules extend to more than two terms and, together with Property 5, ensure that many manipulations familiar from ordinary algebra extend to matrices. For example

$$\begin{aligned} A(2B-3C+D-5E) &= 2AB-3AC+AD-5AE \\ (A+3C-2D)B &= AB+3CB-2DB \end{aligned}$$

Note again that the warning is in effect: For example $A(B-C)$ need *not* equal $AB-CA$. These rules make possible a lot of simplification of matrix expressions.

Example 2.3.8

Simplify the expression $A(BC-CD)+A(C-B)D-AB(C-D)$.

Solution.

$$\begin{aligned} A(BC-CD)+A(C-B)D-AB(C-D) &= A(BC)-A(CD)+(AC-AB)D-(AB)C+(AB)D \\ &= ABC-ACD+ACD-ABD-ABC+ABD \\ &= 0 \end{aligned}$$

Example 2.3.9 and Example 2.3.10 below show how we can use the properties in Theorem 2.3.2 to deduce other facts about matrix multiplication. Matrices A and B are said to **commute** if $AB = BA$.

Example 2.3.9

Suppose that A, B, and C are $n \times n$ matrices and that both A and B commute with C; that is, $AC = CA$ and $BC = CB$. Show that AB commutes with C.

Solution. Showing that AB commutes with C means verifying that $(AB)C = C(AB)$. The computation uses the associative law several times, as well as the given facts that $AC = CA$ and $BC = CB$.

$$(AB)C = A(BC) = A(CB) = (AC)B = (CA)B = C(AB)$$

Example 2.3.10

Show that $AB = BA$ if and only if $(A-B)(A+B) = A^2 - B^2$.

Solution. The following *always* holds:

$$(A-B)(A+B) = A(A+B)-B(A+B) = A^2+AB-BA-B^2 \tag{2.6}$$

Hence if $AB = BA$, then $(A-B)(A+B) = A^2 - B^2$ follows. Conversely, if this last equation holds, then equation (2.6) becomes

$$A^2 - B^2 = A^2 + AB - BA - B^2$$

This gives $0 = AB - BA$, and $AB = BA$ follows.

In Section 2.2 we saw (in Theorem 2.2.1) that every system of linear equations has the form

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

where A is the coefficient matrix, $\mathbf{x}$ is the column of variables, and $\mathbf{b}$ is the constant matrix. Thus the *system* of linear equations becomes a single matrix equation. Matrix multiplication can yield information about such a system.

Example 2.3.11

Consider a system $A\mathbf{x} = \mathbf{b}$ of linear equations where A is an $m \times n$ matrix. Assume that a matrix C exists such that $CA = I_n$. If the system $A\mathbf{x} = \mathbf{b}$ *has* a solution, show that this solution must be $C\mathbf{b}$. Give a condition guaranteeing that $C\mathbf{b}$ *is in fact* a solution.

Solution. Suppose that $\mathbf{x}$ is any solution to the system, so that $A\mathbf{x} = \mathbf{b}$. Multiply both sides of this matrix equation by C to obtain, successively,

$$C(A\mathbf{x}) = C\mathbf{b}, \quad (CA)\mathbf{x} = C\mathbf{b}, \quad I_n\mathbf{x} = C\mathbf{b}, \quad \mathbf{x} = C\mathbf{b}$$

This shows that *if* the system has a solution $\mathbf{x}$, then that solution must be $\mathbf{x} = C\mathbf{b}$, as required. But it does *not* guarantee that the system *has* a solution. However, if we write $\mathbf{x}_1 = C\mathbf{b}$, then

$$A\mathbf{x}_1 = A(C\mathbf{b}) = (AC)\mathbf{b}$$

Thus $\mathbf{x}_1 = C\mathbf{b}$ will be a solution if the condition $AC = I_m$ is satisfied.

The ideas in Example 2.3.11 lead to important information about matrices; this will be pursued in the next section.

Block Multiplication

Definition 2.10 Block Partition of a Matrix

It is often useful to consider matrices whose entries are themselves matrices (called ***blocks****). A matrix viewed in this way is said to be* ***partitioned into blocks****.*

For example, writing a matrix B in the form

$$B = \left[\begin{array}{cccc} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{array} \right] \text{ where the } \mathbf{b}_j \text{ are the columns of } B$$

is such a block partition of B. Here is another example.

