It is important to observe that the product of two matrices need not be defined; the product is defined if and only if the number of columns in the first matrix coincides with the number of rows in the second matrix. Thus it is meaningless to interchange the order of the factors in (a), (b), and (c) above. Frequently we shall write products such as AB without explicitly mentioning the sizes of the factors and in such cases it will be understood that the product is defined. From (d), (e), (f), (g) we find that even when the products AB and BA are both defined it need not be true that $AB = BA$; in other words, matrix multiplication is *not commutative.*

EXAMPLE 11.

(a) If I is the $m \times m$ identity matrix and A is an $m \times n$ matrix, $IA = A$.

(b) If I is the $n \times n$ identity matrix and A is an $m \times n$ matrix, $AI = A$.

(c) If $0^{k,m}$ is the $k \times m$ zero matrix, $0^{k,n} = 0^{k,m}A$. Similarly, $A0^{n,p} = 0^{m,p}$.

EXAMPLE 12. Let A be an $m \times n$ matrix over F. Our earlier shorthand notation, $AX = Y$, for systems of linear equations is consistent with our definition of matrix products. For if

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

with x_i in F, then AX is the $m \times 1$ matrix

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

such that $y_i = A_{i1}x_1 + A_{i2}x_2 + \cdots + A_{in}x_n$.

The use of column matrices suggests a notation which is frequently useful. If B is an $n \times p$ matrix, the columns of B are the $1 \times n$ matrices $B_1, \ldots, B_p$ defined by

$$B_j = \begin{bmatrix} B_{1j} \\ \vdots \\ B_{nj} \end{bmatrix}, \qquad 1 \leq j \leq p.$$

The matrix B is the succession of these columns:

$$B = [B_1, \ldots, B_p].$$

The i, j entry of the product matrix AB is formed from the ith row of A

and the jth column of B. The reader should verify that the jth column of AB is AB_j:

$$AB = [AB_1, \ldots, AB_p].$$

In spite of the fact that a product of matrices depends upon the order in which the factors are written, it is independent of the way in which they are associated, as the next theorem shows.

Theorem 8. *If* A, B, C *are matrices over the field* F *such that the products* BC *and* A(BC) *are defined, then so are the products* AB, (AB)C *and*

$$\mathrm{A(BC)} = \mathrm{(AB)C}.$$

Proof. Suppose B is an $n \times p$ matrix. Since BC is defined, C is a matrix with p rows, and BC has n rows. Because $A(BC)$ is defined we may assume A is an $m \times n$ matrix. Thus the product AB exists and is an $m \times p$ matrix, from which it follows that the product $(AB)C$ exists. To show that $A(BC) = (AB)C$ means to show that

$$[A(BC)]_{ij} = [(AB)C]_{ij}$$

for each i, j. By definition

$$\begin{aligned}
[A(BC)]_{ij} &= \sum_r A_{ir}(BC)_{rj} \\
&= \sum_r A_{ir} \sum_s B_{rs}C_{sj} \\
&= \sum_r \sum_s A_{ir}B_{rs}C_{sj} \\
&= \sum_s \sum_r A_{ir}B_{rs}C_{sj} \\
&= \sum_s (\sum_r A_{ir}B_{rs})C_{sj} \\
&= \sum_s (AB)_{is}C_{sj} \\
&= [(AB)C]_{ij}. \quad \blacksquare
\end{aligned}$$

When A is an $n \times n$ (square) matrix, the product AA is defined. We shall denote this matrix by A^2. By Theorem 8, $(AA)A = A(AA)$ or $A^2A = AA^2$, so that the product AAA is unambiguously defined. This product we denote by A^3. In general, the product $AA \cdots A$ (k times) is unambiguously defined, and we shall denote this product by A^k.

Note that the relation $A(BC) = (AB)C$ implies among other things that linear combinations of linear combinations of the rows of C are again linear combinations of the rows of C.

If B is a given matrix and C is obtained from B by means of an elementary row operation, then each row of C is a linear combination of the rows of B, and hence there is a matrix A such that $AB = C$. In general there are many such matrices A, and among all such it is convenient and possible to choose one having a number of special properties. Before going into this we need to introduce a class of matrices.

