b. Show that B is not invertible.

[*Hint*: Column 3 = 3(column 2) − column 1.]

Exercise 2.4.36 Show that a square matrix A is invertible if and only if it can be left-cancelled: $AB = AC$ implies $B = C$.

Exercise 2.4.37 If $U^2 = I$, show that $I + U$ is not invertible unless $U = I$.

Exercise 2.4.38

a. If J is the 4×4 matrix with every entry 1, show that $I - \frac{1}{2}J$ is self-inverse and symmetric.

b. If X is $n \times m$ and satisfies $X^T X = I_m$, show that $I_n - 2XX^T$ is self-inverse and symmetric.

Exercise 2.4.39 An $n \times n$ matrix P is called an idempotent if $P^2 = P$. Show that:

a. I is the only invertible idempotent.

b. P is an idempotent if and only if $I - 2P$ is self-inverse.

c. U is self-inverse if and only if $U = I - 2P$ for some idempotent P.

d. $I - aP$ is invertible for any $a \neq 1$, and that $(I - aP)^{-1} = I + \left(\frac{a}{1-a}\right)^P$.

Exercise 2.4.40 If $A^2 = kA$, where $k \neq 0$, show that A is invertible if and only if $A = kI$.

Exercise 2.4.41 Let A and B denote $n \times n$ invertible matrices.

a. Show that $A^{-1} + B^{-1} = A^{-1}(A + B)B^{-1}$.

b. If $A + B$ is also invertible, show that $A^{-1} + B^{-1}$ is invertible and find a formula for $(A^{-1} + B^{-1})^{-1}$.

Exercise 2.4.42 Let A and B be $n \times n$ matrices, and let I be the $n \times n$ identity matrix.

a. Verify that $A(I + BA) = (I + AB)A$ and that $(I + BA)B = B(I + AB)$.

b. If $I + AB$ is invertible, verify that $I + BA$ is also invertible and that $(I + BA)^{-1} = I - B(I + AB)^{-1}A$.

2.5 Elementary Matrices

It is now clear that elementary row operations are important in linear algebra: They are essential in solving linear systems (using the gaussian algorithm) and in inverting a matrix (using the matrix inversion algorithm). It turns out that they can be performed by left multiplying by certain invertible matrices. These matrices are the subject of this section.

Definition 2.12 Elementary Matrices

*An $n \times n$ matrix E is called an **elementary matrix** if it can be obtained from the identity matrix I_n by a single elementary row operation (called the operation **corresponding** to E). We say that E is of type I, II, or III if the operation is of that type (see Definition 1.2).*

Hence

$$E_1 = \left[\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right], \quad E_2 = \left[\begin{array}{cc} 1 & 0 \\ 0 & 9 \end{array} \right], \quad \text{and} \quad E_3 = \left[\begin{array}{cc} 1 & 5 \\ 0 & 1 \end{array} \right]$$

are elementary of types I, II, and III, respectively, obtained from the 2×2 identity matrix by interchanging rows 1 and 2, multiplying row 2 by 9, and adding 5 times row 2 to row 1.

Suppose now that the matrix $A = \begin{bmatrix} a & b & c \\ p & q & r \end{bmatrix}$ is left multiplied by the above elementary matrices E_1, E_2, and E_3. The results are:

$$E_1A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ p & q & r \end{bmatrix} = \begin{bmatrix} p & q & r \\ a & b & c \end{bmatrix}$$
$$E_2A = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} a & b & c \\ p & q & r \end{bmatrix} = \begin{bmatrix} a & b & c \\ 9p & 9q & 9r \end{bmatrix}$$
$$E_3A = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ p & q & r \end{bmatrix} = \begin{bmatrix} a+5p & b+5q & c+5r \\ p & q & r \end{bmatrix}$$

In each case, left multiplying A by the elementary matrix has the *same* effect as doing the corresponding row operation to A. This works in general.

Lemma 2.5.1: [10]

If an elementary row operation is performed on an $m \times n$ matrix A, the result is EA where E is the elementary matrix obtained by performing the same operation on the $m \times m$ identity matrix.