Consider the matrices

$$A = \left[\begin{array}{rr|rrr} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 2 & -1 & 4 & 2 & 1 \\ 3 & 1 & -1 & 7 & 5 \end{array} \right] = \left[\begin{array}{cc} I_2 & 0_{23} \\ P & Q \end{array} \right] \quad \text{and} \quad B = \left[\begin{array}{rr} 4 & -2 \\ 5 & 6 \\ \hline 7 & 3 \\ -1 & 0 \\ 1 & 6 \end{array} \right] = \left[\begin{array}{c} X \\ Y \end{array} \right]$$

where the blocks have been labelled as indicated. This is a natural way to partition A into blocks in view of the blocks I_2 and 0_{23} that occur. This notation is particularly useful when we are multiplying the matrices A and B because the product AB can be computed in block form as follows:

$$AB = \left[\begin{array}{cc} I & 0 \\ P & Q \end{array} \right] \left[\begin{array}{c} X \\ Y \end{array} \right] = \left[\begin{array}{c} IX + 0Y \\ PX + QY \end{array} \right] = \left[\begin{array}{c} X \\ PX + QY \end{array} \right] = \left[\begin{array}{rr} 4 & -2 \\ 5 & 6 \\ \hline 30 & 8 \\ 8 & 27 \end{array} \right]$$

This is easily checked to be the product AB, computed in the conventional manner.

In other words, *we can compute the product AB by ordinary matrix multiplication, using blocks as entries*. The only requirement is that the blocks be **compatible**. That is, *the sizes of the blocks must be such that all (matrix) products of blocks that occur make sense*. This means that the number of columns in each block of A must equal the number of rows in the corresponding block of B.

Theorem 2.3.4: Block Multiplication

If matrices A and B are partitioned compatibly into blocks, the product AB can be computed by matrix multiplication using blocks as entries.

We omit the proof.

We have been using two cases of block multiplication. If $B = \left[\begin{array}{cccc} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{array} \right]$ is a matrix where the $\mathbf{b}_j$ are the columns of B, and if the matrix product AB is defined, then we have

$$AB = A \left[\begin{array}{cccc} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{array} \right] = \left[\begin{array}{cccc} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_k \end{array} \right]$$

This is Definition 2.9 and is a block multiplication where $A = [A]$ has only one block. As another illustration,

$$B\mathbf{x} = \left[\begin{array}{cccc} \mathbf{b}_1 & \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{array} \right] \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_k \end{array} \right] = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_k\mathbf{b}_k$$

where $\mathbf{x}$ is any $k \times 1$ column matrix (this is Definition 2.5).

It is not our intention to pursue block multiplication in detail here. However, we give one more example because it will be used below.

Theorem 2.3.5

Suppose matrices $A = \left[\begin{array}{cc} B & X \\ 0 & C \end{array} \right]$ and $A_1 = \left[\begin{array}{cc} B_1 & X_1 \\ 0 & C_1 \end{array} \right]$ are partitioned as shown where B and B_1 are square matrices of the same size, and C and C_1 are also square of the same size. These are compatible partitionings and block multiplication gives

$$AA_1 = \left[\begin{array}{cc} B & X \\ 0 & C \end{array} \right] \left[\begin{array}{cc} B_1 & X_1 \\ 0 & C_1 \end{array} \right] = \left[\begin{array}{cc} BB_1 & BX_1 + XC_1 \\ 0 & CC_1 \end{array} \right]$$

Example 2.3.12

Obtain a formula for A^k where $A = \begin{bmatrix} I & X \\ 0 & 0 \end{bmatrix}$ is square and I is an identity matrix.

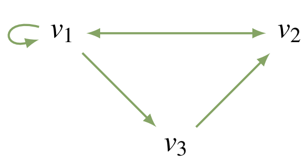
Solution. We have $A^2 = \begin{bmatrix} I & X \\ 0 & 0 \end{bmatrix}\begin{bmatrix} I & X \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} I^2 & IX+X0 \\ 0 & 0^2 \end{bmatrix} = \begin{bmatrix} I & X \\ 0 & 0 \end{bmatrix} = A$. Hence $A^3 = AA^2 = AA = A^2 = A$. Continuing in this way, we see that $A^k = A$ for every $k \geq 1$.

