Section 7.5 The Rank Theorem and Its Implications

Theorem 7.5.1 (***The Rank Theorem***) *The row space and column space of a matrix have the same dimension.*

EXAMPLE 1 Row Space and Column Space Have the Same Dimension

In Example 4 of Section 7.3 we showed that the row space of the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 2 \\ -2 & 1 & -3 & -2 & -4 \\ 0 & 5 & -14 & -9 & 0 \\ 2 & 10 & -28 & -18 & 4 \end{bmatrix} \tag{1}$$

is three-dimensional, so the rank theorem implies that the column space is also three-dimensional. Let us confirm this by finding a basis for the column space.

One way to do this is to transpose A (which converts columns to rows) and then find a basis for the row space of A^T by reducing it to row echelon form and extracting the nonzero row vectors.

Proceeding in this way, we first transpose A to obtain

$$A^T = \begin{bmatrix} 1 & -2 & 0 & 2 \\ 0 & 1 & 5 & 10 \\ 0 & -3 & -14 & -28 \\ 0 & -2 & -9 & -18 \\ 2 & -4 & 0 & 4 \end{bmatrix}$$

and then reduce this matrix to row echelon form to obtain

$$\begin{bmatrix} 1 & -2 & 0 & 2 \\ 0 & 1 & 5 & 10 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{2}$$

(verify). The nonzero row vectors in this matrix form a basis for the row space of A^T, so the column space of A is three-dimensional as anticipated. If desired, a basis of column vectors for the column space of A can be obtained by transposing the row vectors in (2) to obtain

$$\mathbf{c}_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 0 \\ 1 \\ 5 \\ 10 \end{bmatrix}, \quad \mathbf{c}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}$$

■

Theorem 7.5.2 *If A is an $m \times n$ matrix, then*

$$\operatorname{rank}(A) = \operatorname{rank}(A^T) \tag{3}$$

This result has some important implications. For example, if A is an $m \times n$ matrix, then applying Theorem 7.4.1 to A^T yields

$$\operatorname{rank}(A^T) + \operatorname{nullity}(A^T) = m$$

which we can rewrite using (3) as

$$\operatorname{rank}(A) + \operatorname{nullity}(A^T) = m \tag{4}$$

This relationship now makes it possible to express the dimensions of all four fundamental spaces of a matrix in terms of the size and rank of the matrix. Specifically, if A is an $m \times n$ matrix with rank k, then

$$\begin{aligned} &\dim(\operatorname{row}(A)) = k, \quad &&\dim(\operatorname{null}(A)) = n - k \\ &\dim(\operatorname{col}(A)) = k, \quad &&\dim(\operatorname{null}(A^T)) = m - k \end{aligned} \tag{5}$$

Theorem 7.4.1 ***(The Dimension Theorem for Matrices)*** *If A is an $m \times n$ matrix, then*

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = n \tag{2}$$

EXAMPLE 2 Dimensions of the Fundamental Spaces from the Rank

Find the rank of

$$A = \begin{bmatrix} 1 & 2 & -3 & 1 & 1 \\ -3 & 1 & 7 & -1 & 1 \\ -2 & 3 & 4 & 0 & 2 \end{bmatrix}$$

and then use that result to compute the dimensions of the fundamental spaces of A.

Solution The rank of A is the number of nonzero rows in any row echelon form of A, so we will begin by reducing A to row echelon form. Introducing the required zeros in the first column yields

$$\begin{bmatrix} 1 & 2 & -3 & 1 & 1 \\ 0 & 7 & -2 & 2 & 4 \\ 0 & 7 & -2 & 2 & 4 \end{bmatrix}$$

At this point there is no need to go any further, since it is now evident that the row space is two-dimensional. Thus, A has rank 2 and

$$\dim(\text{row}(A)) = \text{rank} = 2, \quad \dim(\text{null}(A)) = \text{number of columns} - \text{rank} = 5 - 2 = 3$$
$$\dim(\text{col}(A)) = \text{rank} = 2, \quad \dim(\text{null}(A^T)) = \text{number of rows} - \text{rank} = 3 - 2 = 1$$

■

The consistency or inconsistency of a linear system $A\mathbf{x} = \mathbf{b}$ is determined by the relationship between the vector $\mathbf{b}$ and the column vectors of A. To see why this so, suppose that the successive column vectors of A are $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, and use Formula (10) of Section 3.1 to rewrite the system as

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{b} \tag{5}$$

We see from this expression that $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ can be expressed as a linear combination of the column vectors of A, and if so, the solutions of the system are given by the coefficients in (5).