Definition. *An* m $\times$ n *matrix is said to be an* **elementary matrix** *if it can be obtained from the* m $\times$ m *identity matrix by means of a single elementary row operation.*

EXAMPLE 13. A 2×2 elementary matrix is necessarily one of the following:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}$$

$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}, \quad c \neq 0, \quad \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}, \quad c \neq 0.$$

Theorem 9. *Let* e *be an elementary row operation and let* E *be the* m $\times$ m *elementary matrix* E = e(I). *Then, for every* m $\times$ n *matrix* A,

$$\mathrm{e(A) = EA}.$$

Proof. The point of the proof is that the entry in the ith row and jth column of the product matrix EA is obtained from the ith row of E and the jth column of A. The three types of elementary row operations should be taken up separately. We shall give a detailed proof for an operation of type (ii). The other two cases are even easier to handle than this one and will be left as exercises. Suppose $r \neq s$ and e is the operation 'replacement of row r by row r plus c times row s.' Then

$$E_{ik} = \begin{cases} \delta_{ik}, & i \neq r \\ \delta_{rk} + c\delta_{sk}, & i = r. \end{cases}$$

Therefore,

$$(EA)_{ij} = \sum_{k=1}^{m} E_{ik}A_{kj} = \begin{cases} A_{ik}, & i \neq r \\ A_{rj} + cA_{sj}, & i = r. \end{cases}$$

In other words $EA = e(A)$. ▌

Corollary. *Let* A *and* B *be* m $\times$ n *matrices over the field* F. *Then* B *is row-equivalent to* A *if and only if* B = PA, *where* P *is a product of* m $\times$ m *elementary matrices.*

Proof. Suppose $B = PA$ where $P = E_s \cdots E_2E_1$ and the E_i are $m \times m$ elementary matrices. Then E_1A is row-equivalent to A, and $E_2(E_1A)$ is row-equivalent to E_1A. So E_2E_1A is row-equivalent to A; and continuing in this way we see that $(E_s \cdots E_1)A$ is row-equivalent to A.

Now suppose that B is row-equivalent to A. Let $E_1, E_2, \ldots, E_s$ be the elementary matrices corresponding to some sequence of elementary row operations which carries A into B. Then $B = (E_s \cdots E_1)A$. ▌

Exercises

1. Let

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}, \qquad C = [1 \quad -1].$$

Compute ABC and CAB.

2. Let

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

Verify directly that $A(AB) = A^2B$.

3. Find two different 2×2 matrices A such that $A^2 = 0$ but $A \neq 0$.

4. For the matrix A of Exercise 2, find elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k \cdots E_2E_1A = I.$$

5. Let

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

Is there a matrix C such that $CA = B$?

6. Let A be an $m \times n$ matrix and B an $n \times k$ matrix. Show that the columns of $C = AB$ are linear combinations of the columns of A. If $\alpha_1, \ldots, \alpha_n$ are the columns of A and $\gamma_1, \ldots, \gamma_k$ are the columns of C, then

$$\gamma_j = \sum_{r=1}^{n} B_{rj}\alpha_r.$$

7. Let A and B be 2×2 matrices such that $AB = I$. Prove that $BA = I$.

8. Let

$$C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

be a 2×2 matrix. We inquire when it is possible to find 2×2 matrices A and B such that $C = AB - BA$. Prove that such matrices can be found if and only if $C_{11} + C_{22} = 0$.

1.6. Invertible Matrices

Suppose P is an $m \times m$ matrix which is a product of elementary matrices. For each $m \times n$ matrix A, the matrix $B = PA$ is row-equivalent to A; hence A is row-equivalent to B and there is a product Q of elementary matrices such that $A = QB$. In particular this is true when A is the

$m \times m$ identity matrix. In other words, there is an $m \times m$ matrix Q, which is itself a product of elementary matrices. such that $QP = I$. As we shall soon see, the existence of a Q with $QP = I$ is equivalent to the fact that P is a product of elementary matrices.