Block multiplication has theoretical uses as we shall see. However, it is also useful in computing products of matrices in a computer with limited memory capacity. The matrices are partitioned into blocks in such a way that each product of blocks can be handled. Then the blocks are stored in auxiliary memory and their products are computed one by one.

Directed Graphs

The study of directed graphs illustrates how matrix multiplication arises in ways other than the study of linear equations or matrix transformations.

A **directed graph** consists of a set of points (called **vertices**) connected by arrows (called **edges**). For example, the vertices could represent cities and the edges available flights. If the graph has n vertices $v_1, v_2, \dots, v_n$, the **adjacency** matrix $A = \left[a_{ij}\right]$ is the $n \times n$ matrix whose (i, j)-entry a_{ij} is 1 if there is an edge from v_j to v_i (note the order), and zero otherwise. For example, the adjacency matrix of the directed graph shown is $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$.



A **path of length** r (or an r-**path**) from vertex j to vertex i is a sequence of r edges leading from v_j to v_i. Thus $v_1 \to v_2 \to v_1 \to v_1 \to v_3$ is a 4-path from v_1 to v_3 in the given graph. The edges are just the paths of length 1, so the (i, j)-entry a_{ij} of the adjacency matrix A is the number of 1-paths from v_j to v_i. This observation has an important extension:

Theorem 2.3.6

If A is the adjacency matrix of a directed graph with n vertices, then the (i, j)-entry of A^r is the number of r-paths $v_j \to v_i$.

As an illustration, consider the adjacency matrix A in the graph shown. Then

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \quad A^2 = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \quad \text{and} \quad A^3 = \begin{bmatrix} 4 & 2 & 1 \\ 3 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

Hence, since the $(2, 1)$-entry of A^2 is 2, there are two 2-paths $v_1 \to v_2$ (in fact they are $v_1 \to v_1 \to v_2$ and $v_1 \to v_3 \to v_2$). Similarly, the $(2, 3)$-entry of A^2 is zero, so there are *no* 2-paths $v_3 \to v_2$, as the reader can verify. The fact that no entry of A^3 is zero shows that it is possible to go from any vertex to any other vertex in exactly three steps.

To see why Theorem 2.3.6 is true, observe that it asserts that

$$\text{the } (i,\ j)\text{-entry of } A^r \text{ equals the number of } r\text{-paths } v_j \to v_i \tag{2.7}$$

holds for each $r \geq 1$. We proceed by induction on r (see Appendix C). The case $r = 1$ is the definition of the adjacency matrix. So assume inductively that (2.7) is true for some $r \geq 1$; we must prove that (2.7) also holds for $r+1$. But every $(r+1)$-path $v_j \to v_i$ is the result of an r-path $v_j \to v_k$ for some k, followed by a 1-path $v_k \to v_i$. Writing $A = \left[a_{ij}\right]$ and $A^r = \left[b_{ij}\right]$, there are b_{kj} paths of the former type (by induction) and a_{ik} of the latter type, and so there are $a_{ik}b_{kj}$ such paths in all. Summing over k, this shows that there are

$$a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \quad (r+1)\text{-paths } v_j \to v_i$$

But this sum is the dot product of the ith row $\left[\begin{array}{cccc} a_{i1} & a_{i2} & \cdots & a_{in} \end{array}\right]$ of A with the jth column $\left[\begin{array}{cccc} b_{1j} & b_{2j} & \cdots & b_{nj} \end{array}\right]^T$ of A^r. As such, it is the $(i,\ j)$-entry of the matrix product $A^rA = A^{r+1}$. This shows that (2.7) holds for $r+1$, as required.

Exercises for 2.3

Exercise 2.3.1 Compute the following matrix products.