This idea can be expressed in a slightly different way: If A is an $m \times n$ matrix, then to say that $\mathbf{b}$ is a linear combination of the column vectors of A is the same as saying that $\mathbf{b}$ is in the subspace of R^m spanned by the column vectors of A. This subspace is called the ***column space*** of A and is denoted by $\text{col}(A)$. The following theorem summarizes this discussion.

Theorem 3.5.5 *A linear system $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ is in the column space of A.*

Theorem 7.5.3 (***The Consistency Theorem***) *If $A\mathbf{x} = \mathbf{b}$ is a linear system of m equations in n unknowns, then the following statements are equivalent.*

(*a*) $A\mathbf{x} = \mathbf{b}$ *is consistent.*

(*b*) $\mathbf{b}$ *is in the column space of A.*

(*c*) *The coefficient matrix A and the augmented matrix $[A \mid \mathbf{b}]$ have the same rank.*

The equivalence of parts (*a*) and (*b*) was given in Theorem 3.5.5, so we need only prove that (*b*) $\Leftrightarrow$ (*c*). The equivalence (*a*) $\Leftrightarrow$ (*c*) will then follow as a logical consequence.

***Proof* (*b*) $\Leftrightarrow$ (*c*)** If $\mathbf{b}$ is in the column space of A, then Theorem 7.2.5 implies that the column spaces of A and $[A \mid \mathbf{b}]$ have the same dimension; that is, the two matrices have the same rank. Conversely, if A and $[A \mid \mathbf{b}]$ have the same rank, then their column spaces have the same dimension, so Theorem 7.2.5 implies that $\mathbf{b}$ is a linear combination of the column vectors of A. ■

Theorem 7.2.5 *Let S be a nonempty set of vectors in R^n, and let S' be a set that results by adding additional vectors in R^n to S.*

(*a*) *If the additional vectors are in* $\text{span}(S)$*, then* $\text{span}(S') = \text{span}(S)$.

(*b*) *If* $\text{span}(S') = \text{span}(S)$*, then the additional vectors are in* $\text{span}(S)$.

(*c*) *If* $\text{span}(S')$ *and* $\text{span}(S)$ *have the same dimension, then the additional vectors are in* $\text{span}(S)$ *and* $\text{span}(S') = \text{span}(S)$.

EXAMPLE 3 Visualizing the Consistency Theorem

To obtain a better understanding of the relationship between the ranks of the coefficient and augmented matrices of a linear system, consider the system

$$\begin{aligned} x_1 - 2x_2 - 3x_3 &= -4 \\ -3x_1 + 7x_2 - x_3 &= -3 \\ 2x_1 - 5x_2 + 4x_3 &= 7 \\ -3x_1 + 6x_2 + 9x_3 &= -1 \end{aligned}$$

The augmented matrix for the system is

$$\left[\begin{array}{rrr|r} 1 & -2 & -3 & -4 \\ -3 & 7 & -1 & -3 \\ 2 & -5 & 4 & 7 \\ -3 & 6 & 9 & -1 \end{array}\right]$$

and the reduced row echelon form of this matrix is (verify)

$$\left[\begin{array}{ccc|c} 1 & 0 & -23 & 0 \\ 0 & 1 & -10 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \tag{6}$$

The "bad" third row in this matrix makes it evident that the system is inconsistent.

However, this row also causes the corresponding row echelon form of the coefficient matrix to have smaller rank than the row echelon form of the augmented matrix [cover the last column of (6) to see this].

This example should make it evident that the augmented matrix and the coefficient matrix of a linear system have the same rank if and only if there are no bad rows in any row echelon form of the augmented matrix, or equivalently, if and only if the system is consistent. ■

Definition 7.5.4 An $m \times n$ matrix A is said to have ***full column rank*** if its column vectors are linearly independent, and it is said to have ***full row rank*** if its row vectors are linearly independent.

