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 $\mathsf{L}_A : \mathsf{F}^n \to \mathsf{F}^m$.*

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 $\mathsf{T} : \mathsf{V} \to \mathsf{W}$ 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

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

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),

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

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

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

Thus

$$\operatorname{rank}(A) = \dim(\mathsf{R}(\mathsf{L}_A)) = \dim(\operatorname{span}(\{a_1, a_2, \ldots, a_n\})). \quad \blacksquare$$

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

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

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, $\operatorname{rank}(A) = \operatorname{rank}(D)$. Clearly, however, $\operatorname{rank}(D) = 3$; so $\operatorname{rank}(A) = 3$. ◆

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 $r = \operatorname{rank}(A)$; 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 $\operatorname{rank}(D) = \operatorname{rank}(A) = 1$ 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.