3

Elementary Matrix Operations and Systems of Linear Equations

This chapter is devoted to two related objectives:

1. the study of certain "rank-preserving" operations on matrices;
2. the application of these operations and the theory of linear transformations to the solution of systems of linear equations.

As a consequence of objective 1, we obtain a simple method for computing the rank of a linear transformation between finite-dimensional vector spaces by applying these rank-preserving matrix operations to a matrix that represents that transformation.

Solving a system of linear equations is probably the most important application of linear algebra. The familiar method of elimination for solving systems of linear equations, which was discussed in Section 1.4, involves the elimination of variables so that a simpler system can be obtained. The technique by which the variables are eliminated utilizes three types of operations:

1. interchanging any two equations in the system;
2. multiplying any equation in the system by a nonzero constant;
3. adding a multiple of one equation to another.

In Section 3.3, we express a system of linear equations as a single matrix equation. In this representation of the system, the three operations above are the "elementary row operations" for matrices. These operations provide a convenient computational method for determining all solutions to a system of linear equations.

3.1 ELEMENTARY MATRIX OPERATIONS AND ELEMENTARY MATRICES

In this section, we define the elementary operations that are used throughout the chapter. In subsequent sections, we use these operations to obtain simple computational methods for determining the rank of a linear transformation and the solution of a system of linear equations. There are two types of elementary matrix operations—row operations and column operations. As we will see, the row operations are more useful. They arise from the three operations that can be used to eliminate variables in a system of linear equations.

Definitions. *Let A be an $m \times n$ matrix. Any one of the following three operations on the rows [columns] of A is called an* ***elementary row [column] operation****:*

(1) *interchanging any two rows [columns] of A;*
(2) *multiplying any row [column] of A by a nonzero scalar;*
(3) *adding any scalar multiple of a row [column] of A to another row [column].*

Any of these three operations is called an ***elementary operation****. Elementary operations are of* ***type 1, type 2,*** *or* ***type 3*** *depending on whether they are obtained by (1), (2), or (3).*

Example 1

Let

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

Interchanging the second row of A with the first row is an example of an elementary row operation of type 1. The resulting matrix is

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

Multiplying the second column of A by 3 is an example of an elementary column operation of type 2. The resulting matrix is

$$C = \begin{pmatrix} 1 & 6 & 3 & 4 \\ 2 & 3 & -1 & 3 \\ 4 & 0 & 1 & 2 \end{pmatrix}.$$

Adding 4 times the third row of A to the first row is an example of an elementary row operation of type 3. In this case, the resulting matrix is

$$M = \begin{pmatrix} 17 & 2 & 7 & 12 \\ 2 & 1 & -1 & 3 \\ 4 & 0 & 1 & 2 \end{pmatrix}. \quad \blacklozenge$$

Notice that if a matrix Q can be obtained from a matrix P by means of an elementary row operation, then P can be obtained from Q by an elementary row operation of the same type. (See Exercise 8.) So, in Example 1, A can be obtained from M by adding -4 times the third row of M to the first row of M.

Definition. *An* $n \times n$ ***elementary matrix*** *is a matrix obtained by performing an elementary operation on* I_n*. The elementary matrix is said to be of* ***type 1, 2,*** *or* ***3*** *according to whether the elementary operation performed on* I_n *is a type 1, 2, or 3 operation, respectively.*

For example, interchanging the first two rows of I_3 produces the elementary matrix

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

Note that E can also be obtained by interchanging the first two columns of I_3. In fact, *any elementary matrix can be obtained in at least two ways*— either by performing an elementary row operation on I_n or by performing an elementary column operation on I_n. (See Exercise 4.) Similarly,

$$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is an elementary matrix since it can be obtained from I_3 by an elementary column operation of type 3 (adding -2 times the first column of I_3 to the third column) or by an elementary row operation of type 3 (adding -2 times the third row to the first row).

Our first theorem shows that performing an elementary row operation on a matrix is equivalent to multiplying the matrix by an elementary matrix.

Theorem 3.1. *Let* $A \in \mathsf{M}_{m\times n}(F)$*, and suppose that* B *is obtained from* A *by performing an elementary row [column] operation. Then there exists an* $m \times m$ $[n \times n]$ *elementary matrix* E *such that* $B = EA$ $[B = AE]$*. In fact,* E *is obtained from* I_m $[I_n]$ *by performing the same elementary row [column] operation as that which was performed on* A *to obtain* B*. Conversely, if* E *is* *an elementary $m \times m$ $[n \times n]$ matrix, then EA $[AE]$ is the matrix obtained from A by performing the same elementary row [column] operation as that which produces E from I_m $[I_n]$.*

The proof, which we omit, requires verifying Theorem 3.1 for each type of elementary row operation. The proof for column operations can then be obtained by using the matrix transpose to transform a column operation into a row operation. The details are left as an exercise. (See Exercise 7.)

The next example illustrates the use of the theorem.

Example 2

Consider the matrices A and B in Example 1. In this case, B is obtained from A by interchanging the first two rows of A. Performing this same operation on I_3, we obtain the elementary matrix

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

Note that $EA = B$.

In the second part of Example 1, C is obtained from A by multiplying the second column of A by 3. Performing this same operation on I_4, we obtain the elementary matrix

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

Observe that $AE = C$. ♦

It is a useful fact that the inverse of an elementary matrix is also an elementary matrix.

Theorem 3.2. *Elementary matrices are invertible, and the inverse of an elementary matrix is an elementary matrix of the same type.*

Proof. Let E be an elementary $n \times n$ matrix. Then E can be obtained by an elementary row operation on I_n. By reversing the steps used to transform I_n into E, we can transform E back into I_n. The result is that I_n can be obtained from E by an elementary row operation of the same type. By Theorem 3.1, there is an elementary matrix $\overline{E}$ such that $\overline{E}E = I_n$. Therefore, by Exercise 10 of Section 2.4, E is invertible and $E^{-1} = \overline{E}$. ∎

EXERCISES

1. Label the following statements as true or false.

(a) An elementary matrix is always square.
(b) The only entries of an elementary matrix are zeros and ones.
(c) The $n \times n$ identity matrix is an elementary matrix.
(d) The product of two $n \times n$ elementary matrices is an elementary matrix.
(e) The inverse of an elementary matrix is an elementary matrix.
(f) The sum of two $n \times n$ elementary matrices is an elementary matrix.
(g) The transpose of an elementary matrix is an elementary matrix.
(h) If B is a matrix that can be obtained by performing an elementary row operation on a matrix A, then B can also be obtained by performing an elementary column operation on A.
(i) If B is a matrix that can be obtained by performing an elementary row operation on a matrix A, then A can be obtained by performing an elementary row operation on B.

2. Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 0 & 3 \\ 1 & -2 & 1 \\ 1 & -3 & 1 \end{pmatrix}, \text{ and } C \neq \begin{pmatrix} 1 & 0 & 3 \\ 0 & -2 & -2 \\ 1 & -3 & 1 \end{pmatrix}.$$

Find an elementary operation that transforms A into B and an elementary operation that transforms B into C. By means of several additional operations, transform C into I_3.

3. Use the proof of Theorem 3.2 to obtain the inverse of each of the following elementary matrices.

(a) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$

4. Prove the assertion made on page 149: Any elementary $n \times n$ matrix can be obtained in at least two ways—either by performing an elementary row operation on I_n or by performing an elementary column operation on I_n.

5. Prove that E is an elementary matrix if and only if E^t is.

6. Let A be an $m \times n$ matrix. Prove that if B can be obtained from A by an elementary row [column] operation, then B^t can be obtained from A^t by the corresponding elementary column [row] operation.

7. Prove Theorem 3.1.

8. Prove that if a matrix Q can be obtained from a matrix P by an elementary row operation, then P can be obtained from Q by an elementary row operation of the same type. *Hint:* Treat each type of elementary row operation separately.

9. Prove that any elementary row [column] operation of type 1 can be obtained by a succession of three elementary row [column] operations of type 3 followed by one elementary row [column] operation of type 2.

10. Prove that any elementary row [column] operation of type 2 can be obtained by *dividing* some row [column] by a nonzero scalar.

11. Prove that any elementary row [column] operation of type 3 can be obtained by *subtracting* a multiple of some row [column] from another row [column].

12. Let A be an $m \times n$ matrix. Prove that there exists a sequence of elementary row operations of types 1 and 3 that transforms A into an upper triangular matrix.

3.2 THE RANK OF A MATRIX AND MATRIX INVERSES

In this section, we define the *rank* of a matrix. We then use elementary operations to compute the rank of a matrix and a linear transformation. The section concludes with a procedure for computing the inverse of an invertible matrix.

Definition. *If $A \in \mathsf{M}_{m\times n}(F)$, we define the* **rank** *of A, denoted* $\operatorname{rank}(A)$, *to be the rank of the linear transformation $\mathsf{L}_A \colon \mathsf{F}^n \to \mathsf{F}^m$.*

Many results about the rank of a matrix follow immediately from the corresponding facts about a linear transformation. An important result of this type, which follows from Fact 3 (p. 100) and Corollary 2 to Theorem 2.18 (p. 102), is that *an $n \times n$ matrix is invertible if and only if its rank is n.*

Every matrix A is the matrix representation of the linear transformation L_A with respect to the appropriate standard ordered bases. Thus the rank of the linear transformation L_A is the same as the rank of one of its matrix representations, namely, A. The next theorem extends this fact to any matrix representation of any linear transformation defined on finite-dimensional vector spaces.

Theorem 3.3. *Let $\mathsf{T} \colon \mathsf{V} \to \mathsf{W}$ be a linear transformation between finite-dimensional vector spaces, and let β and γ be ordered bases for V and W, respectively. Then* $\operatorname{rank}(\mathsf{T}) = \operatorname{rank}([\mathsf{T}]_\beta^\gamma)$.

Proof. This is a restatement of Exercise 20 of Section 2.4. ∎

Now that the problem of finding the rank of a linear transformation has been reduced to the problem of finding the rank of a matrix, we need a result that allows us to perform rank-preserving operations on matrices. The next theorem and its corollary tell us how to do this.

Theorem 3.4. *Let A be an $m \times n$ matrix. If P and Q are invertible $m \times m$ and $n \times n$ matrices, respectively, then*
(a) $\text{rank}(AQ) = \text{rank}(A)$,
(b) $\text{rank}(PA) = \text{rank}(A)$,
and therefore,
(c) $\text{rank}(PAQ) = \text{rank}(A)$.

Proof. First observe that

$$\mathsf{R}(\mathsf{L}_{AQ}) = \mathsf{R}(\mathsf{L}_A\mathsf{L}_Q) = \mathsf{L}_A\mathsf{L}_Q(\mathsf{F}^n) = \mathsf{L}_A(\mathsf{L}_Q(\mathsf{F}^n)) = \mathsf{L}_A(\mathsf{F}^n) = \mathsf{R}(\mathsf{L}_A)$$

since L_Q is onto. Therefore

$$\text{rank}(AQ) = \dim(\mathsf{R}(\mathsf{L}_{AQ})) = \dim(\mathsf{R}(\mathsf{L}_A)) = \text{rank}(A).$$

This establishes (a). To establish (b), apply Exercise 17 of Section 2.4 to $\mathsf{T} = \mathsf{L}_P$. We omit the details. Finally, applying (a) and (b), we have

$$\text{rank}(PAQ) = \text{rank}(PA) = \text{rank}(A). \qquad \blacksquare$$

Corollary. *Elementary row and column operations on a matrix are rank-preserving.*

Proof. If B is obtained from a matrix A by an elementary row operation, then there exists an elementary matrix E such that $B = EA$. By Theorem 3.2 (p. 150), E is invertible, and hence $\text{rank}(B) = \text{rank}(A)$ by Theorem 3.4. The proof that elementary column operations are rank-preserving is left as an exercise. $\blacksquare$

Now that we have a class of matrix operations that preserve rank, we need a way of examining a transformed matrix to ascertain its rank. The next theorem is the first of several in this direction.

Theorem 3.5. *The rank of any matrix equals the maximum number of its linearly independent columns; that is, the rank of a matrix is the dimension of the subspace generated by its columns.*

Proof. For any $A \in \mathsf{M}_{m\times n}(F)$,

$$\text{rank}(A) = \text{rank}(\mathsf{L}_A) = \dim(\mathsf{R}(\mathsf{L}_A)).$$

Let β be the standard ordered basis for F^n. Then β spans F^n and hence, by Theorem 2.2 (p. 68),

$$\mathsf{R}(\mathsf{L}_A) = \text{span}(\mathsf{L}_A(\beta)) = \text{span}\left(\{\mathsf{L}_A(e_1), \mathsf{L}_A(e_2), \ldots, \mathsf{L}_A(e_n)\}\right).$$

But, for any j, we have seen in Theorem 2.13(b) (p. 90) that $\mathsf{L}_A(e_j) = Ae_j = a_j$, where a_j is the jth column of A. Hence

$$\mathsf{R}(\mathsf{L}_A) = \text{span}\left(\{a_1, a_2, \ldots, a_n\}\right).$$

Thus

$$\text{rank}(A) = \dim(\mathsf{R}(\mathsf{L}_A)) = \dim(\text{span}\left(\{a_1, a_2, \ldots, a_n\}\right)). \qquad \blacksquare$$

Example 1

Let

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

Observe that the first and second columns of A are linearly independent and that the third column is a linear combination of the first two. Thus

$$\text{rank}(A) = \dim\left(\text{span}\left(\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}\right)\right) = 2. \quad \blacklozenge$$

To compute the rank of a matrix A, it is frequently useful to postpone the use of Theorem 3.5 until A has been suitably modified by means of appropriate elementary row and column operations so that the number of linearly independent columns is obvious. The corollary to Theorem 3.4 guarantees that the rank of the modified matrix is the same as the rank of A. One such modification of A can be obtained by using elementary row and column operations to introduce zero entries. The next example illustrates this procedure.