→ full column/row rank = "posto coluna/linha máximo".

Since the column vectors of a matrix span the column space and the row vectors span the row space, the column vectors of a matrix with full column rank must be a basis for the column space, and the row vectors of a matrix with full row rank must be a basis for the row space. Thus, we have the following alternative way of viewing the concepts of full column rank and full row rank.

Theorem 7.5.5 *Let A be an $m \times n$ matrix.*

(*a*) *A has full column rank if and only if the column vectors of A form a basis for the column space, that is, if and only if* $\text{rank}(A) = n$.

(*b*) *A has full row rank if and only if the row vectors of A form a basis for the row space, that is, if and only if* $\text{rank}(A) = m$.

EXAMPLE 4

Full Column Rank and Full Row Rank

The matrix

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

has full column rank because the column vectors are not scalar multiples of one another; it does not have full row rank because three vectors in R^2 are linearly dependent. In contrast,

$$A^T = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 1 \end{bmatrix}$$

has full row rank but not full column rank. ■

Theorem 7.5.6 *If A is an $m \times n$ matrix, then the following statements are equivalent.*

(*a*) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(*b*) $A\mathbf{x} = \mathbf{b}$ *has at most one solution for every* $\mathbf{b}$ *in* R^m.

(*c*) *A has full column rank.*

Since the equivalence of parts (*a*) and (*b*) is the content of Theorem 3.5.3, it suffices to show that parts (*a*) and (*c*) are equivalent to complete the proof.

Theorem 3.5.3 *If A is an $m \times n$ matrix, then the following statements are equivalent.*

(*a*) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(*b*) $A\mathbf{x} = \mathbf{b}$ *has at most one solution for every* $\mathbf{b}$ *in* R^m *(i.e., is inconsistent or has a unique solution).*

Theorem 3.5.2 *A general solution of a consistent linear system* $A\mathbf{x} = \mathbf{b}$ *can be obtained by adding a particular solution of* $A\mathbf{x} = \mathbf{b}$ *to a general solution of* $A\mathbf{x} = \mathbf{0}$.

Theorem 7.5.6 *If A is an $m \times n$ matrix, then the following statements are equivalent.*

(a) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(b) $A\mathbf{x} = \mathbf{b}$ *has at most one solution for every* $\mathbf{b}$ *in* R^m.

(c) *A has full column rank.*

If (a) then (b):
For trivial solution, null(A) = 0 and rank(A) = n.
Then, either **b** is in col(A), in an unique way (m = n, columns as basis vectors),
or it is not (m > n, m vectors would be required to fully span the R^m);

If (b), then (a):
If there is more than one solution for every **b**, then the columns of A are LD,
and rank(A) < n and null(A) > 0.

Theorem 7.5.6 *If A is an $m \times n$ matrix, then the following statements are equivalent.*

(*a*) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(*b*) $A\mathbf{x} = \mathbf{b}$ *has at most one solution for every* $\mathbf{b}$ *in* R^m.

(*c*) *A has full column rank.*

Proof (a) ⇔ (c) Let $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$ be the column vectors of A, and write the system $A\mathbf{x} = \mathbf{0}$ in the vector form

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n = \mathbf{0} \tag{7}$$

Thus, to say that $A\mathbf{x} = \mathbf{0}$ has only the trivial solution is equivalent to saying that the n column vectors in (7) are linearly independent; that is, $A\mathbf{x} = \mathbf{0}$ has only the trivial solution if and only if A has full column rank. ■

EXAMPLE 5 Implications of Full Column Rank

We showed in Example 4 that

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

has full column rank. Thus, Theorem 7.5.6 implies that the system $A\mathbf{x} = \mathbf{0}$ has only the trivial solution and that the system $A\mathbf{x} = \mathbf{b}$ has at most one solution for every $\mathbf{b}$ in R^3. We will leave it for you to confirm the first statement by solving the system $A\mathbf{x} = \mathbf{0}$; and we will show that $A\mathbf{x} = \mathbf{b}$ has at most one solution for every vector $\mathbf{b} = (b_1, b_2, b_3)$ in R^3.

