PEARSON NEW INTERNATIONAL EDITION

Linear Algebra

S. Friedberg A. Insel L. Spence

Fourth Edition

ALWAYS LEARNING

PEARSON

Pearson New International Edition

Linear Algebra

S. Friedberg   A. Insel   L. Spence

Fourth Edition

PEARSON

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 $\text{rank}(A) = \text{rank}(B) = r$, $\text{rank}(B') = r-1$. 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) $\text{rank}(A^t) = \text{rank}(A)$.

(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,

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

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

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

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

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

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

(c) By (a),

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

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

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

(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),

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

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 $\text{rank}(A) = 2$.

(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 $\text{rank}(A) = 2$.

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 $\text{rank}(A) = 2$.

(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:

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