Example 2

Let

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

If we subtract the first row of A from rows 2 and 3 (type 3 elementary row operations), the result is

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & -2 & 2 \\ 0 & -1 & 1 \end{pmatrix}.$$

If we now subtract twice the first column from the second and subtract the first column from the third (type 3 elementary column operations), we obtain

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 2 \\ 0 & -1 & 1 \end{pmatrix}.$$

It is now obvious that the maximum number of linearly independent columns of this matrix is 2. Hence the rank of A is 2. ♦

The next theorem uses this process to transform a matrix into a particularly simple form. The power of this theorem can be seen in its corollaries.

Theorem 3.6. *Let A be an $m \times n$ matrix of rank r. Then $r \leq m$, $r \leq n$, and, by means of a finite number of elementary row and column operations, A can be transformed into the matrix*

$$D = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix},$$

where O_1, O_2, and O_3 are zero matrices. Thus $D_{ii} = 1$ for $i \leq r$ and $D_{ij} = 0$ otherwise.

Theorem 3.6 and its corollaries are quite important. Its proof, though easy to understand, is tedious to read. As an aid in following the proof, we first consider an example.

Example 3

Consider the matrix

$$A = \begin{pmatrix} 0 & 2 & 4 & 2 & 2 \\ 4 & 4 & 4 & 8 & 0 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{pmatrix}.$$

By means of a succession of elementary row and column operations, we can transform A into a matrix D as in Theorem 3.6. We list many of the intermediate matrices, but on several occasions a matrix is transformed from the preceding one by means of several elementary operations. The number above each arrow indicates how many elementary operations are involved. Try to identify the nature of each elementary operation (row or column and type) in the following matrix transformations.

$$\begin{pmatrix} 0 & 2 & 4 & 2 & 2 \\ 4 & 4 & 4 & 8 & 0 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 4 & 4 & 4 & 8 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{pmatrix} \xrightarrow{2}$$

$$\begin{pmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 0 & -6 & -8 & -6 & 2 \\ 0 & -3 & -4 & -3 & 1 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 0 & -6 & -8 & -6 & 2 \\ 0 & -3 & -4 & -3 & 1 \end{pmatrix} \xrightarrow{1}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 \\ 0 & -6 & -8 & -6 & 2 \\ 0 & -3 & -4 & -3 & 1 \end{pmatrix} \xrightarrow{2} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 \\ 0 & 0 & 4 & 0 & 8 \\ 0 & 0 & 2 & 0 & 4 \end{pmatrix} \xrightarrow{3} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 8 \\ 0 & 0 & 2 & 0 & 4 \end{pmatrix} \xrightarrow{1}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 & 4 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{1} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = D$$

By the corollary to Theorem 3.4, $\text{rank}(A) = \text{rank}(D)$. Clearly, however, $\text{rank}(D) = 3$; so $\text{rank}(A) = 3$. ♦

Note that the first two elementary operations in Example 3 result in a 1 in the 1,1 position, and the next several operations (type 3) result in 0's everywhere in the first row and first column except for the 1,1 position. Subsequent elementary operations do not change the first row and first column. With this example in mind, we proceed with the proof of Theorem 3.6.

Proof of Theorem 3.6. If A is the zero matrix, $r = 0$ by Exercise 3. In this case, the conclusion follows with $D = A$.

Now suppose that $A \neq O$ and $r = \text{rank}(A)$; then $r > 0$. The proof is by mathematical induction on m, the number of rows of A.

Suppose that $m = 1$. By means of at most one type 1 column operation and at most one type 2 column operation, A can be transformed into a matrix with a 1 in the 1,1 position. By means of at most $n - 1$ type 3 column operations, this matrix can in turn be transformed into the matrix

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

Note that there is one linearly independent column in D. So $\text{rank}(D) = \text{rank}(A) = 1$ by the corollary to Theorem 3.4 and by Theorem 3.5. Thus the theorem is established for $m = 1$.

Next assume that the theorem holds for any matrix with at most $m - 1$ rows (for some $m > 1$). We must prove that the theorem holds for any matrix with m rows.

Suppose that A is any $m \times n$ matrix. If $n = 1$, Theorem 3.6 can be established in a manner analogous to that for $m = 1$ (see Exercise 10).

We now suppose that $n > 1$. Since $A \neq O$, $A_{ij} \neq 0$ for some i, j. By means of at most one elementary row and at most one elementary column operation (each of type 1), we can move the nonzero entry to the 1,1 position (just as was done in Example 3). By means of at most one additional type 2 operation, we can assure a 1 in the 1,1 position. (Look at the second operation in Example 3.) By means of at most $m-1$ type 3 row operations and at most $n-1$ type 3 column operations, we can eliminate all nonzero entries in the first row and the first column with the exception of the 1 in the 1,1 position. (In Example 3, we used two row and three column operations to do this.)

Thus, with a finite number of elementary operations, A can be transformed into a matrix

$$B = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{array}\right),$$

where B' is an $(m-1) \times (n-1)$ matrix. In Example 3, for instance,

$$B' = \begin{pmatrix} 2 & 4 & 2 & 2 \\ -6 & -8 & -6 & 2 \\ -3 & -4 & -3 & 1 \end{pmatrix}.$$

By Exercise 11, B' has rank one less than B. Since $\operatorname{rank}(A) = \operatorname{rank}(B) = r$, $\operatorname{rank}(B') = r-1$. Therefore $r-1 \leq m-1$ and $r-1 \leq n-1$ by the induction hypothesis. Hence $r \leq m$ and $r \leq n$.

Also by the induction hypothesis, B' can be transformed by a finite number of elementary row and column operations into the $(m-1)\times(n-1)$ matrix D' such that

$$D' = \begin{pmatrix} I_{r-1} & O_4 \\ O_5 & O_6 \end{pmatrix},$$

where O_4, O_5, and O_6 are zero matrices. That is, D' consists of all zeros except for its first $r-1$ diagonal entries, which are ones. Let

$$D = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & D' & \\ 0 & & & \end{array}\right).$$

We see that the theorem now follows once we show that D can be obtained from B by means of a finite number of elementary row and column operations. However this follows by repeated applications of Exercise 12.

Thus, since A can be transformed into B and B can be transformed into D, each by a finite number of elementary operations, A can be transformed into D by a finite number of elementary operations.

Finally, since D' contains ones as its first $r-1$ diagonal entries, D contains ones as its first r diagonal entries and zeros elsewhere. This establishes the theorem. ∎

Corollary 1. *Let A be an $m \times n$ matrix of rank r. Then there exist invertible matrices B and C of sizes $m \times m$ and $n \times n$, respectively, such that $D = BAC$, where*

$$D = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$$

is the $m \times n$ matrix in which O_1, O_2, and O_3 are zero matrices.

Proof. By Theorem 3.6, A can be transformed by means of a finite number of elementary row and column operations into the matrix D. We can appeal to Theorem 3.1 (p. 149) each time we perform an elementary operation. Thus there exist elementary $m \times m$ matrices $E_1, E_2, \ldots, E_p$ and elementary $n \times n$ matrices $G_1, G_2, \ldots, G_q$ such that

$$D = E_p E_{p-1} \cdots E_2 E_1 A G_1 G_2 \cdots G_q.$$

By Theorem 3.2 (p. 150), each E_j and G_j is invertible. Let $B = E_p E_{p-1} \cdots E_1$ and $C = G_1 G_2 \cdots G_q$. Then B and C are invertible by Exercise 4 of Section 2.4, and $D = BAC$. ∎

Corollary 2. *Let A be an $m \times n$ matrix. Then*

(a) $\operatorname{rank}(A^t) = \operatorname{rank}(A)$.

(b) *The rank of any matrix equals the maximum number of its linearly independent rows; that is, the rank of a matrix is the dimension of the subspace generated by its rows.*

(c) *The rows and columns of any matrix generate subspaces of the same dimension, numerically equal to the rank of the matrix.*

Proof. (a) By Corollary 1, there exist invertible matrices B and C such that $D = BAC$, where D satisfies the stated conditions of the corollary. Taking transposes, we have

$$D^t = (BAC)^t = C^t A^t B^t.$$

Since B and C are invertible, so are B^t and C^t by Exercise 5 of Section 2.4. Hence by Theorem 3.4,

$$\operatorname{rank}(A^t) = \operatorname{rank}(C^t A^t B^t) = \operatorname{rank}(D^t).$$

Suppose that $r = \operatorname{rank}(A)$. Then D^t is an $n \times m$ matrix with the form of the matrix D in Corollary 1, and hence $\operatorname{rank}(D^t) = r$ by Theorem 3.5. Thus

$$\operatorname{rank}(A^t) = \operatorname{rank}(D^t) = r = \operatorname{rank}(A).$$

This establishes (a).

The proofs of (b) and (c) are left as exercises. (See Exercise 13.) ∎

Corollary 3. *Every invertible matrix is a product of elementary matrices.*

Proof. If A is an invertible $n \times n$ matrix, then $\text{rank}(A) = n$. Hence the matrix D in Corollary 1 equals I_n, and there exist invertible matrices B and C such that $I_n = BAC$.

As in the proof of Corollary 1, note that $B = E_p E_{p-1} \cdots E_1$ and $C = G_1 G_2 \cdots G_q$, where the E_i's and G_i's are elementary matrices. Thus $A = B^{-1} I_n C^{-1} = B^{-1} C^{-1}$, so that

$$A = E_1^{-1} E_2^{-1} \cdots E_p^{-1} G_q^{-1} G_{q-1}^{-1} \cdots G_1^{-1}.$$

The inverses of elementary matrices are elementary matrices, however, and hence A is the product of elementary matrices. ∎

We now use Corollary 2 to relate the rank of a matrix product to the rank of each factor. Notice how the proof exploits the relationship between the rank of a matrix and the rank of a linear transformation.

Theorem 3.7. *Let $\mathsf{T}\colon \mathsf{V} \to \mathsf{W}$ and $\mathsf{U}\colon \mathsf{W} \to \mathsf{Z}$ be linear transformations on finite-dimensional vector spaces V, W, and Z, and let A and B be matrices such that the product AB is defined. Then*

(a) $\text{rank}(\mathsf{UT}) \leq \text{rank}(\mathsf{U})$.
(b) $\text{rank}(\mathsf{UT}) \leq \text{rank}(\mathsf{T})$.
(c) $\text{rank}(AB) \leq \text{rank}(A)$.
(d) $\text{rank}(AB) \leq \text{rank}(B)$.

Proof. We prove these items in the order: (a), (c), (d), and (b).

(a) Clearly, $\mathsf{R}(\mathsf{T}) \subseteq \mathsf{W}$. Hence

$$\mathsf{R}(\mathsf{UT}) = \mathsf{UT}(\mathsf{V}) = \mathsf{U}(\mathsf{T}(\mathsf{V})) = \mathsf{U}(\mathsf{R}(\mathsf{T})) \subseteq \mathsf{U}(\mathsf{W}) = \mathsf{R}(\mathsf{U}).$$

Thus

$$\text{rank}(\mathsf{UT}) = \dim(\mathsf{R}(\mathsf{UT})) \leq \dim(\mathsf{R}(\mathsf{U})) = \text{rank}(\mathsf{U}).$$

(c) By (a),

$$\text{rank}(AB) = \text{rank}(\mathsf{L}_{AB}) = \text{rank}(\mathsf{L}_A \mathsf{L}_B) \leq \text{rank}(\mathsf{L}_A) = \text{rank}(A).$$

(d) By (c) and Corollary 2 to Theorem 3.6,

$$\text{rank}(AB) = \text{rank}((AB)^t) = \text{rank}(B^t A^t) \leq \text{rank}(B^t) = \text{rank}(B).$$

(b) Let α, β, and γ be ordered bases for V, W, and Z, respectively, and let $A' = [\mathsf{U}]_\beta^\gamma$ and $B' = [\mathsf{T}]_\alpha^\beta$. Then $A'B' = [\mathsf{UT}]_\alpha^\gamma$ by Theorem 2.11 (p. 88). Hence, by Theorem 3.3 and (d),

$$\text{rank}(\mathsf{UT}) = \text{rank}(A'B') \leq \text{rank}(B') = \text{rank}(\mathsf{T}).$$ ∎

It is important to be able to compute the rank of any matrix. We can use the corollary to Theorem 3.4, Theorems 3.5 and 3.6, and Corollary 2 to Theorem 3.6 to accomplish this goal.