Reducing the augmented matrix $[A \mid \mathbf{b}]$ until the left side is in reduced row echelon form yields

$$\left[\begin{array}{cc|c} 1 & 0 & b_1 \\ 0 & 1 & b_2 - 2b_1 \\ 0 & 0 & b_3 - b_2 + 5b_1 \end{array}\right]$$

(verify), so there are two possibilities: $b_3 - b_2 + 5b_1 \neq 0$ or $b_3 - b_2 + 5b_1 = 0$. In the first case the system is inconsistent, and in the second case the system has the unique solution $x_1 = b_1$, $x_2 = b_2 - 2b_1$. In either case it is correct to say that there is at most one solution. ■

OPTIONAL *Proof of Theorem 7.5.1*

We want to prove that the row space and column space of an $m \times n$ matrix A have the same dimension. For this purpose, assume that A has rank k, which implies that the reduced row echelon form R has exactly k nonzero row vectors, say $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_k$. Since A and R have the same row space by Theorem 7.3.7, it follows that the row vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m$ of A can be expressed as linear combinations of the row vectors of R, say

$$\begin{aligned}
\mathbf{a}_1 &= c_{11}\mathbf{r}_1 + c_{12}\mathbf{r}_2 + c_{13}\mathbf{r}_3 + \cdots + c_{1k}\mathbf{r}_k \\
\mathbf{a}_2 &= c_{21}\mathbf{r}_1 + c_{22}\mathbf{r}_2 + c_{23}\mathbf{r}_3 + \cdots + c_{2k}\mathbf{r}_k \\
&\;\;\vdots \\
\mathbf{a}_m &= c_{m1}\mathbf{r}_1 + c_{m2}\mathbf{r}_2 + c_{m3}\mathbf{r}_3 + \cdots + c_{mk}\mathbf{r}_k
\end{aligned} \tag{10}$$

$$
\begin{aligned}
\mathbf{a}_1 &= c_{11}\mathbf{r}_1 + c_{12}\mathbf{r}_2 + c_{13}\mathbf{r}_3 + \cdots + c_{1k}\mathbf{r}_k \\
\mathbf{a}_2 &= c_{21}\mathbf{r}_1 + c_{22}\mathbf{r}_2 + c_{23}\mathbf{r}_3 + \cdots + c_{2k}\mathbf{r}_k \\
\vdots \quad & \quad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\
\mathbf{a}_m &= c_{m1}\mathbf{r}_1 + c_{m2}\mathbf{r}_2 + c_{m3}\mathbf{r}_3 + \cdots + c_{mk}\mathbf{r}_k
\end{aligned}
\tag{10}
$$

Next we equate corresponding components on the two sides of each equation. For this purpose let a_{ij} be the jth component of $\mathbf{a}_i$, and let r_{ij} be the jth component of $\mathbf{r}_i$. Thus, the relationships between the jth components on the two sides of (10) are

$$
\begin{aligned}
a_{1j} &= c_{11}r_{1j} + c_{12}r_{2j} + c_{13}r_{3j} + \cdots + c_{1k}r_{kj} \\
a_{2j} &= c_{21}r_{1j} + c_{22}r_{2j} + c_{23}r_{3j} + \cdots + c_{2k}r_{kj} \\
\vdots \quad & \quad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\
a_{mj} &= c_{m1}r_{1j} + c_{m2}r_{2j} + c_{m3}r_{3j} + \cdots + c_{mk}r_{kj}
\end{aligned}
$$

$$
\begin{aligned}
a_{1j} &= c_{11}r_{1j} + c_{12}r_{2j} + c_{13}r_{3j} + \cdots + c_{1k}r_{kj} \\
a_{2j} &= c_{21}r_{1j} + c_{22}r_{2j} + c_{23}r_{3j} + \cdots + c_{2k}r_{kj} \\
&\vdots \\
a_{mj} &= c_{m1}r_{1j} + c_{m2}r_{2j} + c_{m3}r_{3j} + \cdots + c_{mk}r_{kj}
\end{aligned}
$$

which we can rewrite in matrix form as

$$
\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} = r_{1j} \begin{bmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{m1} \end{bmatrix} + r_{2j} \begin{bmatrix} c_{12} \\ c_{22} \\ \vdots \\ c_{m2} \end{bmatrix} + r_{3j} \begin{bmatrix} c_{13} \\ c_{23} \\ \vdots \\ c_{m3} \end{bmatrix} + \cdots + r_{kj} \begin{bmatrix} c_{1k} \\ c_{2k} \\ \vdots \\ c_{mk} \end{bmatrix}
$$