Proof. We prove it for operations of type III; the proofs for types I and II are left as exercises. Let E be the elementary matrix corresponding to the operation that adds k times row p to row $q \neq p$. The proof depends on the fact that each row of EA is equal to the corresponding row of E times A. Let $K_1, K_2, \dots, K_m$ denote the rows of I_m. Then row i of E is K_i if $i \neq q$, while row q of E is $K_q + kK_p$. Hence:

$$\begin{aligned} \text{If } i \neq q \text{ then row } i \text{ of } EA = \; & K_iA = (\text{row } i \text{ of } A). \\ \text{Row } q \text{ of } EA = (K_q + kK_p)A = \; & K_qA + k(K_pA) \\ = \; & (\text{row } q \text{ of } A) \text{ plus } k \,(\text{row } p \text{ of } A). \end{aligned}$$

Thus EA is the result of adding k times row p of A to row q, as required. □

The effect of an elementary row operation can be reversed by another such operation (called its inverse) which is also elementary of the same type (see the discussion following (Example 1.1.3). It follows that each elementary matrix E is invertible. In fact, if a row operation on I produces E, then the inverse operation carries E back to I. If F is the elementary matrix corresponding to the inverse operation, this means $FE = I$ (by Lemma 2.5.1). Thus $F = E^{-1}$ and we have proved

Lemma 2.5.2

Every elementary matrix E is invertible, and E^{-1} is also a elementary matrix (of the same type). Moreover, E^{-1} corresponds to the inverse of the row operation that produces E.

The following table gives the inverse of each type of elementary row operation:

Type	Operation	Inverse Operation
I	Interchange rows p and q	Interchange rows p and q
II	Multiply row p by $k \neq 0$	Multiply row p by $1/k$, $k \neq 0$
III	Add k times row p to row $q \neq p$	Subtract k times row p from row q, $q \neq p$

[10] A *lemma* is an auxiliary theorem used in the proof of other theorems.

Note that elementary matrices of type I are self-inverse.

Example 2.5.1

Find the inverse of each of the elementary matrices

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 9 \end{bmatrix}, \quad \text{and} \quad E_3 = \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solution. E_1, E_2, and E_3 are of type I, II, and III respectively, so the table gives

$$E_1^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = E_1, \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{9} \end{bmatrix}, \quad \text{and} \quad E_3^{-1} = \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Inverses and Elementary Matrices

Suppose that an $m \times n$ matrix A is carried to a matrix B (written $A \to B$) by a series of k elementary row operations. Let $E_1, E_2, \dots, E_k$ denote the corresponding elementary matrices. By Lemma 2.5.1, the reduction becomes

$$A \to E_1A \to E_2E_1A \to E_3E_2E_1A \to \cdots \to E_kE_{k-1}\cdots E_2E_1A = B$$

In other words,

$$A \to UA = B \quad \text{where } U = E_kE_{k-1}\cdots E_2E_1$$

The matrix $U = E_kE_{k-1}\cdots E_2E_1$ is invertible, being a product of invertible matrices by Lemma 2.5.2. Moreover, U can be computed without finding the E_i as follows: If the above series of operations carrying $A \to B$ is performed on I_m in place of A, the result is $I_m \to UI_m = U$. Hence this series of operations carries the block matrix $\begin{bmatrix} A & I_m \end{bmatrix} \to \begin{bmatrix} B & U \end{bmatrix}$. This, together with the above discussion, proves

Theorem 2.5.1

Suppose A is $m \times n$ and $A \to B$ by elementary row operations.

1. *$B = UA$ where U is an $m \times m$ invertible matrix.*
2. *U can be computed by $\begin{bmatrix} A & I_m \end{bmatrix} \to \begin{bmatrix} B & U \end{bmatrix}$ using the operations carrying $A \to B$.*
3. *$U = E_kE_{k-1}\cdots E_2E_1$ where $E_1, E_2, \dots, E_k$ are the elementary matrices corresponding (in order) to the elementary row operations carrying A to B.*

Example 2.5.2

If $A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 1 \end{bmatrix}$, express the reduced row-echelon form R of A as $R = UA$ where U is invertible.