The object is to perform elementary row and column operations on a matrix to "simplify" it (so that the transformed matrix has many zero entries) to the point where a simple observation enables us to determine how many linearly independent rows or columns the matrix has, and thus to determine its rank.

Example 4

(a) Let

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

Note that the first and second rows of A are linearly independent since one is not a multiple of the other. Thus $\operatorname{rank}(A) = 2$.

(b) Let

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

In this case, there are several ways to proceed. Suppose that we begin with an elementary row operation to obtain a zero in the 2,1 position. Subtracting the first row from the second row, we obtain

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

Now note that the third row is a multiple of the second row, and the first and second rows are linearly independent. Thus $\operatorname{rank}(A) = 2$.

As an alternative method, note that the first, third, and fourth columns of A are identical and that the first and second columns of A are linearly independent. Hence $\operatorname{rank}(A) = 2$.

(c) Let

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

Using elementary row operations, we can transform A as follows:

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

It is clear that the last matrix has three linearly independent rows and hence has rank 3. ♦

In summary, perform row and column operations until the matrix is simplified enough so that the maximum number of linearly independent rows or columns is obvious.

The Inverse of a Matrix

We have remarked that an $n \times n$ matrix is invertible if and only if its rank is n. Since we know how to compute the rank of any matrix, we can always test a matrix to determine whether it is invertible. We now provide a simple technique for computing the inverse of a matrix that utilizes elementary row operations.

Definition. *Let A and B be $m \times n$ and $m \times p$ matrices, respectively. By the* **augmented matrix** $(A|B)$, *we mean the $m \times (n+p)$ matrix $(A\ B)$, that is, the matrix whose first n columns are the columns of A, and whose last p columns are the columns of B.*

Let A be an invertible $n \times n$ matrix, and consider the $n \times 2n$ augmented matrix $C = (A|I_n)$. By Exercise 15, we have

$$A^{-1}C = (A^{-1}A|A^{-1}I_n) = (I_n|A^{-1}). \tag{1}$$

By Corollary 3 to Theorem 3.6, A^{-1} is the product of elementary matrices, say $A^{-1} = E_pE_{p-1}\cdots E_1$. Thus (1) becomes

$$E_pE_{p-1}\cdots E_1(A|I_n) = A^{-1}C = (I_n|A^{-1}).$$

Because multiplying a matrix on the left by an elementary matrix transforms the matrix by an elementary row operation (Theorem 3.1 p. 149), we have the following result: *If A is an invertible $n \times n$ matrix, then it is possible to transform the matrix $(A|I_n)$ into the matrix $(I_n|A^{-1})$ by means of a finite number of elementary row operations.*

Conversely, suppose that A is invertible and that, for some $n \times n$ matrix B, the matrix $(A|I_n)$ can be transformed into the matrix $(I_n|B)$ by a finite number of elementary row operations. Let $E_1, E_2, \ldots, E_p$ be the elementary matrices associated with these elementary row operations as in Theorem 3.1; then

$$E_pE_{p-1}\cdots E_1(A|I_n) = (I_n|B). \tag{2}$$

Letting $M = E_pE_{p-1}\cdots E_1$, we have from (2) that

$$(MA|M) = M(A|I_n) = (I_n|B).$$

Hence $MA = I_n$ and $M = B$. It follows that $M = A^{-1}$. So $B = A^{-1}$. Thus we have the following result: *If A is an invertible $n \times n$ matrix, and the matrix $(A|I_n)$ is transformed into a matrix of the form $(I_n|B)$ by means of a finite number of elementary row operations, then $B = A^{-1}$.*

If, on the other hand, A is an $n \times n$ matrix that is not invertible, then $\operatorname{rank}(A) < n$. Hence any attempt to transform $(A|I_n)$ into a matrix of the form $(I_n|B)$ by means of elementary row operations must fail because otherwise A can be transformed into I_n using the same row operations. This is impossible, however, because elementary row operations preserve rank. In fact, A can be transformed into a matrix with a row containing only zero entries, yielding the following result: *If A is an $n \times n$ matrix that is not invertible, then any attempt to transform $(A|I_n)$ into a matrix of the form $(I_n|B)$ produces a row whose first n entries are zeros.*

The next two examples demonstrate these comments.

Example 5

We determine whether the matrix

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

is invertible, and if it is, we compute its inverse.

We attempt to use elementary row operations to transform

$$(A|I) = \left(\begin{array}{ccc|ccc} 0 & 2 & 4 & 1 & 0 & 0 \\ 2 & 4 & 2 & 0 & 1 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array}\right)$$

into a matrix of the form $(I|B)$. One method for accomplishing this transformation is to change each column of A successively, beginning with the first column, into the corresponding column of I. Since we need a nonzero entry in the 1,1 position, we begin by interchanging rows 1 and 2. The result is

$$\left(\begin{array}{ccc|ccc} 2 & 4 & 2 & 0 & 1 & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array}\right).$$

In order to place a 1 in the 1,1 position, we must multiply the first row by $\frac{1}{2}$; this operation yields

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 3 & 3 & 1 & 0 & 0 & 1 \end{array}\right).$$

We now complete work in the first column by adding -3 times row 1 to row 3 to obtain

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & 2 & 4 & 1 & 0 & 0 \\ 0 & -3 & -2 & 0 & -\frac{3}{2} & 1 \end{array}\right).$$

In order to change the second column of the preceding matrix into the second column of I, we multiply row 2 by $\frac{1}{2}$ to obtain a 1 in the 2,2 position. This operation produces

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0 \\ 0 & -3 & -2 & 0 & -\frac{3}{2} & 1 \end{array}\right).$$

We now complete our work on the second column by adding -2 times row 2 to row 1 and 3 times row 2 to row 3. The result is

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -3 & -1 & \frac{1}{2} & 0 \\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 4 & \frac{3}{2} & -\frac{3}{2} & 1 \end{array}\right).$$

Only the third column remains to be changed. In order to place a 1 in the 3,3 position, we multiply row 3 by $\frac{1}{4}$; this operation yields

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -3 & -1 & \frac{1}{2} & 0 \\ 0 & 1 & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & \frac{3}{8} & -\frac{3}{8} & \frac{1}{4} \end{array}\right).$$

Adding appropriate multiples of row 3 to rows 1 and 2 completes the process and gives

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{8} & -\frac{5}{8} & \frac{3}{4} \\ 0 & 1 & 0 & -\frac{1}{4} & \frac{3}{4} & -\frac{1}{2} \\ 0 & 0 & 1 & \frac{3}{8} & -\frac{3}{8} & \frac{1}{4} \end{array}\right).$$

Thus A is invertible, and

$$A^{-1} = \begin{pmatrix} \frac{1}{8} & -\frac{5}{8} & \frac{3}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{2} \\ \frac{3}{8} & -\frac{3}{8} & \frac{1}{4} \end{pmatrix}. \quad \blacklozenge$$

Example 6

We determine whether the matrix

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

is invertible, and if it is, we compute its inverse. Using a strategy similar to the one used in Example 5, we attempt to use elementary row operations to transform

$$(A|I) = \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 1 & 0 \\ 1 & 5 & 4 & 0 & 0 & 1 \end{array}\right)$$

into a matrix of the form $(I|B)$. We first add -2 times row 1 to row 2 and -1 times row 1 to row 3. We then add row 2 to row 3. The result,

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 1 & 0 \\ 1 & 5 & 4 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -3 & -3 & -2 & 1 & 0 \\ 0 & 3 & 3 & -1 & 0 & 1 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & -3 & -3 & -2 & 1 & 0 \\ 0 & 0 & 0 & -3 & 1 & 1 \end{array}\right),$$

is a matrix with a row whose first 3 entries are zeros. Therefore A is not invertible. ♦

Being able to test for invertibility and compute the inverse of a matrix allows us, with the help of Theorem 2.18 (p. 101) and its corollaries, to test for invertibility and compute the inverse of a linear transformation. The next example demonstrates this technique.

Example 7

Let $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_2(R)$ be defined by $\mathsf{T}(f(x)) = f(x) + f'(x) + f''(x)$, where $f'(x)$ and $f''(x)$ denote the first and second derivatives of $f(x)$. We use Corollary 1 of Theorem 2.18 (p. 102) to test T for invertibility and compute the inverse if T is invertible. Taking β to be the standard ordered basis of $\mathsf{P}_2(R)$, we have

$$[\mathsf{T}]_\beta = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Using the method of Examples 5 and 6, we can show that $[\mathsf{T}]_\beta$ is invertible with inverse

$$([\mathsf{T}]_\beta)^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus T is invertible, and $([\mathsf{T}]_\beta)^{-1} = [\mathsf{T}^{-1}]_\beta$. Hence by Theorem 2.14 (p. 91), we have

$$\begin{aligned}[\mathsf{T}^{-1}(a_0 + a_1x + a_2x^2)]_\beta &= \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} \\ &= \begin{pmatrix} a_0 - a_1 \\ a_1 - 2a_2 \\ a_2 \end{pmatrix}.\end{aligned}$$

Therefore

$$\mathsf{T}^{-1}(a_0 + a_1x + a_2x^2) = (a_0 - a_1) + (a_1 - 2a_2)x + a_2x^2. \quad \blacklozenge$$

EXERCISES

1. Label the following statements as true or false.
   - **(a)** The rank of a matrix is equal to the number of its nonzero columns.
   - **(b)** The product of two matrices always has rank equal to the lesser of the ranks of the two matrices.
   - **(c)** The $m \times n$ zero matrix is the only $m \times n$ matrix having rank 0.
   - **(d)** Elementary row operations preserve rank.
   - **(e)** Elementary column operations do not necessarily preserve rank.
   - **(f)** The rank of a matrix is equal to the maximum number of linearly independent rows in the matrix.
   - **(g)** The inverse of a matrix can be computed exclusively by means of elementary row operations.
   - **(h)** The rank of an $n \times n$ matrix is at most n.
   - **(i)** An $n \times n$ matrix having rank n is invertible.

2. Find the rank of the following matrices.

   (a) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ **(b)** $\begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ **(c)** $\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 4 \end{pmatrix}$

(d) $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & 2 & 3 & 1 & 1 \\ 1 & 4 & 0 & 1 & 2 \\ 0 & 2 & -3 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$

(f) $\begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 2 & 4 & 1 & 3 & 0 \\ 3 & 6 & 2 & 5 & 1 \\ -4 & -8 & 1 & -3 & 1 \end{pmatrix}$ (g) $\begin{pmatrix} 1 & 1 & 0 & 1 \\ 2 & 2 & 0 & 2 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$

3. Prove that for any $m \times n$ matrix A, $\text{rank}(A) = 0$ if and only if A is the zero matrix.

4. Use elementary row and column operations to transform each of the following matrices into a matrix D satisfying the conditions of Theorem 3.6, and then determine the rank of each matrix.

(a) $\begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 0 & -1 & 2 \\ 1 & 1 & 1 & 2 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & 1 \\ -1 & 2 \\ 2 & 1 \end{pmatrix}$

5. For each of the following matrices, compute the rank and the inverse if it exists.

(a) $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 3 & 4 \\ 2 & 3 & -1 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & -2 & 4 \\ 1 & 1 & -1 \\ 2 & 4 & -5 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 5 & 5 & 1 \\ -2 & -3 & 0 & 3 \\ 3 & 4 & -2 & -3 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 1 & -1 & 2 \\ 2 & 0 & 1 & 0 \\ 0 & -1 & 1 & -3 \end{pmatrix}$

6. For each of the following linear transformations T, determine whether T is invertible, and compute T^{-1} if it exists.

(a) $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_2(R)$ defined by $\mathsf{T}(f(x)) = f''(x) + 2f'(x) - f(x)$.
(b) $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{P}_2(R)$ defined by $\mathsf{T}(f(x)) = (x+1)f'(x)$.
(c) $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^3$ defined by

$$\mathsf{T}(a_1, a_2, a_3) = (a_1 + 2a_2 + a_3, -a_1 + a_2 + 2a_3, a_1 + a_3).$$

(d) $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{P}_2(R)$ defined by

$$\mathsf{T}(a_1, a_2, a_3) = (a_1 + a_2 + a_3) + (a_1 - a_2 + a_3)x + a_1x^2.$$

(e) $\mathsf{T}\colon \mathsf{P}_2(R) \to \mathsf{R}^3$ defined by $\mathsf{T}(f(x)) = (f(-1), f(0), f(1))$.

(f) $\mathsf{T}\colon \mathsf{M}_{2\times 2}(R) \to \mathsf{R}^4$ defined by

$$\mathsf{T}(A) = (\operatorname{tr}(A), \operatorname{tr}(A^t), \operatorname{tr}(EA), \operatorname{tr}(AE)),$$

where

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

7. Express the invertible matrix

$$\begin{pmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

as a product of elementary matrices.

8. Let A be an $m \times n$ matrix. Prove that if c is any nonzero scalar, then $\operatorname{rank}(cA) = \operatorname{rank}(A)$.