Definition. *Let* A *be an* $n \times n$ *(square) matrix over the field* F. *An* $n \times n$ *matrix* B *such that* $BA = I$ *is called a* **left inverse** *of* A; *an* $n \times n$ *matrix* B *such that* $AB = I$ *is called a* **right inverse** *of* A. *If* $AB = BA = I$, *then* B *is called a* **two-sided inverse** *of* A *and* A *is said to be* **invertible.**

Lemma. *If* A *has a left inverse* B *and a right inverse* C, *then* $B = C$.

Proof. Suppose $BA = I$ and $AC = I$. Then

$$B = BI = B(AC) = (BA)C = IC = C. \quad \blacksquare$$

Thus if A has a left and a right inverse, A is invertible and has a unique two-sided inverse, which we shall denote by A^{-1} and simply call **the inverse** of A.

Theorem 10. *Let* A *and* B *be* $n \times n$ *matrices over* F.

(i) *If* A *is invertible, so is* A^{-1} *and* $(A^{-1})^{-1} = A$.
(ii) *If both* A *and* B *are invertible, so is* AB, *and* $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. The first statement is evident from the symmetry of the definition. The second follows upon verification of the relations

$$(AB)(B^{-1}A^{-1}) = (B^{-1}A^{-1})(AB) = I. \quad \blacksquare$$

Corollary. *A product of invertible matrices is invertible.*

Theorem 11. *An elementary matrix is invertible.*

Proof. Let E be an elementary matrix corresponding to the elementary row operation e. If e_1 is the inverse operation of e (Theorem 2) and $E_1 = e_1(I)$, then

$$EE_1 = e(E_1) = e(e_1(I)) = I$$

and

$$E_1E = e_1(E) = e_1(e(I)) = I$$

so that E is invertible and $E_1 = E^{-1}$. $\blacksquare$

EXAMPLE 14.

(a)
$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(b)
$$\begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -c \\ 0 & 1 \end{bmatrix}$$

(c)
$$\begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -c & 1 \end{bmatrix}$$

(d) When $c \neq 0$,
$$\begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} c^{-1} & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & c^{-1} \end{bmatrix}.$$

Theorem 12. *If A is an $n \times n$ matrix, the following are equivalent.*

(i) *A is invertible.*
(ii) *A is row-equivalent to the $n \times n$ identity matrix.*
(iii) *A is a product of elementary matrices.*

Proof. Let R be a row-reduced echelon matrix which is row-equivalent to A. By Theorem 9 (or its corollary),
$$R = E_k \cdots E_2 E_1 A$$
where $E_1, \ldots, E_k$ are elementary matrices. Each E_j is invertible, and so
$$A = E_1^{-1} \cdots E_k^{-1} R.$$
Since products of invertible matrices are invertible, we see that A is invertible if and only if R is invertible. Since R is a (square) row-reduced echelon matrix, R is invertible if and only if each row of R contains a non-zero entry, that is, if and only if $R = I$. We have now shown that A is invertible if and only if $R = I$, and if $R = I$ then $A = E_k^{-1} \cdots E_1^{-1}$. It should now be apparent that (i), (ii), and (iii) are equivalent statements about A. ∎

Corollary. *If A is an invertible $n \times n$ matrix and if a sequence of elementary row operations reduces A to the identity, then that same sequence of operations when applied to I yields A^{-1}.*

Corollary. *Let A and B be $m \times n$ matrices. Then B is row-equivalent to A if and only if $B = PA$ where P is an invertible $m \times m$ matrix.*

Theorem 13. *For an $n \times n$ matrix A, the following are equivalent.*

(i) *A is invertible.*
(ii) *The homogeneous system $AX = 0$ has only the trivial solution $X = 0$.*
(iii) *The system of equations $AX = Y$ has a solution X for each $n \times 1$ matrix Y.*

Proof. According to Theorem 7, condition (ii) is equivalent to the fact that A is row-equivalent to the identity matrix. By Theorem 12, (i) and (ii) are therefore equivalent. If A is invertible, the solution of $AX = Y$ is $X = A^{-1}Y$. Conversely, suppose $AX = Y$ has a solution for each given Y. Let R be a row-reduced echelon matrix which is row-