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

b. $\left[\begin{array}{rrr} 1 & -1 & 2 \\ 2 & 0 & 4 \end{array}\right]\left[\begin{array}{rrr} 2 & 3 & 1 \\ 1 & 9 & 7 \\ -1 & 0 & 2 \end{array}\right]$

c. $\left[\begin{array}{rrr} 5 & 0 & -7 \\ 1 & 5 & 9 \end{array}\right]\left[\begin{array}{r} 3 \\ 1 \\ -1 \end{array}\right]$

d. $\left[\begin{array}{rrr} 1 & 3 & -3 \end{array}\right]\left[\begin{array}{rr} 3 & 0 \\ -2 & 1 \\ 0 & 6 \end{array}\right]$

e. $\left[\begin{array}{rrr} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]\left[\begin{array}{rr} 3 & -2 \\ 5 & -7 \\ 9 & 7 \end{array}\right]$

f. $\left[\begin{array}{rrr} 1 & -1 & 3 \end{array}\right]\left[\begin{array}{r} 2 \\ 1 \\ -8 \end{array}\right]$

g. $\left[\begin{array}{r} 2 \\ 1 \\ -7 \end{array}\right]\left[\begin{array}{rrr} 1 & -1 & 3 \end{array}\right]$

h. $\left[\begin{array}{rr} 3 & 1 \\ 5 & 2 \end{array}\right]\left[\begin{array}{rr} 2 & -1 \\ -5 & 3 \end{array}\right]$

i. $\left[\begin{array}{rrr} 2 & 3 & 1 \\ 5 & 7 & 4 \end{array}\right]\left[\begin{array}{ccc} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{array}\right]$

j. $\left[\begin{array}{ccc} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{array}\right]\left[\begin{array}{ccc} a' & 0 & 0 \\ 0 & b' & 0 \\ 0 & 0 & c' \end{array}\right]$

Exercise 2.3.2 In each of the following cases, find all possible products A^2, AB, AC, and so on.

a. $A = \left[\begin{array}{rrr} 1 & 2 & 3 \\ -1 & 0 & 0 \end{array}\right]$, $B = \left[\begin{array}{rr} 1 & -2 \\ \frac{1}{2} & 3 \end{array}\right]$,
$C = \left[\begin{array}{rr} -1 & 0 \\ 2 & 5 \\ 0 & 5 \end{array}\right]$

b. $A = \left[\begin{array}{rrr} 1 & 2 & 4 \\ 0 & 1 & -1 \end{array}\right]$, $B = \left[\begin{array}{rr} -1 & 6 \\ 1 & 0 \end{array}\right]$,
$C = \left[\begin{array}{rr} 2 & 0 \\ -1 & 1 \\ 1 & 2 \end{array}\right]$

Exercise 2.3.3 Find a, b, a_1, and b_1 if:

a. $\left[\begin{array}{cc} a & b \\ a_1 & b_1 \end{array}\right]\left[\begin{array}{rr} 3 & -5 \\ -1 & 2 \end{array}\right]=\left[\begin{array}{rr} 1 & -1 \\ 2 & 0 \end{array}\right]$

b. $\left[\begin{array}{rr} 2 & 1 \\ -1 & 2 \end{array}\right]\left[\begin{array}{cc} a & b \\ a_1 & b_1 \end{array}\right]=\left[\begin{array}{rr} 7 & 2 \\ -1 & 4 \end{array}\right]$

Exercise 2.3.4 Verify that $A^2-A-6I=0$ if:

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

Exercise 2.3.5

Given $A=\left[\begin{array}{rr} 1 & -1 \\ 0 & 1 \end{array}\right]$, $B=\left[\begin{array}{rrr} 1 & 0 & -2 \\ 3 & 1 & 0 \end{array}\right]$, $C=\left[\begin{array}{cc} 1 & 0 \\ 2 & 1 \\ 5 & 8 \end{array}\right]$, and $D=\left[\begin{array}{rrr} 3 & -1 & 2 \\ 1 & 0 & 5 \end{array}\right]$, verify the following facts from Theorem 2.3.1.

a. $A(B-D)=AB-AD$ b. $A(BC)=(AB)C$

c. $(CD)^T=D^TC^T$

Exercise 2.3.6 Let A be a 2×2 matrix.

a. If A commutes with $\left[\begin{array}{cc} 0 & 1 \\ 0 & 0 \end{array}\right]$, show that $A=\left[\begin{array}{cc} a & b \\ 0 & a \end{array}\right]$ for some a and b.

b. If A commutes with $\left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array}\right]$, show that $A=\left[\begin{array}{cc} a & 0 \\ c & a \end{array}\right]$ for some a and c.

c. Show that A commutes with *every* 2×2 matrix if and only if $A=\left[\begin{array}{cc} a & 0 \\ 0 & a \end{array}\right]$ for some a.