9. Complete the proof of the corollary to Theorem 3.4 by showing that elementary column operations preserve rank.

10. Prove Theorem 3.6 for the case that A is an $m \times 1$ matrix.

11. Let

$$B = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{array}\right),$$

where B' is an $m \times n$ submatrix of B. Prove that if $\operatorname{rank}(B) = r$, then $\operatorname{rank}(B') = r - 1$.

12. Let B' and D' be $m \times n$ matrices, and let B and D be $(m+1) \times (n+1)$ matrices respectively defined by

$$B = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B' & \\ 0 & & & \end{array}\right) \quad \text{and} \quad D = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & D' & \\ 0 & & & \end{array}\right).$$

Prove that if B' can be transformed into D' by an elementary row [column] operation, then B can be transformed into D by an elementary row [column] operation.

13. Prove (b) and (c) of Corollary 2 to Theorem 3.6.

14. Let $\mathsf{T}, \mathsf{U}: \mathsf{V} \to \mathsf{W}$ be linear transformations.

(a) Prove that $\mathsf{R}(\mathsf{T}+\mathsf{U}) \subseteq \mathsf{R}(\mathsf{T})+\mathsf{R}(\mathsf{U})$. (See the definition of the sum of subsets of a vector space on page 22.)

(b) Prove that if W is finite-dimensional, then $\operatorname{rank}(\mathsf{T}+\mathsf{U}) \leq \operatorname{rank}(\mathsf{T})+\operatorname{rank}(\mathsf{U})$.

(c) Deduce from (b) that $\operatorname{rank}(A+B) \leq \operatorname{rank}(A)+\operatorname{rank}(B)$ for any $m \times n$ matrices A and B.

15. Suppose that A and B are matrices having n rows. Prove that $M(A|B) = (MA|MB)$ for any $m \times n$ matrix M.

16. Supply the details to the proof of (b) of Theorem 3.4.

17. Prove that if B is a 3×1 matrix and C is a 1×3 matrix, then the 3×3 matrix BC has rank at most 1. Conversely, show that if A is any 3×3 matrix having rank 1, then there exist a 3×1 matrix B and a 1×3 matrix C such that $A = BC$.

18. Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Prove that AB can be written as a sum of n matrices of rank at most one.

19. Let A be an $m \times n$ matrix with rank m and B be an $n \times p$ matrix with rank n. Determine the rank of AB. Justify your answer.

20. Let

$$A = \begin{pmatrix} 1 & 0 & -1 & 2 & 1 \\ -1 & 1 & 3 & -1 & 0 \\ -2 & 1 & 4 & -1 & 3 \\ 3 & -1 & -5 & 1 & -6 \end{pmatrix}.$$

(a) Find a 5×5 matrix M with rank 2 such that $AM = O$, where O is the 4×5 zero matrix.

(b) Suppose that B is a 5×5 matrix such that $AB = O$. Prove that $\operatorname{rank}(B) \leq 2$.

21. Let A be an $m \times n$ matrix with rank m. Prove that there exists an $n \times m$ matrix B such that $AB = I_m$.

22. Let B be an $n \times m$ matrix with rank m. Prove that there exists an $m \times n$ matrix A such that $AB = I_m$.

3.3 SYSTEMS OF LINEAR EQUATIONS—THEORETICAL ASPECTS

This section and the next are devoted to the study of systems of linear equations, which arise naturally in both the physical and social sciences. In this section, we apply results from Chapter 2 to describe the solution sets of systems of linear equations as subsets of a vector space. In Section 3.4, elementary row operations are used to provide a computational method for finding all solutions to such systems.

The system of equations

$$(S)\qquad \begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned}$$

where a_{ij} and b_i ($1 \leq i \leq m$ and $1 \leq j \leq n$) are scalars in a field F and $x_1, x_2, \ldots, x_n$ are n variables taking values in F, is called a **system of m linear equations in n unknowns over the field F.**

The $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

is called the **coefficient matrix** of the system (S).

If we let

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

then the system (S) may be rewritten as a single matrix equation

$$Ax = b.$$

To exploit the results that we have developed, we often consider a system of linear equations as a single matrix equation.

A **solution** to the system (S) is an n-tuple

$$s = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \in \mathsf{F}^n$$

such that $As = b$. The set of all solutions to the system (S) is called the **solution set** of the system. System (S) is called **consistent** if its solution set is nonempty; otherwise it is called **inconsistent**.

Example 1

(a) Consider the system

$$\begin{aligned} x_1 + x_2 &= 3 \\ x_1 - x_2 &= 1. \end{aligned}$$

By use of familiar techniques, we can solve the preceding system and conclude that there is only one solution: $x_1 = 2$, $x_2 = 1$; that is,

$$s = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

In matrix form, the system can be written

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix};$$

so

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

(b) Consider

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 1 \\ x_1 - x_2 + 2x_3 &= 6; \end{aligned}$$

that is,

$$\begin{pmatrix} 2 & 3 & 1 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 6 \end{pmatrix}.$$

This system has many solutions, such as

$$s = \begin{pmatrix} -6 \\ 2 \\ 7 \end{pmatrix} \quad \text{and} \quad s = \begin{pmatrix} 8 \\ -4 \\ -3 \end{pmatrix}.$$

(c) Consider

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 + x_2 &= 1; \end{aligned}$$

that is,

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

It is evident that this system has no solutions. Thus we see that a system of linear equations can have one, many, or no solutions. ♦

We must be able to recognize when a system has a solution and then be able to describe all its solutions. This section and the next are devoted to this end.

We begin our study of systems of linear equations by examining the class of *homogeneous* systems of linear equations. Our first result (Theorem 3.8) shows that the set of solutions to a homogeneous system of m linear equations in n unknowns forms a subspace of F^n. We can then apply the theory of vector spaces to this set of solutions. For example, a basis for the solution space can be found, and any solution can be expressed as a linear combination of the vectors in the basis.

Definitions. *A system $Ax = b$ of m linear equations in n unknowns is said to be* ***homogeneous*** *if $b = 0$. Otherwise the system is said to be* ***nonhomogeneous.***

Any homogeneous system has at least one solution, namely, the zero vector. The next result gives further information about the set of solutions to a homogeneous system.

Theorem 3.8. *Let $Ax = 0$ be a homogeneous system of m linear equations in n unknowns over a field F. Let K denote the set of all solutions to $Ax = 0$. Then $\mathsf{K} = \mathsf{N}(\mathsf{L}_A)$; hence K is a subspace of F^n of dimension $n - \operatorname{rank}(\mathsf{L}_A) = n - \operatorname{rank}(A)$.*

Proof. Clearly, $\mathsf{K} = \{s \in \mathsf{F}^n : As = 0\} = \mathsf{N}(\mathsf{L}_A)$. The second part now follows from the dimension theorem (p. 70). ∎

Corollary. *If $m < n$, the system $Ax = 0$ has a nonzero solution.*

Proof. Suppose that $m < n$. Then $\operatorname{rank}(A) = \operatorname{rank}(\mathsf{L}_A) \leq m$. Hence

$$\dim(\mathsf{K}) = n - \operatorname{rank}(\mathsf{L}_A) \geq n - m > 0,$$

where $\mathsf{K} = \mathsf{N}(\mathsf{L}_A)$. Since $\dim(\mathsf{K}) > 0$, $\mathsf{K} \neq \{0\}$. Thus there exists a nonzero vector $s \in \mathsf{K}$; so s is a nonzero solution to $Ax = 0$. ∎

Example 2

(a) Consider the system

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 0 \\ x_1 - x_2 - x_3 &= 0. \end{aligned}$$

Let

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & -1 & -1 \end{pmatrix}$$

be the coefficient matrix of this system. It is clear that $\operatorname{rank}(A) = 2$. If K is the solution set of this system, then $\dim(\mathsf{K}) = 3 - 2 = 1$. Thus any nonzero solution constitutes a basis for K. For example, since

$$\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$$

is a solution to the given system,

$$\left\{ \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \right\}$$

is a basis for K. Thus any vector in K is of the form

$$t \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} t \\ -2t \\ 3t \end{pmatrix},$$

where $t \in R$.

(b) Consider the system $x_1 - 2x_2 + x_3 = 0$ of one equation in three unknowns. If $A = \begin{pmatrix} 1 & -2 & 1 \end{pmatrix}$ is the coefficient matrix, then $\operatorname{rank}(A) = 1$. Hence if K is the solution set, then $\dim(\mathsf{K}) = 3 - 1 = 2$. Note that

$$\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

are linearly independent vectors in K. Thus they constitute a basis for K, so that

$$\mathsf{K} = \left\{ t_1 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} : t_1, t_2 \in R \right\}. \quad \blacklozenge$$

In Section 3.4, explicit computational methods for finding a basis for the solution set of a homogeneous system are discussed.

We now turn to the study of nonhomogeneous systems. Our next result shows that the solution set of a nonhomogeneous system $Ax = b$ can be described in terms of the solution set of the homogeneous system $Ax = 0$. We refer to the equation $Ax = 0$ as the **homogeneous system corresponding to** $Ax = b$.

Theorem 3.9. *Let K be the solution set of a system of linear equations $Ax = b$, and let $\mathsf{K_H}$ be the solution set of the corresponding homogeneous system $Ax = 0$. Then for any solution s to $Ax = b$*

$$K = \{s\} + \mathsf{K_H} = \{s + k \colon k \in \mathsf{K_H}\}.$$

Proof. Let s be any solution to $Ax = b$. We must show that $K = \{s\} + \mathsf{K_H}$. If $w \in K$, then $Aw = b$. Hence

$$A(w - s) = Aw - As = b - b = 0.$$

So $w - s \in \mathsf{K_H}$. Thus there exists $k \in \mathsf{K_H}$ such that $w - s = k$. It follows that $w = s + k \in \{s\} + \mathsf{K_H}$, and therefore

$$K \subseteq \{s\} + \mathsf{K_H}.$$

Conversely, suppose that $w \in \{s\} + \mathsf{K_H}$; then $w = s + k$ for some $k \in \mathsf{K_H}$. But then $Aw = A(s + k) = As + Ak = b + 0 = b$; so $w \in K$. Therefore $\{s\} + \mathsf{K_H} \subseteq K$, and thus $K = \{s\} + \mathsf{K_H}$. ∎

Example 3

(a) Consider the system

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 7 \\ x_1 - x_2 - x_3 &= -4. \end{aligned}$$

The corresponding homogeneous system is the system in Example 2(a). It is easily verified that

$$s = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}$$

is a solution to the preceding nonhomogeneous system. So the solution set of the system is

$$\mathsf{K} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} + t \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} : t \in R \right\}$$

by Theorem 3.9.

(b) Consider the system $x_1 - 2x_2 + x_3 = 4$. The corresponding homogeneous system is the system in Example 2(b). Since

$$s = \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}$$

is a solution to the given system, the solution set K can be written as

$$\mathsf{K} = \left\{ \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} : t_1, t_2 \in R \right\}. \quad \blacklozenge$$

The following theorem provides us with a means of computing solutions to certain systems of linear equations.

Theorem 3.10. *Let $Ax = b$ be a system of n linear equations in n unknowns. If A is invertible, then the system has exactly one solution, namely, $A^{-1}b$. Conversely, if the system has exactly one solution, then A is invertible.*

Proof. Suppose that A is invertible. Substituting $A^{-1}b$ into the system, we have $A(A^{-1}b) = (AA^{-1})b = b$. Thus $A^{-1}b$ is a solution. If s is an arbitrary solution, then $As = b$. Multiplying both sides by A^{-1} gives $s = A^{-1}b$. Thus the system has one and only one solution, namely, $A^{-1}b$.

Conversely, suppose that the system has exactly one solution s. Let $\mathsf{K_H}$ denote the solution set for the corresponding homogeneous system $Ax = 0$. By Theorem 3.9, $\{s\} = \{s\} + \mathsf{K_H}$. But this is so only if $\mathsf{K_H} = \{0\}$. Thus $\mathsf{N}(\mathsf{L}_A) = \{0\}$, and hence A is invertible. ∎

Example 4

Consider the following system of three linear equations in three unknowns:

$$\begin{aligned} 2x_2 + 4x_3 &= 2 \\ 2x_1 + 4x_2 + 2x_3 &= 3 \\ 3x_1 + 3x_2 + x_3 &= 1. \end{aligned}$$

In Example 5 of Section 3.2, we computed the inverse of the coefficient matrix A of this system. Thus the system has exactly one solution, namely,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A^{-1}b = \begin{pmatrix} \frac{1}{8} & -\frac{5}{8} & \frac{3}{4} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{2} \\ \frac{3}{8} & -\frac{3}{8} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{7}{8} \\ \frac{5}{4} \\ -\frac{1}{8} \end{pmatrix}. \quad \blacklozenge$$

We use this technique for solving systems of linear equations having invertible coefficient matrices in the application that concludes this section.

In Example 1(c), we saw a system of linear equations that has no solutions. We now establish a criterion for determining when a system has solutions. This criterion involves the rank of the coefficient matrix of the system $Ax = b$ and the rank of the matrix $(A|b)$. The matrix $(A|b)$ is called the **augmented matrix of the system** $Ax = b$.