Since the left side of this equation is the jth column vector of A,

we have shown that the k column vectors on the right side of the equation

span the column space of A.

Thus, the dimension of the column space of A is at most k; that is,

$$\dim(\mathrm{col}(A)) \leq \dim(\mathrm{row}(A)) \tag{11}$$

It follows from this that

$$\dim(\mathrm{col}(A^T)) \leq \dim(\mathrm{row}(A^T))$$

→ If (11) holds for a generic matrix A, it also holds for A^T.

or

$$\dim(\mathrm{row}(A)) \leq \dim(\mathrm{col}(A)) \tag{12}$$

We can conclude from (11) and (12) that $\dim(\mathrm{row}(A)) = \dim(\mathrm{col}(A))$. ■

OVERDETERMINED AND UNDERDETERMINED LINEAR SYSTEMS

In engineering applications, the equations in a linear system $A\mathbf{x} = \mathbf{b}$ are often mathematical formulations of physical constraints on a set of variables, and engineers generally try to match the number of variables and constraints.

However, this is not always possible, so an engineer may be faced with a linear system that has more equations than unknowns (called an ***overdetermined system***) or a linear system that has fewer equations than unknowns (called an ***underdetermined system***).

The occurrence of an overdetermined or underdetermined linear system in applications often signals that some undesirable physical phenomenon may occur. The following theorem explains why.

Theorem 7.5.7 *Let A be an $m \times n$ matrix.*

(a) (***Overdetermined Case***) *If $m > n$, then the system $A\mathbf{x} = \mathbf{b}$ is inconsistent for some vector $\mathbf{b}$ in R^m.*

(b) (***Underdetermined Case***) *If $m < n$, then for every vector $\mathbf{b}$ in R^m the system $A\mathbf{x} = \mathbf{b}$ is either inconsistent or has infinitely many solutions.*

Proof (a) If $m > n$, then the column vectors of A cannot span R^m. Thus, there is at least one vector $\mathbf{b}$ in R^m that is not a linear combination of the column vectors of A, and for such a $\mathbf{b}$ the system $A\mathbf{x} = \mathbf{b}$ has no solution.

Proof (b) If $m < n$, then the column vectors of A must be linearly dependent (n vectors in R^m). This implies that $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions, so the result follows from Theorem 3.5.2. ■

In proof (b), by Theorem 7.5.6, A**x** = **0** will have more than the trivial solution (the columns of A are LD, rank(A) < n and null(A) > 0).
Then, if A**x** = **0** has infinitely many solutions, either A**x** = **b** will follow the same way, or the system does not have a particular solution (inconsistent).

Theorem 3.5.2 *A general solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$ can be obtained by adding a particular solution of $A\mathbf{x} = \mathbf{b}$ to a general solution of $A\mathbf{x} = \mathbf{0}$.*

EXAMPLE 6
A Misbehaving Robot

To express Theorem 7.5.7 in transformation terms, think of $A\mathbf{x}$ as a matrix transformation from R^n to R^m, and think of the vector $\mathbf{b}$ in the equation $A\mathbf{x} = \mathbf{b}$ as some output that we would like the transformation to produce in response to some input $\mathbf{x}$.

Part (*a*) of Theorem 7.5.7 states that if $m > n$, then there is some output that cannot be produced by any input, and part (*b*) states that if $m < n$, then for each possible output $\mathbf{b}$ there is either no input that produces that output or there are infinitely many inputs that produce that output.