Exercise 2.3.7

a. If A^2 can be formed, what can be said about the size of A?

b. If AB and BA can both be formed, describe the sizes of A and B.

c. If ABC can be formed, A is 3×3, and C is 5×5, what size is B?

Exercise 2.3.8

a. Find two 2×2 matrices A such that $A^2=0$.

b. Find three 2×2 matrices A such that (i) $A^2=I$; (ii) $A^2=A$.

c. Find 2×2 matrices A and B such that $AB=0$ but $BA\neq 0$.

Exercise 2.3.9 Write $P=\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array}\right]$, and let A be $3\times n$ and B be $m\times 3$.

a. Describe PA in terms of the rows of A.

b. Describe BP in terms of the columns of B.

Exercise 2.3.10 Let A, B, and C be as in Exercise 2.3.5. Find the $(3,\ 1)$-entry of CAB using exactly six numerical multiplications.

Exercise 2.3.11 Compute AB, using the indicated block partitioning.

$$A=\left[\begin{array}{rr|rr} 2 & -1 & 3 & 1 \\ 1 & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right] \quad B=\left[\begin{array}{rr|r} 1 & 2 & 0 \\ -1 & 0 & 0 \\ \hline 0 & 5 & 1 \\ 1 & -1 & 0 \end{array}\right]$$

Exercise 2.3.12 In each case give formulas for all powers A, A^2, A^3, $\dots$ of A using the block decomposition indicated.

a. $A=\left[\begin{array}{r|rr} 1 & 0 & 0 \\ \hline 1 & 1 & -1 \\ 1 & -1 & 1 \end{array}\right]$

b. $A=\left[\begin{array}{rr|rr} 1 & -1 & 2 & -1 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right]$

Exercise 2.3.13 Compute the following using block multiplication (all blocks are $k\times k$).

a. $\left[\begin{array}{rr} I & X \\ -Y & I \end{array}\right]\left[\begin{array}{cc} I & 0 \\ Y & I \end{array}\right]$ b. $\left[\begin{array}{cc} I & X \\ 0 & I \end{array}\right]\left[\begin{array}{rr} I & -X \\ 0 & I \end{array}\right]$

c. $\left[\begin{array}{cc} I & X \end{array}\right]\left[\begin{array}{cc} I & X \end{array}\right]^T$ d. $\left[\begin{array}{cc} I & X^T \end{array}\right]\left[\begin{array}{cc} -X & I \end{array}\right]^T$

e. $\left[\begin{array}{rr} I & X \\ 0 & -I \end{array}\right]^n$ any $n\geq 1$

f. $\left[\begin{array}{cc} 0 & X \\ I & 0 \end{array}\right]^n$ any $n\geq 1$

Exercise 2.3.14 Let A denote an $m \times n$ matrix.

a. If $AX = 0$ for every $n \times 1$ matrix X, show that $A = 0$.

b. If $YA = 0$ for every $1 \times m$ matrix Y, show that $A = 0$.

Exercise 2.3.15

a. If $U = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$, and $AU = 0$, show that $A = 0$.

b. Let U be such that $AU = 0$ implies that $A = 0$. If $PU = QU$, show that $P = Q$.

Exercise 2.3.16 Simplify the following expressions where A, B, and C represent matrices.

a. $A(3B - C) + (A - 2B)C + 2B(C + 2A)$

b. $A(B + C - D) + B(C - A + D) - (A + B)C + (A - B)D$

c. $AB(BC - CB) + (CA - AB)BC + CA(A - B)C$

d. $(A - B)(C - A) + (C - B)(A - C) + (C - A)^2$

Exercise 2.3.17 If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ where $a \neq 0$, show that A factors in the form $A = \begin{bmatrix} 1 & 0 \\ x & 1 \end{bmatrix} \begin{bmatrix} y & z \\ 0 & w \end{bmatrix}$.