Theorem 3.11. *Let $Ax = b$ be a system of linear equations. Then the system is consistent if and only if* $\operatorname{rank}(A) = \operatorname{rank}(A|b)$.

Proof. To say that $Ax = b$ has a solution is equivalent to saying that $b \in \mathsf{R}(\mathsf{L}_A)$. (See Exercise 9.) In the proof of Theorem 3.5 (p. 153), we saw that

$$\mathsf{R}(\mathsf{L}_A) = \operatorname{span}(\{a_1, a_2, \ldots, a_n\}),$$

the span of the columns of A. Thus $Ax = b$ has a solution if and only if $b \in \text{span}(\{a_1, a_2, \ldots, a_n\})$. But $b \in \text{span}(\{a_1, a_2, \ldots, a_n\})$ if and only if $\text{span}(\{a_1, a_2, \ldots, a_n\}) = \text{span}(\{a_1, a_2, \ldots, a_n, b\})$. This last statement is equivalent to

$$\dim(\text{span}(\{a_1, a_2, \ldots, a_n\})) = \dim(\text{span}(\{a_1, a_2, \ldots, a_n, b\})).$$

So by Theorem 3.5, the preceding equation reduces to

$$\text{rank}(A) = \text{rank}(A|b). \qquad \blacksquare$$

Example 5

Recall the system of equations

$$\begin{aligned} x_1 + x_2 &= 0 \\ x_1 + x_2 &= 1 \end{aligned}$$

in Example 1(c).

Since

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad (A|b) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix},$$

$\text{rank}(A) = 1$ and $\text{rank}(A|b) = 2$. Because the two ranks are unequal, the system has no solutions. ♦

Example 6

We can use Theorem 3.11 to determine whether $(3, 3, 2)$ is in the range of the linear transformation $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^3$ defined by

$$\mathsf{T}(a_1, a_2, a_3) = (a_1 + a_2 + a_3, a_1 - a_2 + a_3, a_1 + a_3).$$

Now $(3, 3, 2) \in \mathsf{R}(\mathsf{T})$ if and only if there exists a vector $s = (x_1, x_2, x_3)$ in R^3 such that $\mathsf{T}(s) = (3, 3, 2)$. Such a vector s must be a solution to the system

$$\begin{aligned} x_1 + x_2 + x_3 &= 3 \\ x_1 - x_2 + x_3 &= 3 \\ x_1 \qquad + x_3 &= 2. \end{aligned}$$

Since the ranks of the coefficient matrix and the augmented matrix of this system are 2 and 3, respectively, it follows that this system has no solutions. Hence $(3, 3, 2) \notin \mathsf{R}(\mathsf{T})$. ♦

An Application

In 1973, Wassily Leontief won the Nobel prize in economics for his work in developing a mathematical model that can be used to describe various economic phenomena. We close this section by applying some of the ideas we have studied to illustrate two special cases of his work.

We begin by considering a simple society composed of three people (industries)—a farmer who grows all the food, a tailor who makes all the clothing, and a carpenter who builds all the housing. We assume that each person sells to and buys from a central pool and that everything produced is consumed. Since no commodities either enter or leave the system, this case is referred to as the **closed model**.

Each of these three individuals consumes all three of the commodities produced in the society. Suppose that the proportion of each of the commodities consumed by each person is given in the following table. Notice that each of the columns of the table must sum to 1.

	Food	Clothing	Housing
Farmer	0.40	0.20	0.20
Tailor	0.10	0.70	0.20
Carpenter	0.50	0.10	0.60

Let p_1, p_2, and p_3 denote the incomes of the farmer, tailor, and carpenter, respectively. To ensure that this society survives, we require that the consumption of each individual equals his or her income. Note that the farmer consumes 20% of the clothing. Because the total cost of all clothing is p_2, the tailor's income, the amount spent by the farmer on clothing is $0.20p_2$. Moreover, the amount spent by the farmer on food, clothing, and housing must equal the farmer's income, and so we obtain the equation

$$0.40p_1 + 0.20p_2 + 0.20p_3 = p_1.$$

Similar equations describing the expenditures of the tailor and carpenter produce the following system of linear equations:

$$\begin{aligned} 0.40p_1 + 0.20p_2 + 0.20p_3 &= p_1 \\ 0.10p_1 + 0.70p_2 + 0.20p_3 &= p_2 \\ 0.50p_1 + 0.10p_2 + 0.60p_3 &= p_3. \end{aligned}$$

This system can be written as $Ap = p$, where

$$p = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

and A is the coefficient matrix of the system. In this context, A is called the **input–output (or consumption) matrix**, and $Ap = p$ is called the **equilibrium condition.**

For vectors $b = (b_1, b_2, \ldots, b_n)$ and $c = (c_1, c_2, \ldots, c_n)$ in R^n, we use the notation $b \geq c$ $[b > c]$ to mean $b_i \geq c_i$ $[b_i > c_i]$ for all i. The vector b is called **nonnegative [positive]** if $b \geq 0$ $[b > 0]$.

At first, it may seem reasonable to replace the equilibrium condition by the inequality $Ap \leq p$, that is, the requirement that consumption not exceed production. But, in fact, $Ap \leq p$ implies that $Ap = p$ in the closed model. For otherwise, there exists a k for which

$$p_k > \sum_j A_{kj} p_j.$$

Hence, since the columns of A sum to 1,

$$\sum_i p_i > \sum_i \sum_j A_{ij} p_j = \sum_j \left(\sum_i A_{ij} \right) p_j = \sum_j p_j,$$

which is a contradiction.

One solution to the homogeneous system $(I - A)x = 0$, which is equivalent to the equilibrium condition, is

$$p = \begin{pmatrix} 0.25 \\ 0.35 \\ 0.40 \end{pmatrix}.$$

We may interpret this to mean that the society survives if the farmer, tailor, and carpenter have incomes in the proportions $25 : 35 : 40$ (or $5 : 7 : 8$).

Notice that we are not simply interested in any nonzero solution to the system, but in one that is nonnegative. Thus we must consider the question of whether the system $(I - A)x = 0$ has a nonnegative solution, where A is a matrix with nonnegative entries whose columns sum to 1. A useful theorem in this direction (whose proof may be found in "Applications of Matrices to Economic Models and Social Science Relationships," by Ben Noble, *Proceedings of the Summer Conference for College Teachers on Applied Mathematics*, 1971, CUPM, Berkeley, California) is stated below.

Theorem 3.12. *Let A be an $n \times n$ input–output matrix having the form*

$$A = \begin{pmatrix} B & C \\ D & E \end{pmatrix},$$

where D is a $1 \times (n-1)$ positive vector and C is an $(n-1) \times 1$ positive vector. Then $(I - A)x = 0$ has a one-dimensional solution set that is generated by a nonnegative vector.

Observe that any input–output matrix with all positive entries satisfies the hypothesis of this theorem. The following matrix does also:

$$\begin{pmatrix} 0.75 & 0.50 & 0.65 \\ 0 & 0.25 & 0.35 \\ 0.25 & 0.25 & 0 \end{pmatrix}.$$

In the **open model**, we assume that there is an outside demand for each of the commodities produced. Returning to our simple society, let x_1, x_2, and x_3 be the monetary values of food, clothing, and housing produced with respective outside demands d_1, d_2, and d_3. Let A be the 3×3 matrix such that A_{ij} represents the amount (in a fixed monetary unit such as the dollar) of commodity i required to produce one monetary unit of commodity j. Then the value of the surplus of food in the society is

$$x_1 - (A_{11}x_1 + A_{12}x_2 + A_{13}x_3),$$

that is, the value of food produced minus the value of food consumed while producing the three commodities. The assumption that everything produced is consumed gives us a similar equilibrium condition for the open model, namely, that the surplus of each of the three commodities must equal the corresponding outside demands. Hence

$$x_i - \sum_{j=1}^{3} A_{ij}x_j = d_i \quad \text{for } i = 1, 2, 3.$$

In general, we must find a nonnegative solution to $(I - A)x = d$, where A is a matrix with nonnegative entries such that the sum of the entries of each column of A does not exceed one, and $d \geq 0$. It is easy to see that if $(I - A)^{-1}$ exists and is nonnegative, then the desired solution is $(I - A)^{-1}d$.

Recall that for a real number a, the series $1 + a + a^2 + \cdots$ converges to $(1 - a)^{-1}$ if $|a| < 1$. Similarly, it can be shown (using the concept of convergence of matrices developed in Section 5.3) that the series $I + A + A^2 + \cdots$ converges to $(I - A)^{-1}$ if $\{A^n\}$ converges to the zero matrix. In this case, $(I - A)^{-1}$ is nonnegative since the matrices $I, A, A^2, \ldots$ are nonnegative.

To illustrate the open model, suppose that 30 cents worth of food, 10 cents worth of clothing, and 30 cents worth of housing are required for the production of \$1 worth of food. Similarly, suppose that 20 cents worth of food, 40 cents worth of clothing, and 20 cents worth of housing are required for the production of \$1 of clothing. Finally, suppose that 30 cents worth of food, 10 cents worth of clothing, and 30 cents worth of housing are required for the production of \$1 worth of housing. Then the input–output matrix is

$$A = \begin{pmatrix} 0.30 & 0.20 & 0.30 \\ 0.10 & 0.40 & 0.10 \\ 0.30 & 0.20 & 0.30 \end{pmatrix};$$

so

$$I-A=\begin{pmatrix} 0.70 & -0.20 & -0.30\\ -0.10 & 0.60 & -0.10\\ -0.30 & -0.20 & 0.70\end{pmatrix} \quad\text{and}\quad (I-A)^{-1}=\begin{pmatrix} 2.0 & 1.0 & 1.0\\ 0.5 & 2.0 & 0.5\\ 1.0 & 1.0 & 2.0\end{pmatrix}.$$

Since $(I-A)^{-1}$ is nonnegative, we can find a (unique) nonnegative solution to $(I-A)x=d$ for any demand d. For example, suppose that there are outside demands for \$30 billion in food, \$20 billion in clothing, and \$10 billion in housing. If we set

$$d=\begin{pmatrix}30\\20\\10\end{pmatrix},$$

then

$$x=(I-A)^{-1}d=\begin{pmatrix}90\\60\\70\end{pmatrix}.$$

So a gross production of \$90 billion of food, \$60 billion of clothing, and \$70 billion of housing is necessary to meet the required demands.

EXERCISES

1. Label the following statements as true or false.
   (a) Any system of linear equations has at least one solution.
   (b) Any system of linear equations has at most one solution.
   (c) Any homogeneous system of linear equations has at least one solution.
   (d) Any system of n linear equations in n unknowns has at most one solution.
   (e) Any system of n linear equations in n unknowns has at least one solution.
   (f) If the homogeneous system corresponding to a given system of linear equations has a solution, then the given system has a solution.
   (g) If the coefficient matrix of a homogeneous system of n linear equations in n unknowns is invertible, then the system has no nonzero solutions.
   (h) The solution set of any system of m linear equations in n unknowns is a subspace of F^n.

2. For each of the following homogeneous systems of linear equations, find the dimension of and a basis for the solution set.

(a) $\begin{aligned} x_1 + 3x_2 &= 0 \\ 2x_1 + 6x_2 &= 0 \end{aligned}$

(b) $\begin{aligned} x_1 + x_2 - x_3 &= 0 \\ 4x_1 + x_2 - 2x_3 &= 0 \end{aligned}$

(c) $\begin{aligned} x_1 + 2x_2 - x_3 &= 0 \\ 2x_1 + x_2 + x_3 &= 0 \end{aligned}$

(d) $\begin{aligned} 2x_1 + x_2 - x_3 &= 0 \\ x_1 - x_2 + x_3 &= 0 \\ x_1 + 2x_2 - 2x_3 &= 0 \end{aligned}$

(e) $x_1 + 2x_2 - 3x_3 + x_4 = 0$

(f) $\begin{aligned} x_1 + 2x_2 &= 0 \\ x_1 - x_2 &= 0 \end{aligned}$

(g) $\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 0 \\ x_2 - x_3 + x_4 &= 0 \end{aligned}$

3. Using the results of Exercise 2, find all solutions to the following systems.

(a) $\begin{aligned} x_1 + 3x_2 &= 5 \\ 2x_1 + 6x_2 &= 10 \end{aligned}$

(b) $\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ 4x_1 + x_2 - 2x_3 &= 3 \end{aligned}$

(c) $\begin{aligned} x_1 + 2x_2 - x_3 &= 3 \\ 2x_1 + x_2 + x_3 &= 6 \end{aligned}$

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

(e) $x_1 + 2x_2 - 3x_3 + x_4 = 1$

(f) $\begin{aligned} x_1 + 2x_2 &= 5 \\ x_1 - x_2 &= -1 \end{aligned}$

(g) $\begin{aligned} x_1 + 2x_2 + x_3 + x_4 &= 1 \\ x_2 - x_3 + x_4 &= 1 \end{aligned}$

4. For each system of linear equations with the invertible coefficient matrix A,

(1) Compute A^{-1}.

(2) Use A^{-1} to solve the system.