Solution. Reduce the double matrix $\begin{bmatrix} A & I \end{bmatrix} \to \begin{bmatrix} R & U \end{bmatrix}$ as follows:

$$\begin{bmatrix} A & I \end{bmatrix} = \left[\begin{array}{ccc|cc} 2 & 3 & 1 & 1 & 0 \\ 1 & 2 & 1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|cc} 1 & 2 & 1 & 0 & 1 \\ 2 & 3 & 1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{ccc|cc} 1 & 2 & 1 & 0 & 1 \\ 0 & -1 & -1 & 1 & -2 \end{array}\right]$$
$$\to \left[\begin{array}{ccc|cc} 1 & 0 & -1 & 2 & -3 \\ 0 & 1 & 1 & -1 & 2 \end{array}\right]$$

Hence $R = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$ and $U = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$.

Now suppose that A is invertible. We know that $A \to I$ by Theorem 2.4.5, so taking $B = I$ in Theorem 2.5.1 gives $\begin{bmatrix} A & I \end{bmatrix} \to \begin{bmatrix} I & U \end{bmatrix}$ where $I = UA$. Thus $U = A^{-1}$, so we have $\begin{bmatrix} A & I \end{bmatrix} \to \begin{bmatrix} I & A^{-1} \end{bmatrix}$. This is the matrix inversion algorithm in Section 2.4. However, more is true: Theorem 2.5.1 gives $A^{-1} = U = E_k E_{k-1} \cdots E_2 E_1$ where $E_1, E_2, \dots, E_k$ are the elementary matrices corresponding (in order) to the row operations carrying $A \to I$. Hence

$$A = \left(A^{-1}\right)^{-1} = (E_k E_{k-1} \cdots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \cdots E_{k-1}^{-1} E_k^{-1} \tag{2.10}$$

By Lemma 2.5.2, this shows that every invertible matrix A is a product of elementary matrices. Since elementary matrices are invertible (again by Lemma 2.5.2), this proves the following important characterization of invertible matrices.

Theorem 2.5.2

A square matrix is invertible if and only if it is a product of elementary matrices.

It follows from Theorem 2.5.1 that $A \to B$ by row operations if and only if $B = UA$ for some invertible matrix B. In this case we say that A and B are **row-equivalent**. (See Exercise 2.5.17.)

Example 2.5.3

Express $A = \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix}$ as a product of elementary matrices.

Solution. Using Lemma 2.5.1, the reduction of $A \to I$ is as follows:

$$A = \begin{bmatrix} -2 & 3 \\ 1 & 0 \end{bmatrix} \to E_1 A = \begin{bmatrix} 1 & 0 \\ -2 & 3 \end{bmatrix} \to E_2 E_1 A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \to E_3 E_2 E_1 A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

where the corresponding elementary matrices are

$$E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

Hence $(E_3\ E_2\ E_1)A = I$, so:

$$A = (E_3E_2E_1)^{-1} = E_1^{-1}E_2^{-1}E_3^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

Smith Normal Form

Let A be an $m \times n$ matrix of rank r, and let R be the reduced row-echelon form of A. Theorem 2.5.1 shows that $R = UA$ where U is invertible, and that U can be found from $\begin{bmatrix} A & I_m \end{bmatrix} \to \begin{bmatrix} R & U \end{bmatrix}$.