Exercise 2.3.18 If A and B commute with C, show that the same is true of:

a. $A + B$

b. kA, k any scalar

Exercise 2.3.19 If A is any matrix, show that both AA^T and A^TA are symmetric.

Exercise 2.3.20 If A and B are symmetric, show that AB is symmetric if and only if $AB = BA$.

Exercise 2.3.21 If A is a 2×2 matrix, show that $A^TA = AA^T$ if and only if A is symmetric or $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ for some a and b.

Exercise 2.3.22

a. Find all symmetric 2×2 matrices A such that $A^2 = 0$.

b. Repeat (a) if A is 3×3.

c. Repeat (a) if A is $n \times n$.

Exercise 2.3.23 Show that there exist no 2×2 matrices A and B such that $AB - BA = I$. [*Hint*: Examine the (1, 1)- and (2, 2)-entries.]

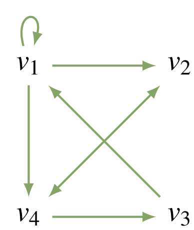
Exercise 2.3.24 Let B be an $n \times n$ matrix. Suppose $AB = 0$ for some nonzero $m \times n$ matrix A. Show that no $n \times n$ matrix C exists such that $BC = I$.

Exercise 2.3.25 An autoparts manufacturer makes fenders, doors, and hoods. Each requires assembly and packaging carried out at factories: Plant 1, Plant 2, and Plant 3. Matrix A below gives the number of hours for assembly and packaging, and matrix B gives the hourly rates at the three plants. Explain the meaning of the (3, 2)-entry in the matrix AB. Which plant is the most economical to operate? Give reasons.

$$\begin{array}{l} \\ \text{Fenders} \\ \text{Doors} \\ \text{Hoods} \end{array} \begin{array}{c} \begin{array}{cc} \text{Assembly} & \text{Packaging} \end{array} \\ \begin{bmatrix} 12 & 2 \\ 21 & 3 \\ 10 & 2 \end{bmatrix} \end{array} = A$$

$$\begin{array}{l} \\ \text{Assembly} \\ \text{Packaging} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Plant 1} & \text{Plant 2} & \text{Plant 3} \end{array} \\ \begin{bmatrix} 21 & 18 & 20 \\ 14 & 10 & 13 \end{bmatrix} \end{array} = B$$

Exercise 2.3.26 For the directed graph below, find the adjacency matrix A, compute A^3, and determine the number of paths of length 3 from v_1 to v_4 and from v_2 to v_3.



Exercise 2.3.27 In each case either show the statement is true, or give an example showing that it is false.

a. If $A^2 = I$, then $A = I$.

b. If $AJ = A$, then $J = I$.

c. If A is square, then $(A^T)^3 = (A^3)^T$.

d. If A is symmetric, then $I + A$ is symmetric.

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

f. If $A \neq 0$, then $A^2 \neq 0$.

g. If A has a row of zeros, so also does BA for all B.

h. If A commutes with $A+B$, then A commutes with B.

i. If B has a column of zeros, so also does AB.

j. If AB has a column of zeros, so also does B.

k. If A has a row of zeros, so also does AB.

l. If AB has a row of zeros, so also does A.

Exercise 2.3.28

a. If A and B are 2×2 matrices whose rows sum to 1, show that the rows of AB also sum to 1.

b. Repeat part (a) for the case where A and B are $n \times n$.

Exercise 2.3.29 Let A and B be $n \times n$ matrices for which the systems of equations $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ each have only the trivial solution $\mathbf{x} = \mathbf{0}$. Show that the system $(AB)\mathbf{x} = \mathbf{0}$ has only the trivial solution.