(a) $\begin{aligned} x_1 + 3x_2 &= 4 \\ 2x_1 + 5x_2 &= 3 \end{aligned}$

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

5. Give an example of a system of n linear equations in n unknowns with infinitely many solutions.

6. Let $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^2$ be defined by $\mathsf{T}(a, b, c) = (a + b, 2a - c)$. Determine $\mathsf{T}^{-1}(1, 11)$.

7. Determine which of the following systems of linear equations has a solution.

(a) $$\begin{aligned} x_1 + x_2 - x_3 + 2x_4 &= 2 \\ x_1 + x_2 + 2x_3 &= 1 \\ 2x_1 + 2x_2 + x_3 + 2x_4 &= 4 \end{aligned}$$

(b) $$\begin{aligned} x_1 + x_2 - x_3 &= 1 \\ 2x_1 + x_2 + 3x_3 &= 2 \end{aligned}$$

(c) $$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 1 \\ x_1 + x_2 - x_3 &= 0 \\ x_1 + 2x_2 + x_3 &= 3 \end{aligned}$$

(d) $$\begin{aligned} x_1 + x_2 + 3x_3 - x_4 &= 0 \\ x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 - 2x_2 + x_3 - x_4 &= 1 \\ 4x_1 + x_2 + 8x_3 - x_4 &= 0 \end{aligned}$$

(e) $$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ 2x_1 + x_2 + 2x_3 &= 3 \\ x_1 - 4x_2 + 7x_3 &= 4 \end{aligned}$$

8. Let $\mathsf{T}\colon \mathsf{R}^3 \to \mathsf{R}^3$ be defined by $\mathsf{T}(a,b,c) = (a+b, b-2c, a+2c)$. For each vector v in R^3, determine whether $v \in \mathsf{R}(\mathsf{T})$.

   (a) $v = (1,3,-2)$ (b) $v = (2,1,1)$

9. Prove that the system of linear equations $Ax = b$ has a solution if and only if $b \in \mathsf{R}(\mathsf{L}_A)$.

10. Prove or give a counterexample to the following statement: If the coefficient matrix of a system of m linear equations in n unknowns has rank m, then the system has a solution.

11. In the closed model of Leontief with food, clothing, and housing as the basic industries, suppose that the input–output matrix is

$$A = \begin{pmatrix} \frac{7}{16} & \frac{1}{2} & \frac{3}{16} \\ \frac{5}{16} & \frac{1}{6} & \frac{5}{16} \\ \frac{1}{4} & \frac{1}{3} & \frac{1}{2} \end{pmatrix}.$$

   At what ratio must the farmer, tailor, and carpenter produce in order for equilibrium to be attained?

12. A certain economy consists of two sectors: goods and services. Suppose that 60% of all goods and 30% of all services are used in the production of goods. What proportion of the total economic output is used in the production of goods?

13. In the notation of the open model of Leontief, suppose that

$$A = \begin{pmatrix} \frac{1}{2} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{5} \end{pmatrix} \quad \text{and} \quad d = \begin{pmatrix} 2 \\ 5 \end{pmatrix}$$

   are the input–output matrix and the demand vector, respectively. How much of each commodity must be produced to satisfy this demand?

14. A certain economy consisting of the two sectors of goods and services supports a defense system that consumes \$90 billion worth of goods and \$20 billion worth of services from the economy but does not contribute to economic production. Suppose that 50 cents worth of goods and 20 cents worth of services are required to produce \$1 worth of goods and that 30 cents worth of goods and 60 cents worth of services are required to produce \$1 worth of services. What must the total output of the economic system be to support this defense system?

3.4 SYSTEMS OF LINEAR EQUATIONS— COMPUTATIONAL ASPECTS

In Section 3.3, we obtained a necessary and sufficient condition for a system of linear equations to have solutions (Theorem 3.11 p. 174) and learned how to express the solutions to a nonhomogeneous system in terms of solutions to the corresponding homogeneous system (Theorem 3.9 p. 172). The latter result enables us to determine all the solutions to a given system if we can find one solution to the given system and a basis for the solution set of the corresponding homogeneous system. In this section, we use elementary row operations to accomplish these two objectives simultaneously. The essence of this technique is to transform a given system of linear equations into a system having the same solutions, but which is easier to solve (as in Section 1.4).

Definition. *Two systems of linear equations are called* **equivalent** *if they have the same solution set.*

The following theorem and corollary give a useful method for obtaining equivalent systems.

Theorem 3.13. *Let $Ax = b$ be a system of m linear equations in n unknowns, and let C be an invertible $m \times m$ matrix. Then the system $(CA)x = Cb$ is equivalent to $Ax = b$.*

Proof. Let K be the solution set for $Ax = b$ and K' the solution set for $(CA)x = Cb$. If $w \in K$, then $Aw = b$. So $(CA)w = Cb$, and hence $w \in K'$. Thus $K \subseteq K'$.

Conversely, if $w \in K'$, then $(CA)w = Cb$. Hence

$$Aw = C^{-1}(CAw) = C^{-1}(Cb) = b;$$

so $w \in K$. Thus $K' \subseteq K$, and therefore, $K = K'$. ∎

Corollary. *Let $Ax = b$ be a system of m linear equations in n unknowns. If $(A'|b')$ is obtained from $(A|b)$ by a finite number of elementary row operations, then the system $A'x = b'$ is equivalent to the original system.*

Proof. Suppose that $(A'|b')$ is obtained from $(A|b)$ by elementary row operations. These may be executed by multiplying $(A|b)$ by elementary $m \times m$ matrices $E_1, E_2, \ldots, E_p$. Let $C = E_p \cdots E_2 E_1$; then

$$(A'|b') = C(A|b) = (CA|Cb).$$

Since each E_i is invertible, so is C. Now $A' = CA$ and $b' = Cb$. Thus by Theorem 3.13, the system $A'x = b'$ is equivalent to the system $Ax = b$. ∎

We now describe a method for solving any system of linear equations. Consider, for example, the system of linear equations

$$\begin{aligned} 3x_1 + 2x_2 + 3x_3 - 2x_4 &= 1 \\ x_1 + x_2 + x_3 &= 3 \\ x_1 + 2x_2 + x_3 - x_4 &= 2. \end{aligned}$$

First, we form the augmented matrix

$$\left(\begin{array}{rrrr|r} 3 & 2 & 3 & -2 & 1 \\ 1 & 1 & 1 & 0 & 3 \\ 1 & 2 & 1 & -1 & 2 \end{array}\right).$$

By using elementary row operations, we transform the augmented matrix into an upper triangular matrix in which the first nonzero entry of each row is 1, and it occurs in a column to the right of the first nonzero entry of each preceding row. (Recall that matrix A is upper triangular if $A_{ij} = 0$ whenever $i > j$.)

1. *In the leftmost nonzero column, create a 1 in the first row.* In our example, we can accomplish this step by interchanging the first and third rows. The resulting matrix is

$$\left(\begin{array}{rrrr|r} 1 & 2 & 1 & -1 & 2 \\ 1 & 1 & 1 & 0 & 3 \\ 3 & 2 & 3 & -2 & 1 \end{array}\right).$$

2. *By means of type 3 row operations, use the first row to obtain zeros in the remaining positions of the leftmost nonzero column.* In our example, we must add -1 times the first row to the second row and then add -3 times the first row to the third row to obtain

$$\left(\begin{array}{rrrr|r} 1 & 2 & 1 & -1 & 2 \\ 0 & -1 & 0 & 1 & 1 \\ 0 & -4 & 0 & 1 & -5 \end{array}\right).$$

3. *Create a 1 in the next row in the leftmost possible column, without using previous row(s).* In our example, the second column is the leftmost

possible column, and we can obtain a 1 in the second row, second column by multiplying the second row by -1. This operation produces

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & -4 & 0 & 1 & -5 \end{array}\right).$$

4. *Now use type 3 elementary row operations to obtain zeros below the 1 created in the preceding step.* In our example, we must add four times the second row to the third row. The resulting matrix is

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & -3 & -9 \end{array}\right).$$

5. *Repeat steps 3 and 4 on each succeeding row until no nonzero rows remain.* (This creates zeros above the first nonzero entry in each row.) In our example, this can be accomplished by multiplying the third row by $-\frac{1}{3}$. This operation produces

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right).$$

We have now obtained the desired matrix. To complete the simplification of the augmented matrix, we must make the first nonzero entry in each row the only nonzero entry in its column. (This corresponds to eliminating certain unknowns from all but one of the equations.)

6. *Work upward, beginning with the last nonzero row, and add multiples of each row to the rows above.* (This creates zeros above the first nonzero entry in each row.) In our example, the third row is the last nonzero row, and the first nonzero entry of this row lies in column 4. Hence we add the third row to the first and second rows to obtain zeros in row 1, column 4 and row 2, column 4. The resulting matrix is

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 5 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right).$$

7. *Repeat the process described in step 6 for each preceding row until it is performed with the second row, at which time the reduction process is complete.* In our example, we must add -2 times the second row to the first row in order to make the first row, second column entry become zero. This operation produces

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right).$$

We have now obtained the desired reduction of the augmented matrix. This matrix corresponds to the system of linear equations

$$\begin{aligned} x_1 + \quad x_3 \quad &= 1 \\ x_2 \qquad\quad &= 2 \\ x_4 &= 3. \end{aligned}$$

Recall that, by the corollary to Theorem 3.13, this system is equivalent to the original system. But this system is easily solved. Obviously $x_2 = 2$ and $x_4 = 3$. Moreover, x_1 and x_3 can have any values provided their sum is 1. Letting $x_3 = t$, we then have $x_1 = 1 - t$. Thus an arbitrary solution to the original system has the form

$$\begin{pmatrix} 1-t \\ 2 \\ t \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 3 \end{pmatrix} + t \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Observe that

$$\left\{ \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$

is a basis for the homogeneous system of equations corresponding to the given system.

In the preceding example we performed elementary row operations on the augmented matrix of the system until we obtained the augmented matrix of a system having properties 1, 2, and 3 on page 27. Such a matrix has a special name.

Definition. *A matrix is said to be in* **reduced row echelon form** *if the following three conditions are satisfied.*

(a) *Any row containing a nonzero entry precedes any row in which all the entries are zero (if any).*

(b) *The first nonzero entry in each row is the only nonzero entry in its column.*

(c) *The first nonzero entry in each row is 1 and it occurs in a column to the right of the first nonzero entry in the preceding row.*

Example 1

(a) The matrix on page 184 is in reduced row echelon form. Note that the first nonzero entry of each row is 1 and that the column containing each such entry has all zeros otherwise. Also note that each time we move downward to a new row, we must move to the right one or more columns to find the first nonzero entry of the new row.

(b) The matrix

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

is *not* in reduced row echelon form, because the first column, which contains the first nonzero entry in row 1, contains another nonzero entry. Similarly, the matrix

$$\begin{pmatrix} 0 & 1 & 0 & 2 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

is not in reduced row echelon form, because the first nonzero entry of the second row is not to the right of the first nonzero entry of the first row. Finally, the matrix

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

is not in reduced row echelon form, because the first nonzero entry of the first row is not 1. ♦

It can be shown (see the corollary to Theorem 3.16) that the reduced row echelon form of a matrix is unique; that is, if different sequences of elementary row operations are used to transform a matrix into matrices Q and Q' in reduced row echelon form, then $Q = Q'$. Thus, although there are many different sequences of elementary row operations that can be used to transform a given matrix into reduced row echelon form, they all produce the same result.

The procedure described on pages 183–185 for reducing an augmented matrix to reduced row echelon form is called **Gaussian elimination**. It consists of two separate parts.

1. In the *forward pass* (steps 1-5), the augmented matrix is transformed into an upper triangular matrix in which the first nonzero entry of each row is 1, and it occurs in a column to the right of the first nonzero entry of each preceding row.
2. In the *backward pass* or *back-substitution* (steps 6-7), the upper triangular matrix is transformed into reduced row echelon form by making the first nonzero entry of each row the only nonzero entry of its column.

Of all the methods for transforming a matrix into its reduced row echelon form, Gaussian elimination requires the fewest arithmetic operations. (For large matrices, it requires approximately 50% fewer operations than the Gauss-Jordan method, in which the matrix is transformed into reduced row echelon form by using the first nonzero entry in each row to make zero all other entries in its column.) Because of this efficiency, Gaussian elimination is the preferred method when solving systems of linear equations on a computer. In this context, the Gaussian elimination procedure is usually modified in order to minimize roundoff errors. Since discussion of these techniques is inappropriate here, readers who are interested in such matters are referred to books on numerical analysis.

When a matrix is in reduced row echelon form, the corresponding system of linear equations is easy to solve. We present below a procedure for solving any system of linear equations for which the augmented matrix is in reduced row echelon form. First, however, we note that every matrix can be transformed into reduced row echelon form by Gaussian elimination. In the forward pass, we satisfy conditions (a) and (c) in the definition of reduced row echelon form and thereby make zero all entries below the first nonzero entry in each row. Then in the backward pass, we make zero all entries above the first nonzero entry in each row, thereby satisfying condition (b) in the definition of reduced row echelon form.