The matrix R has r leading ones (since $\operatorname{rank} A = r$) so, as R is reduced, the $n \times m$ matrix R^T contains each row of I_r in the first r columns. Thus row operations will carry $R^T \to \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times m}$. Hence Theorem 2.5.1 (again) shows that $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times m} = U_1R^T$ where U_1 is an $n \times n$ invertible matrix. Writing $V = U_1^T$, we obtain

$$UAV = RV = RU_1^T = \left(U_1R^T\right)^T = \left(\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times m}\right)^T = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{m\times n}$$

Moreover, the matrix $U_1 = V^T$ can be computed by $\begin{bmatrix} R^T & I_n \end{bmatrix} \to \left[\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times m} \; V^T\right]$. This proves

Theorem 2.5.3

Let A be an $m \times n$ matrix of rank r. There exist invertible matrices U and V of size $m \times m$ and $n \times n$, respectively, such that

$$UAV = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{m\times n}$$

Moreover, if R is the reduced row-echelon form of A, then:

1. *U can be computed by $\begin{bmatrix} A & I_m \end{bmatrix} \to \begin{bmatrix} R & U \end{bmatrix}$;*
2. *V can be computed by $\begin{bmatrix} R^T & I_n \end{bmatrix} \to \left[\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}_{n\times m} \; V^T\right]$.*

If A is an $m \times n$ matrix of rank r, the matrix $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ is called the **Smith normal form**[11] of A. Whereas the reduced row-echelon form of A is the "nicest" matrix to which A can be carried by row operations, the Smith canonical form is the "nicest" matrix to which A can be carried by *row and column* operations. This is because doing row operations to R^T amounts to doing *column* operations to R and then transposing.

[11] Named after Henry John Stephen Smith (1826–83).

Example 2.5.4

Given $A = \begin{bmatrix} 1 & -1 & 1 & 2 \\ 2 & -2 & 1 & -1 \\ -1 & 1 & 0 & 3 \end{bmatrix}$, find invertible matrices U and V such that $UAV = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, where $r = \operatorname{rank} A$.

Solution. The matrix U and the reduced row-echelon form R of A are computed by the row reduction $\begin{bmatrix} A & I_3 \end{bmatrix} \to \begin{bmatrix} R & U \end{bmatrix}$:

$$\left[\begin{array}{rrrr|rrr} 1 & -1 & 1 & 2 & 1 & 0 & 0 \\ 2 & -2 & 1 & -1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 3 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrrr|rrr} 1 & -1 & 0 & -3 & -1 & 1 & 0 \\ 0 & 0 & 1 & 5 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 1 \end{array}\right]$$

Hence

$$R = \begin{bmatrix} 1 & -1 & 0 & -3 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad U = \begin{bmatrix} -1 & 1 & 0 \\ 2 & -1 & 0 \\ -1 & 1 & 1 \end{bmatrix}$$

In particular, $r = \operatorname{rank} R = 2$. Now row-reduce $\begin{bmatrix} R^T & I_4 \end{bmatrix} \to \begin{bmatrix} \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} & V^T \end{bmatrix}$:

$$\left[\begin{array}{rrr|rrrr} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ -3 & 5 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{rrr|rrrr} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & -5 & 1 \end{array}\right]$$

whence

$$V^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & -5 & -1 \end{bmatrix} \quad \text{so} \quad V = \begin{bmatrix} 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Then $UAV = \begin{bmatrix} I_2 & 0 \\ 0 & 0 \end{bmatrix}$ as is easily verified.

Uniqueness of the Reduced Row-echelon Form

In this short subsection, Theorem 2.5.1 is used to prove the following important theorem.

Theorem 2.5.4

If a matrix A is carried to reduced row-echelon matrices R and S by row operations, then $R = S$.

Proof. Observe first that $UR = S$ for some invertible matrix U (by Theorem 2.5.1 there exist invertible matrices P and Q such that $R = PA$ and $S = QA$; take $U = QP^{-1}$). We show that $R = S$ by induction on the number m of rows of R and S. The case $m = 1$ is left to the reader. If R_j and S_j denote column j in R and S respectively, the fact that $UR = S$ gives

$$UR_j = S_j \quad \text{for each } j \tag{2.11}$$

Since U is invertible, this shows that R and S have the same zero columns. Hence, by passing to the matrices obtained by deleting the zero columns from R and S, we may assume that R and S have no zero columns.