Exercise 2.3.30 The **trace** of a square matrix A, denoted $\operatorname{tr} A$, is the sum of the elements on the main diagonal of A. Show that, if A and B are $n \times n$ matrices:

a. $\operatorname{tr}(A+B) = \operatorname{tr} A + \operatorname{tr} B$.

b. $\operatorname{tr}(kA) = k \operatorname{tr}(A)$ for any number k.

c. $\operatorname{tr}(A^T) = \operatorname{tr}(A)$.

d. $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

e. $\operatorname{tr}(AA^T)$ is the sum of the squares of all entries of A.

Exercise 2.3.31 Show that $AB - BA = I$ is impossible.

[*Hint*: See the preceding exercise.]

Exercise 2.3.32 A square matrix P is called an **idempotent** if $P^2 = P$. Show that:

a. 0 and I are idempotents.

b. $\left[\begin{array}{cc} 1 & 1 \\ 0 & 0 \end{array}\right]$, $\left[\begin{array}{cc} 1 & 0 \\ 1 & 0 \end{array}\right]$, and $\frac{1}{2}\left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array}\right]$, are idempotents.

c. If P is an idempotent, so is $I - P$. Show further that $P(I-P) = 0$.

d. If P is an idempotent, so is P^T.

e. If P is an idempotent, so is $Q = P + AP - PAP$ for any square matrix A (of the same size as P).

f. If A is $n \times m$ and B is $m \times n$, and if $AB = I_n$, then BA is an idempotent.

Exercise 2.3.33 Let A and B be $n \times n$ **diagonal matrices** (all entries off the main diagonal are zero).

a. Show that AB is diagonal and $AB = BA$.

b. Formulate a rule for calculating XA if X is $m \times n$.

c. Formulate a rule for calculating AY if Y is $n \times k$.

Exercise 2.3.34 If A and B are $n \times n$ matrices, show that:

a. $AB = BA$ if and only if

$$(A+B)^2 = A^2 + 2AB + B^2$$

b. $AB = BA$ if and only if

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

Exercise 2.3.35 In Theorem 2.3.3, prove

a. part 3;

b. part 5.

2.4 Matrix Inverses

Three basic operations on matrices, addition, multiplication, and subtraction, are analogs for matrices of the same operations for numbers. In this section we introduce the matrix analog of numerical division.

To begin, consider how a numerical equation $ax = b$ is solved when a and b are known numbers. If $a = 0$, there is no solution (unless $b = 0$). But if $a \neq 0$, we can multiply both sides by the inverse $a^{-1} = \frac{1}{a}$ to obtain the solution $x = a^{-1}b$. Of course multiplying by a^{-1} is just dividing by a, and the property of a^{-1} that makes this work is that $a^{-1}a = 1$. Moreover, we saw in Section 2.2 that the role that 1 plays in arithmetic is played in matrix algebra by the identity matrix I. This suggests the following definition.

Definition 2.11 Matrix Inverses

If A is a square matrix, a matrix B is called an ***inverse*** *of A if and only if*

$$AB = I \quad \text{and} \quad BA = I$$

A matrix A that has an inverse is called an ***invertible matrix***.[8]

Example 2.4.1

Show that $B = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}$ is an inverse of $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$.

Solution. Compute AB and BA.

$$AB = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad BA = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Hence $AB = I = BA$, so B is indeed an inverse of A.

Example 2.4.2

Show that $A = \begin{bmatrix} 0 & 0 \\ 1 & 3 \end{bmatrix}$ has no inverse.

Solution. Let $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ denote an arbitrary 2×2 matrix. Then

$$AB = \begin{bmatrix} 0 & 0 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ a+3c & b+3d \end{bmatrix}$$

so AB has a row of zeros. Hence AB cannot equal I for any B.

[8] Only square matrices have inverses. Even though it is plausible that nonsquare matrices A and B could exist such that $AB = I_m$ and $BA = I_n$, where A is $m \times n$ and B is $n \times m$, we claim that this forces $n = m$. Indeed, if $m < n$ there exists a nonzero column $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$ (by Theorem 1.3.1), so $\mathbf{x} = I_n\mathbf{x} = (BA)\mathbf{x} = B(A\mathbf{x}) = B(\mathbf{0}) = \mathbf{0}$, a contradiction. Hence $m \geq n$. Similarly, the condition $AB = I_m$ implies that $n \geq m$. Hence $m = n$ so A is square.