Theorem 3.14. *Gaussian elimination transforms any matrix into its reduced row echelon form.*

We now describe a method for solving a system in which the augmented matrix is in reduced row echelon form. To illustrate this procedure, we consider the system

$$\begin{aligned}
2x_1 + 3x_2 + x_3 + 4x_4 - 9x_5 &= 17\\
x_1 + x_2 + x_3 + x_4 - 3x_5 &= 6\\
x_1 + x_2 + x_3 + 2x_4 - 5x_5 &= 8\\
2x_1 + 2x_2 + 2x_3 + 3x_4 - 8x_5 &= 14,
\end{aligned}$$

for which the augmented matrix is

$$\left(\begin{array}{ccccc|c} 2 & 3 & 1 & 4 & -9 & 17\\ 1 & 1 & 1 & 1 & -3 & 6\\ 1 & 1 & 1 & 2 & -5 & 8\\ 2 & 2 & 2 & 3 & -8 & 14 \end{array}\right).$$

Applying Gaussian elimination to the augmented matrix of the system produces the following sequence of matrices.

$$\left(\begin{array}{ccccc|c} 2 & 3 & 1 & 4 & -9 & 17\\ 1 & 1 & 1 & 1 & -3 & 6\\ 1 & 1 & 1 & 2 & -5 & 8\\ 2 & 2 & 2 & 3 & -8 & 14 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & -3 & 6\\ 2 & 3 & 1 & 4 & -9 & 17\\ 1 & 1 & 1 & 2 & -5 & 8\\ 2 & 2 & 2 & 3 & -8 & 14 \end{array}\right) \longrightarrow$$

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & -3 & 6 \\ 0 & 1 & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 1 & -2 & 2 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & -3 & 6 \\ 0 & 1 & -1 & 2 & -3 & 5 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \longrightarrow$$

$$\left(\begin{array}{ccccc|c} 1 & 1 & 1 & 0 & -1 & 4 \\ 0 & 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{ccccc|c} 1 & 0 & 2 & 0 & -2 & 3 \\ 0 & 1 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

The system of linear equations corresponding to this last matrix is

$$\begin{aligned} x_1 \quad\quad + 2x_3 \quad\quad - 2x_5 &= 3 \\ x_2 - x_3 \quad\quad + x_5 &= 1 \\ x_4 - 2x_5 &= 2. \end{aligned}$$

Notice that we have ignored the last row since it consists entirely of zeros.

To solve a system for which the augmented matrix is in reduced row echelon form, divide the variables into two sets. The first set consists of those variables that appear as leftmost variables in one of the equations of the system (in this case the set is $\{x_1, x_2, x_4\}$). The second set consists of all the remaining variables (in this case, $\{x_3, x_5\}$). To each variable in the second set, assign a parametric value $t_1, t_2, \ldots$ ($x_3 = t_1$, $x_5 = t_2$), and then solve for the variables of the first set in terms of those in the second set:

$$\begin{aligned} x_1 &= -2x_3 + 2x_5 + 3 = -2t_1 + 2t_2 + 3 \\ x_2 &= x_3 - x_5 + 1 = t_1 - t_2 + 1 \\ x_4 &= \phantom{-2x_3 + {}} 2x_5 + 2 = \phantom{-2t_1 + {}} 2t_2 + 2. \end{aligned}$$

Thus an arbitrary solution is of the form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2t_1 + 2t_2 + 3 \\ t_1 - t_2 + 1 \\ t_1 \\ 2t_2 + 2 \\ t_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} + t_1 \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{pmatrix},$$

where $t_1, t_2 \in R$. Notice that

$$\left\{ \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \\ 0 \\ 2 \\ 1 \end{pmatrix} \right\}$$

is a basis for the solution set of the corresponding homogeneous system of equations and

$$\begin{pmatrix} 3 \\ 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}$$

is a particular solution to the original system.

Therefore, in simplifying the augmented matrix of the system to reduced row echelon form, we are in effect simultaneously finding a particular solution to the original system and a basis for the solution set of the associated homogeneous system. Moreover, this procedure detects when a system is inconsistent, for by Exercise 3, solutions exist if and only if, in the reduction of the augmented matrix to reduced row echelon form, we do not obtain a row in which the only nonzero entry lies in the last column.

Thus to use this procedure for solving a system $Ax = b$ of m linear equations in n unknowns, we need only begin to transform the augmented matrix $(A|b)$ into its reduced row echelon form $(A'|b')$ by means of Gaussian elimination. If a row is obtained in which the only nonzero entry lies in the last column, then the original system is inconsistent. Otherwise, discard any zero rows from $(A'|b')$, and write the corresponding system of equations. Solve this system as described above to obtain an arbitrary solution of the form

$$s = s_0 + t_1u_1 + t_2u_2 + \cdots + t_{n-r}u_{n-r},$$

where r is the number of nonzero rows in A' $(r \leq m)$. The preceding equation is called a **general solution** of the system $Ax = b$. It expresses an arbitrary solution s of $Ax = b$ in terms of $n - r$ parameters. The following theorem states that s cannot be expressed in fewer than $n - r$ parameters.

Theorem 3.15. *Let $Ax = b$ be a system of r nonzero equations in n unknowns. Suppose that* $\operatorname{rank}(A) = \operatorname{rank}(A|b)$ *and that $(A|b)$ is in reduced row echelon form. Then*

(a) $\operatorname{rank}(A) = r$.

(b) *If the general solution obtained by the procedure above is of the form*

$$s = s_0 + t_1u_1 + t_2u_2 + \cdots + t_{n-r}u_{n-r},$$

then $\{u_1, u_2, \ldots, u_{n-r}\}$ is a basis for the solution set of the corresponding homogeneous system, and s_0 is a solution to the original system.

Proof. Since $(A|b)$ is in reduced row echelon form, $(A|b)$ must have r nonzero rows. Clearly these rows are linearly independent by the definition of the reduced row echelon form, and so $\operatorname{rank}(A|b) = r$. Thus $\operatorname{rank}(A) = r$.

Let K be the solution set for $Ax = b$, and let K_H be the solution set for $Ax = 0$. Setting $t_1 = t_2 = \cdots = t_{n-r} = 0$, we see that $s = s_0 \in K$. But by Theorem 3.9 (p. 172), $K = \{s_0\} + \mathsf{K}_\mathsf{H}$. Hence

$$\mathsf{K}_\mathsf{H} = \{-s_0\} + K = \mathrm{span}(\{u_1, u_2, \ldots, u_{n-r}\}).$$

Because $\mathrm{rank}(A) = r$, we have $\dim(\mathsf{K}_\mathsf{H}) = n - r$. Thus since $\dim(\mathsf{K}_\mathsf{H}) = n - r$ and K_H is generated by a set $\{u_1, u_2, \ldots, u_{n-r}\}$ containing at most $n - r$ vectors, we conclude that this set is a basis for K_H. ∎

An Interpretation of the Reduced Row Echelon Form

Let A be an $m \times n$ matrix with columns $a_1, a_2, \ldots, a_n$, and let B be the reduced row echelon form of A. Denote the columns of B by $b_1, b_2, \ldots, b_n$. If the rank of A is r, then the rank of B is also r by the corollary to Theorem 3.4 (p. 153). Because B is in reduced row echelon form, no nonzero row of B can be a linear combination of the other rows of B. Hence B must have exactly r nonzero rows, and if $r \geq 1$, the vectors $e_1, e_2, \ldots, e_r$ must occur among the columns of B. For $i = 1, 2, \ldots, r$, let j_i denote a column number of B such that $b_{j_i} = e_i$. We claim that $a_{j_1}, a_{j_2}, \ldots, a_{j_r}$, the columns of A corresponding to these columns of B, are linearly independent. For suppose that there are scalars $c_1, c_2, \ldots, c_r$ such that

$$c_1 a_{j_1} + c_2 a_{j_2} + \cdots + c_r a_{j_r} = 0.$$

Because B can be obtained from A by a sequence of elementary row operations, there exists (as in the proof of the corollary to Theorem 3.13) an invertible $m \times m$ matrix M such that $MA = B$. Multiplying the preceding equation by M yields

$$c_1 M a_{j_1} + c_2 M a_{j_2} + \cdots + c_r M a_{j_r} = 0.$$

Since $M a_{j_i} = b_{j_i} = e_i$, it follows that

$$c_1 e_1 + c_2 e_2 + \cdots + c_r e_r = 0.$$

Hence $c_1 = c_2 = \cdots = c_r = 0$, proving that the vectors $a_{j_1}, a_{j_2}, \ldots, a_{j_r}$ are linearly independent.

Because B has only r nonzero rows, every column of B has the form

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

for scalars $d_1, d_2, \ldots, d_r$. The corresponding column of A must be

$$\begin{aligned} M^{-1}(d_1e_1 + d_2e_2 + \cdots + d_re_r) &= d_1M^{-1}e_1 + d_2M^{-1}e_2 + \cdots + d_rM^{-1}e_r \\ &= d_1M^{-1}b_{j_1} + d_2M^{-1}b_{j_2} + \cdots + d_rM^{-1}b_{j_r} \\ &= d_1a_{j_1} + d_2a_{j_2} + \cdots + d_ra_{j_r}. \end{aligned}$$

The next theorem summarizes these results.

Theorem 3.16. *Let A be an $m \times n$ matrix of rank r, where $r > 0$, and let B be the reduced row echelon form of A. Then*

(a) *The number of nonzero rows in B is r.*
(b) *For each $i = 1, 2, \ldots, r$, there is a column b_{j_i} of B such that $b_{j_i} = e_i$.*
(c) *The columns of A numbered $j_1, j_2, \ldots, j_r$ are linearly independent.*
(d) *For each $k = 1, 2, \ldots n$, if column k of B is $d_1e_1 + d_2e_2 + \cdots + d_re_r$, then column k of A is $d_1a_{j_1} + d_2a_{j_2} + \cdots + d_ra_{j_r}$.*

Corollary. *The reduced row echelon form of a matrix is unique.*

Proof. Exercise. (See Exercise15.) ∎

Example 2

Let

$$A = \begin{pmatrix} 2 & 4 & 6 & 2 & 4 \\ 1 & 2 & 3 & 1 & 1 \\ 2 & 4 & 8 & 0 & 0 \\ 3 & 6 & 7 & 5 & 9 \end{pmatrix}.$$

The reduced row echelon form of A is

$$B = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since B has three nonzero rows, the rank of A is 3. The first, third, and fifth columns of B are e_1, e_2, and e_3; so Theorem 3.16(c) asserts that the first, third, and fifth columns of A are linearly independent.

Let the columns of A be denoted a_1, a_2, a_3, a_4, and a_5. Because the second column of B is $2e_1$, it follows from Theorem 3.16(d) that $a_2 = 2a_1$, as is easily checked. Moreover, since the fourth column of B is $4e_1 + (-1)e_2$, the same result shows that

$$a_4 = 4a_1 + (-1)a_3. \quad ♦$$

In Example 6 of Section 1.6, we extracted a basis for R^3 from the generating set

$$S = \{(2,-3,5),(8,-12,20),(1,0,-2),(0,2,-1),(7,2,0)\}.$$

The procedure described there can be streamlined by using Theorem 3.16. We begin by noting that if S were linearly independent, then S would be a basis for R^3. In this case, it is clear that S is linearly dependent because S contains more than $\dim(\mathsf{R}^3) = 3$ vectors. Nevertheless, it is instructive to consider the calculation that is needed to determine whether S is linearly dependent or linearly independent. Recall that S is linearly dependent if there are scalars c_1, c_2, c_3, c_4, and c_5, not all zero, such that

$$c_1(2,-3,5)+c_2(8,-12,20)+c_3(1,0,-2)+c_4(0,2,-1)+c_5(7,2,0) = (0,0,0).$$

Thus S is linearly dependent if and only if the system of linear equations

$$\begin{aligned} 2c_1 + 8c_2 + c_3 \phantom{{}+2c_4} + 7c_5 &= 0\\ -3c_1 - 12c_2 \phantom{{}+c_3} + 2c_4 + 2c_5 &= 0\\ 5c_1 + 20c_2 - 2c_3 - c_4 \phantom{{}+2c_5} &= 0 \end{aligned}$$

has a nonzero solution. The augmented matrix of this system of equations is

$$A = \begin{pmatrix} 2 & 8 & 1 & 0 & 7 & 0\\ -3 & -12 & 0 & 2 & 2 & 0\\ 5 & 20 & -2 & -1 & 0 & 0 \end{pmatrix},$$

and its reduced row echelon form is

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

Using the technique described earlier in this section, we can find nonzero solutions of the preceding system, confirming that S is linearly dependent. However, Theorem 3.16(c) gives us additional information. Since the first, third, and fourth columns of B are e_1, e_2, and e_3, we conclude that the first, third, and fourth columns of A are linearly independent. But the columns of A other than the last column (which is the zero vector) are vectors in S. Hence

$$\beta = \{(2,-3,5),(1,0,-2),(0,2,-1)\}$$

is a linearly independent subset of S. If follows from (b) of Corollary 2 to the replacement theorem (p. 47) that β is a basis for R^3.

Because every finite-dimensional vector space over F is isomorphic to F^n for some n, a similar approach can be used to reduce any finite generating set to a basis. This technique is illustrated in the next example.