But then the first column of R and S is the first column of I_m because R and S are row-echelon, so (2.11) shows that the first column of U is column 1 of I_m. Now write U, R, and S in block form as follows.

$$U = \left[\begin{array}{cc} 1 & X \\ 0 & V \end{array} \right], \quad R = \left[\begin{array}{cc} 1 & X \\ 0 & R' \end{array} \right], \quad \text{and} \quad S = \left[\begin{array}{cc} 1 & Z \\ 0 & S' \end{array} \right]$$

Since $UR = S$, block multiplication gives $VR' = S'$ so, since V is invertible (U is invertible) and both R' and S' are reduced row-echelon, we obtain $R' = S'$ by induction. Hence R and S have the same number (say r) of leading 1s, and so both have m–r zero rows.

In fact, R and S have leading ones in the same columns, say r of them. Applying (2.11) to these columns shows that the first r columns of U are the first r columns of I_m. Hence we can write U, R, and S in block form as follows:

$$U = \left[\begin{array}{cc} I_r & M \\ 0 & W \end{array} \right], \quad R = \left[\begin{array}{cc} R_1 & R_2 \\ 0 & 0 \end{array} \right], \quad \text{and} \quad S = \left[\begin{array}{cc} S_1 & S_2 \\ 0 & 0 \end{array} \right]$$

where R_1 and S_1 are $r \times r$. Then block multiplication gives $UR = R$; that is, $S = R$. This completes the proof. □

Exercises for 2.5

Exercise 2.5.1 For each of the following elementary matrices, describe the corresponding elementary row operation and write the inverse.

a. $E = \left[\begin{array}{ccc} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right]$

b. $E = \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{array} \right]$

c. $E = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{array} \right]$

d. $E = \left[\begin{array}{rrr} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right]$

e. $E = \left[\begin{array}{ccc} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right]$

f. $E = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{array} \right]$

Exercise 2.5.2 In each case find an elementary matrix E such that $B = EA$.

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

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

c. $A = \left[\begin{array}{rr} 1 & 1 \\ -1 & 2 \end{array} \right]$, $B = \left[\begin{array}{rr} -1 & 2 \\ 1 & 1 \end{array} \right]$

d. $A = \left[\begin{array}{rr} 4 & 1 \\ 3 & 2 \end{array} \right]$, $B = \left[\begin{array}{rr} 1 & -1 \\ 3 & 2 \end{array} \right]$

e. $A = \left[\begin{array}{rr} -1 & 1 \\ 1 & -1 \end{array} \right]$, $B = \left[\begin{array}{rr} -1 & 1 \\ -1 & 1 \end{array} \right]$

f. $A = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}, B = \begin{bmatrix} -1 & 3 \\ 2 & 1 \end{bmatrix}$

Exercise 2.5.3 Let $A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$ and $C = \begin{bmatrix} -1 & 1 \\ 2 & 1 \end{bmatrix}$.

a. Find elementary matrices E_1 and E_2 such that $C = E_2E_1A$.

b. Show that there is *no* elementary matrix E such that $C = EA$.

Exercise 2.5.4 If E is elementary, show that A and EA differ in at most two rows.

Exercise 2.5.5

a. Is I an elementary matrix? Explain.

b. Is 0 an elementary matrix? Explain.

Exercise 2.5.6 In each case find an invertible matrix U such that $UA = R$ is in reduced row-echelon form, and express U as a product of elementary matrices.

a. $A = \begin{bmatrix} 1 & -1 & 2 \\ -2 & 1 & 0 \end{bmatrix}$

b. $A = \begin{bmatrix} 1 & 2 & 1 \\ 5 & 12 & -1 \end{bmatrix}$

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

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

Exercise 2.5.7 In each case find an invertible matrix U such that $UA = B$, and express U as a product of elementary matrices.

a. $A = \begin{bmatrix} 2 & 1 & 3 \\ -1 & 1 & 2 \end{bmatrix}, B = \begin{bmatrix} 1 & -1 & -2 \\ 3 & 0 & 1 \end{bmatrix}$

b. $A = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 3 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix}$

Exercise 2.5.8 In each case factor A as a product of elementary matrices.

a. $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$

b. $A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$

c. $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 2 & 1 & 6 \end{bmatrix}$

d. $A = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 4 \\ -2 & 2 & 15 \end{bmatrix}$

Exercise 2.5.9 Let E be an elementary matrix.

a. Show that E^T is also elementary of the same type.

b. Show that $E^T = E$ if E is of type I or II.

Exercise 2.5.10 Show that every matrix A can be factored as $A = UR$ where U is invertible and R is in reduced row-echelon form.

Exercise 2.5.11 If $A = \begin{bmatrix} 1 & 2 \\ 1 & -3 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 2 \\ -5 & -3 \end{bmatrix}$ find an elementary matrix F such that $AF = B$.

[*Hint*: See Exercise 2.5.9.]

Exercise 2.5.12 In each case find invertible U and V such that $UAV = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, where $r = \text{rank } A$.

a. $A = \begin{bmatrix} 1 & 1 & -1 \\ -2 & -2 & 4 \end{bmatrix}$

b. $A = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$

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

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

Exercise 2.5.13 Prove Lemma 2.5.1 for elementary matrices of:

a. type I;

b. type II.

Exercise 2.5.14 While trying to invert A, $\begin{bmatrix} A & I \end{bmatrix}$ is carried to $\begin{bmatrix} P & Q \end{bmatrix}$ by row operations. Show that $P = QA$.

Exercise 2.5.15 If A and B are $n \times n$ matrices and AB is a product of elementary matrices, show that the same is true of A.

Exercise 2.5.16 If U is invertible, show that the reduced row-echelon form of a matrix $\left[\begin{array}{cc} U & A \end{array} \right]$ is $\left[\begin{array}{cc} I & U^{-1}A \end{array} \right]$.

Exercise 2.5.17 Two matrices A and B are called **row-equivalent** (written $A \overset{r}{\sim} B$) if there is a sequence of elementary row operations carrying A to B.

a. Show that $A \overset{r}{\sim} B$ if and only if $A = UB$ for some invertible matrix U.

b. Show that:

    i. $A \overset{r}{\sim} A$ for all matrices A.

    ii. If $A \overset{r}{\sim} B$, then $B \overset{r}{\sim} A$

    iii. If $A \overset{r}{\sim} B$ and $B \overset{r}{\sim} C$, then $A \overset{r}{\sim} C$.

c. Show that, if A and B are both row-equivalent to some third matrix, then $A \overset{r}{\sim} B$.

d. Show that $\left[\begin{array}{rrrr} 1 & -1 & 3 & 2 \\ 0 & 1 & 4 & 1 \\ 1 & 0 & 8 & 6 \end{array} \right]$ and $\left[\begin{array}{rrrr} 1 & -1 & 4 & 5 \\ -2 & 1 & -11 & -8 \\ -1 & 2 & 2 & 2 \end{array} \right]$ are row-equivalent. [*Hint*: Consider (c) and Theorem 1.2.1.]

Exercise 2.5.18 If U and V are invertible $n \times n$ matrices, show that $U \overset{r}{\sim} V$. (See Exercise 2.5.17.)

Exercise 2.5.19 (See Exercise 2.5.17.) Find all matrices that are row-equivalent to:

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

b. $\left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 1 \end{array} \right]$

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

d. $\left[\begin{array}{ccc} 1 & 2 & 0 \\ 0 & 0 & 1 \end{array} \right]$

Exercise 2.5.20 Let A and B be $m \times n$ and $n \times m$ matrices, respectively. If $m > n$, show that AB is not invertible. [*Hint*: Use Theorem 1.3.1 to find $\mathbf{x} \neq \mathbf{0}$ with $B\mathbf{x} = \mathbf{0}$.]