Example 3

The set

$$S=\{2+x+2x^2+3x^3, 4+2x+4x^2+6x^3, 6+3x+8x^2+7x^3, 2+x+5x^3, 4+x+9x^3\}$$

generates a subspace V of $\mathsf{P}_3(R)$. To find a subset of S that is a basis for V, we consider the subset

$$S' = \{(2,1,2,3), (4,2,4,6), (6,3,8,7), (2,1,0,5), (4,1,0,9)\}$$

consisting of the images of the polynomials in S under the standard representation of $\mathsf{P}_3(R)$ with respect to the standard ordered basis. Note that the 4×5 matrix in which the columns are the vectors in S' is the matrix A in Example 2. From the reduced row echelon form of A, which is the matrix B in Example 2, we see that the first, third, and fifth columns of A are linearly independent and the second and fourth columns of A are linear combinations of the first, third, and fifth columns. Hence

$$\{(2,1,2,3), (6,3,8,7), (4,1,0,9)\}$$

is a basis for the subspace of R^4 that is generated by S'. It follows that

$$\{2+x+2x^2+3x^3, 6+3x+8x^2+7x^3, 4+x+9x^3\}$$

is a basis for the subspace V of $\mathsf{P}_3(R)$. ♦

We conclude this section by describing a method for extending a linearly independent subset S of a finite-dimensional vector space V to a basis for V. Recall that this is always possible by (c) of Corollary 2 to the replacement theorem (p. 47). Our approach is based on the replacement theorem and assumes that we can find an explicit basis β for V. Let S' be the ordered set consisting of the vectors in S followed by those in β. Since $\beta \subseteq S'$, the set S' generates V. We can then apply the technique described above to reduce this generating set to a basis for V containing S.

Example 4

Let

$$V = \{(x_1, x_2, x_3, x_4, x_5) \in \mathsf{R}^5 \colon x_1 + 7x_2 + 5x_3 - 4x_4 + 2x_5 = 0\}.$$

It is easily verified that V is a subspace of R^5 and that

$$S = \{(-2,0,0,-1,-1), (1,1,-2,-1,-1), (-5,1,0,1,1)\}$$

is a linearly independent subset of V.

To extend S to a basis for V, we first obtain a basis β for V. To do so, we solve the system of linear equations that defines V. Since in this case V is defined by a single equation, we need only write the equation as

$$x_1 = -7x_2 - 5x_3 + 4x_4 - 2x_5$$

and assign parametric values to x_2, x_3, x_4, and x_5. If $x_2 = t_1$, $x_3 = t_2$, $x_4 = t_3$, and $x_5 = t_4$, then the vectors in V have the form

$$\begin{aligned}(x_1,x_2,x_3,x_4,x_5) &= (-7t_1 - 5t_2 + 4t_3 - 2t_4, t_1, t_2, t_3, t_4)\\ &= t_1(-7,1,0,0,0) + t_2(-5,0,1,0,0) + t_3(4,0,0,1,0) + t_4(-2,0,0,0,1).\end{aligned}$$

Hence

$$\beta = \{(-7,1,0,0,0), (-5,0,1,0,0), (4,0,0,1,0), (-2,0,0,0,1)\}$$

is a basis for V by Theorem 3.15.

The matrix whose columns consist of the vectors in S followed by those in β is

$$\begin{pmatrix} -2 & 1 & -5 & -7 & -5 & 4 & -2 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 & 1 & 0 \\ -1 & -1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix},$$

and its reduced row echelon form is

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & 1 & 0 & 0 & -.5 & 0 & 0 \\ 0 & 0 & 1 & 1 & .5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus

$$\{(-2,0,0,-1,-1), (1,1,-2,-1,-1), (-5,1,0,1,1), (4,0,0,1,0)\}$$

is a basis for V containing S. ◆

EXERCISES

1. Label the following statements as true or false.

(a) If $(A'|b')$ is obtained from $(A|b)$ by a finite sequence of elementary column operations, then the systems $Ax = b$ and $A'x = b'$ are equivalent.

(b) If $(A'|b')$ is obtained from $(A|b)$ by a finite sequence of elementary row operations, then the systems $Ax = b$ and $A'x = b'$ are equivalent.

(c) If A is an $n \times n$ matrix with rank n, then the reduced row echelon form of A is I_n.

(d) Any matrix can be put in reduced row echelon form by means of a finite sequence of elementary row operations.

(e) If $(A|b)$ is in reduced row echelon form, then the system $Ax = b$ is consistent.

(f) Let $Ax = b$ be a system of m linear equations in n unknowns for which the augmented matrix is in reduced row echelon form. If this system is consistent, then the dimension of the solution set of $Ax = 0$ is $n - r$, where r equals the number of nonzero rows in A.

(g) If a matrix A is transformed by elementary row operations into a matrix A' in reduced row echelon form, then the number of nonzero rows in A' equals the rank of A.

2. Use Gaussian elimination to solve the following systems of linear equations.

(a) $$\begin{aligned} x_1 + 2x_2 - x_3 &= -1 \\ 2x_1 + 2x_2 + x_3 &= 1 \\ 3x_1 + 5x_2 - 2x_3 &= -1 \end{aligned}$$

(b) $$\begin{aligned} x_1 - 2x_2 - x_3 &= 1 \\ 2x_1 - 3x_2 + x_3 &= 6 \\ 3x_1 - 5x_2 &= 7 \\ x_1 + 5x_3 &= 9 \end{aligned}$$

(c) $$\begin{aligned} x_1 + 2x_2 + 2x_4 &= 6 \\ 3x_1 + 5x_2 - x_3 + 6x_4 &= 17 \\ 2x_1 + 4x_2 + x_3 + 2x_4 &= 12 \\ 2x_1 - 7x_3 + 11x_4 &= 7 \end{aligned}$$

(d) $$\begin{aligned} x_1 - x_2 - 2x_3 + 3x_4 &= -7 \\ 2x_1 - x_2 + 6x_3 + 6x_4 &= -2 \\ -2x_1 + x_2 - 4x_3 - 3x_4 &= 0 \\ 3x_1 - 2x_2 + 9x_3 + 10x_4 &= -5 \end{aligned}$$

(e) $$\begin{aligned} x_1 - 4x_2 - x_3 + x_4 &= 3 \\ 2x_1 - 8x_2 + x_3 - 4x_4 &= 9 \\ -x_1 + 4x_2 - 2x_3 + 5x_4 &= -6 \end{aligned}$$

(f) $$\begin{aligned} x_1 + 2x_2 - x_3 + 3x_4 &= 2 \\ 2x_1 + 4x_2 - x_3 + 6x_4 &= 5 \\ x_2 + 2x_4 &= 3 \end{aligned}$$

(g) $$\begin{aligned} 2x_1 - 2x_2 - x_3 + 6x_4 - 2x_5 &= 1 \\ x_1 - x_2 + x_3 + 2x_4 - x_5 &= 2 \\ 4x_1 - 4x_2 + 5x_3 + 7x_4 - x_5 &= 6 \end{aligned}$$

(h) $$\begin{aligned} 3x_1 - x_2 + x_3 - x_4 + 2x_5 &= 5 \\ x_1 - x_2 - x_3 - 2x_4 - x_5 &= 2 \\ 5x_1 - 2x_2 + x_3 - 3x_4 + 3x_5 &= 10 \\ 2x_1 - x_2 - 2x_4 + x_5 &= 5 \end{aligned}$$

(i) $$\begin{aligned} 3x_1 - x_2 + 2x_3 + 4x_4 + x_5 &= 2 \\ x_1 - x_2 + 2x_3 + 3x_4 + x_5 &= -1 \\ 2x_1 - 3x_2 + 6x_3 + 9x_4 + 4x_5 &= -5 \\ 7x_1 - 2x_2 + 4x_3 + 8x_4 + x_5 &= 6 \end{aligned}$$

(j) $$\begin{aligned} 2x_1 \qquad + 3x_3 \qquad - 4x_5 &= 5 \\ 3x_1 - 4x_2 + 8x_3 + 3x_4 \qquad &= 8 \\ x_1 - x_2 + 2x_3 + x_4 - x_5 &= 2 \\ -2x_1 + 5x_2 - 9x_3 - 3x_4 - 5x_5 &= -8 \end{aligned}$$

3. Suppose that the augmented matrix of a system $Ax = b$ is transformed into a matrix $(A'|b')$ in reduced row echelon form by a finite sequence of elementary row operations.
   (a) Prove that $\text{rank}(A') \neq \text{rank}(A'|b')$ if and only if $(A'|b')$ contains a row in which the only nonzero entry lies in the last column.
   (b) Deduce that $Ax = b$ is consistent if and only if $(A'|b')$ contains no row in which the only nonzero entry lies in the last column.

4. For each of the systems that follow, apply Exercise 3 to determine whether the system is consistent. If the system is consistent, find all solutions. Finally, find a basis for the solution set of the corresponding homogeneous system.

(a) $$\begin{aligned} x_1 + 2x_2 - x_3 + x_4 &= 2 \\ 2x_1 + x_2 + x_3 - x_4 &= 3 \\ x_1 + 2x_2 - 3x_3 + 2x_4 &= 2 \end{aligned}$$

(b) $$\begin{aligned} x_1 + x_2 - 3x_3 + x_4 &= -2 \\ x_1 + x_2 + x_3 - x_4 &= 2 \\ x_1 + x_2 - x_3 \qquad &= 0 \end{aligned}$$

(c) $$\begin{aligned} x_1 + x_2 - 3x_3 + x_4 &= 1 \\ x_1 + x_2 + x_3 - x_4 &= 2 \\ x_1 + x_2 - x_3 \qquad &= 0 \end{aligned}$$

5. Let the reduced row echelon form of A be

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

Determine A if the first, second, and fourth columns of A are

$$\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix},$$

respectively.

6. Let the reduced row echelon form of A be

$$\begin{pmatrix} 1 & -3 & 0 & 4 & 0 & 5 \\ 0 & 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Determine A if the first, third, and sixth columns of A are

$$\begin{pmatrix} 1 \\ -2 \\ -1 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ 2 \\ -4 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 3 \\ -9 \\ 2 \\ 5 \end{pmatrix},$$

respectively.

7. It can be shown that the vectors $u_1 = (2, -3, 1)$, $u_2 = (1, 4, -2)$, $u_3 = (-8, 12, -4)$, $u_4 = (1, 37, -17)$, and $u_5 = (-3, -5, 8)$ generate R^3. Find a subset of $\{u_1, u_2, u_3, u_4, u_5\}$ that is a basis for R^3.

8. Let W denote the subspace of R^5 consisting of all vectors having coordinates that sum to zero. The vectors

$$\begin{aligned} u_1 &= (2, -3, 4, -5, 2), & u_2 &= (-6, 9, -12, 15, -6), \\ u_3 &= (3, -2, 7, -9, 1), & u_4 &= (2, -8, 2, -2, 6), \\ u_5 &= (-1, 1, 2, 1, -3), & u_6 &= (0, -3, -18, 9, 12), \\ u_7 &= (1, 0, -2, 3, -2), \quad \text{and} & u_8 &= (2, -1, 1, -9, 7) \end{aligned}$$

generate W. Find a subset of $\{u_1, u_2, \ldots, u_8\}$ that is a basis for W.

9. Let W be the subspace of $\mathsf{M}_{2\times 2}(R)$ consisting of the symmetric 2×2 matrices. The set

$$S = \left\{ \begin{pmatrix} 0 & -1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 9 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ -2 & 4 \end{pmatrix}, \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix} \right\}$$

generates W. Find a subset of S that is a basis for W.

10. Let

$$\mathsf{V} = \{(x_1, x_2, x_3, x_4, x_5) \in \mathsf{R}^5 \colon x_1 - 2x_2 + 3x_3 - x_4 + 2x_5 = 0\}.$$

(a) Show that $S = \{(0, 1, 1, 1, 0)\}$ is a linearly independent subset of V.
(b) Extend S to a basis for V.

11. Let V be as in Exercise 10.

(a) Show that $S = \{(1, 2, 1, 0, 0)\}$ is a linearly independent subset of V.
(b) Extend S to a basis for V.

12. Let V denote the set of all solutions to the system of linear equations

$$\begin{aligned} x_1 - x_2 \phantom{{}- x_3} + 2x_4 - 3x_5 + x_6 &= 0 \\ 2x_1 - x_2 - x_3 + 3x_4 - 4x_5 + 4x_6 &= 0. \end{aligned}$$

(a) Show that $S = \{(0,-1,0,1,1,0),(1,0,1,1,1,0)\}$ is a linearly independent subset of V.
(b) Extend S to a basis for V.

13. Let V be as in Exercise 12.
(a) Show that $S = \{(1,0,1,1,1,0),(0,2,1,1,0,0)\}$ is a linearly independent subset of V.
(b) Extend S to a basis for V.

14. If $(A|b)$ is in reduced row echelon form, prove that A is also in reduced row echelon form.

15. Prove the corollary to Theorem 3.16: The reduced row echelon form of a matrix is unique.

INDEX OF DEFINITIONS FOR CHAPTER 3