Exercise 2.5.21 Define an *elementary column operation* on a matrix to be one of the following: (I) Interchange two columns. (II) Multiply a column by a nonzero scalar. (III) Add a multiple of a column to another column. Show that:

a. If an elementary column operation is done to an $m \times n$ matrix A, the result is AF, where F is an $n \times n$ elementary matrix.

b. Given any $m \times n$ matrix A, there exist $m \times m$ elementary matrices $E_1, \dots, E_k$ and $n \times n$ elementary matrices $F_1, \dots, F_p$ such that, in block form,

$$E_k \cdots E_1 A F_1 \cdots F_p = \left[\begin{array}{cc} I_r & 0 \\ 0 & 0 \end{array} \right]$$

Exercise 2.5.22 Suppose B is obtained from A by:

a. interchanging rows i and j;

b. multiplying row i by $k \neq 0$;

c. adding k times row i to row j $(i \neq j)$.

In each case describe how to obtain B^{-1} from A^{-1}. [*Hint*: See part (a) of the preceding exercise.]

Exercise 2.5.23 Two $m \times n$ matrices A and B are called **equivalent** (written $A \overset{e}{\sim} B$) if there exist invertible matrices U and V (sizes $m \times m$ and $n \times n$) such that $A = UBV$.

a. Prove the following the properties of equivalence.

    i. $A \overset{e}{\sim} A$ for all $m \times n$ matrices A.

    ii. If $A \overset{e}{\sim} B$, then $B \overset{e}{\sim} A$.

    iii. If $A \overset{e}{\sim} B$ and $B \overset{e}{\sim} C$, then $A \overset{e}{\sim} C$.

b. Prove that two $m \times n$ matrices are equivalent if they have the same rank. [*Hint*: Use part (a) and Theorem 2.5.3.]

2.6 Linear Transformations

If A is an $m \times n$ matrix, recall that the transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ defined by

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

is called the *matrix transformation induced* by A. In Section 2.2, we saw that many important geometric transformations were in fact matrix transformations. These transformations can be characterized in a different way. The new idea is that of a linear transformation, one of the basic notions in linear algebra. We define these transformations in this section, and show that they are really just the matrix transformations looked at in another way. Having these two ways to view them turns out to be useful because, in a given situation, one perspective or the other may be preferable.

Linear Transformations

Definition 2.13 Linear Transformations $\mathbb{R}^n \to \mathbb{R}^m$

A transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is called a ***linear transformation*** *if it satisfies the following two conditions for all vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$ and all scalars a:*

T1 $\quad T(\mathbf{x}+\mathbf{y}) = T(\mathbf{x}) + T(\mathbf{y})$

T2 $\quad T(a\mathbf{x}) = aT(\mathbf{x})$

Of course, $\mathbf{x}+\mathbf{y}$ and $a\mathbf{x}$ here are computed in $\mathbb{R}^n$, while $T(\mathbf{x})+T(\mathbf{y})$ and $aT(\mathbf{x})$ are in $\mathbb{R}^m$. We say that T *preserves addition* if T1 holds, and that T *preserves scalar multiplication* if T2 holds. Moreover, taking $a = 0$ and $a = -1$ in T2 gives

$$T(\mathbf{0}) = \mathbf{0} \quad \text{and} \quad T(-\mathbf{x}) = -T(\mathbf{x}) \quad \text{for all } \mathbf{x}$$

Hence T preserves the zero vector and the negative of a vector. Even more is true.

Recall that a vector $\mathbf{y}$ in $\mathbb{R}^n$ is called a **linear combination** of vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ if $\mathbf{y}$ has the form

$$\mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k$$

for some scalars $a_1, a_2, \dots, a_k$. Conditions T1 and T2 combine to show that every linear transformation T *preserves linear combinations* in the sense of the following theorem. This result is used repeatedly in linear algebra.

Theorem 2.6.1: Linearity Theorem

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then for each $k = 1, 2, \dots$

$$T(a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k) = a_1T(\mathbf{x}_1) + a_2T(\mathbf{x}_2) + \cdots + a_kT(\mathbf{x}_k)$$

for all scalars a_i and all vectors $\mathbf{x}_i$ in $\mathbb{R}^n$.