Thus, for example, if the input $\mathbf{x}$ is a vector of voltages to the driving motors of a robot, and if the output $\mathbf{b}$ is a vector of speeds and position coordinates that describe the action of the robot in response to the input, then

an overdetermined system governs a robot that cannot achieve certain desired actions, and an underdetermined system governs a robot for which certain actions can be achieved in infinitely many ways, which may not be desirable. ■

MATRICES OF THE FORM A^TA AND AA^T

Matrices of the form A^TA and AA^T play an important role in many applications, so we will now focus our attention on matrices of this form.

To start, recall from Formula (9) of Section 3.6 that if A is an $m \times n$ matrix with column vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, then

$$A^TA = \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1 & \mathbf{a}_1 \cdot \mathbf{a}_2 & \cdots & \mathbf{a}_1 \cdot \mathbf{a}_n \\ \mathbf{a}_2 \cdot \mathbf{a}_1 & \mathbf{a}_2 \cdot \mathbf{a}_2 & \cdots & \mathbf{a}_2 \cdot \mathbf{a}_n \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_n \cdot \mathbf{a}_1 & \mathbf{a}_n \cdot \mathbf{a}_2 & \cdots & \mathbf{a}_n \cdot \mathbf{a}_n \end{bmatrix} \tag{8}$$

It is symmetric.

Since transposing a matrix converts columns to rows and rows to columns, it follows from (8) that if $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m$ are the row vectors of A, then

$$AA^T = \begin{bmatrix} \mathbf{r}_1 \cdot \mathbf{r}_1 & \mathbf{r}_1 \cdot \mathbf{r}_2 & \cdots & \mathbf{r}_1 \cdot \mathbf{r}_m \\ \mathbf{r}_2 \cdot \mathbf{r}_1 & \mathbf{r}_2 \cdot \mathbf{r}_2 & \cdots & \mathbf{r}_2 \cdot \mathbf{r}_m \\ \vdots & \vdots & & \vdots \\ \mathbf{r}_m \cdot \mathbf{r}_1 & \mathbf{r}_m \cdot \mathbf{r}_2 & \cdots & \mathbf{r}_m \cdot \mathbf{r}_m \end{bmatrix} \tag{9}$$

It is also symmetric.

The next theorem provides some important links between properties of a general matrix A, its transpose A^T, and the square symmetric matrix A^TA.

Theorem 7.5.8 *If A is an $m \times n$ matrix, then:*

(a) *A and A^TA have the same null space.*

(b) *A and A^TA have the same row space.*

(c) *A^T and A^TA have the same column space.*

(d) *A and A^TA have the same rank.*

We will prove part (*a*) and leave the remaining proofs for the exercises.

Proof **(*a*)** We must show that every solution of $A\mathbf{x} = \mathbf{0}$ is a solution of $A^TA\mathbf{x} = \mathbf{0}$, and conversely. If $\mathbf{x}_0$ is any solution of $A\mathbf{x} = \mathbf{0}$, then $\mathbf{x}_0$ is also a solution of $A^TA\mathbf{x} = \mathbf{0}$ since

$$A^TA\mathbf{x}_0 = A^T(A\mathbf{x}_0) = A^T\mathbf{0} = \mathbf{0}$$

Conversely, if $\mathbf{x}_0$ is any solution of $A^TA\mathbf{x} = \mathbf{0}$, then $\mathbf{x}_0$ is in the null space of A^TA and hence is orthogonal to every vector in the row space of A^TA by Theorem 3.5.6.

However, A^TA is symmetric, so $\mathbf{x}_0$ is also orthogonal to every vector in the column space of A^TA. In particular, $\mathbf{x}_0$ must be orthogonal to the vector $A^TA\mathbf{x}_0$; that is, $\mathbf{x}_0 \cdot (A^TA\mathbf{x}_0) = 0$.

It is known, from Eq. (12) below that $x_0.(A^TAx_0) = (Ax_0).(Ax_0)$.

This implies that $A\mathbf{x}_0 \cdot A\mathbf{x}_0 = 0$, so $A\mathbf{x}_0 = \mathbf{0}$ by part (*d*) of Theorem 1.2.6. This proves that $\mathbf{x}_0$ is a solution of $A\mathbf{x} = \mathbf{0}$. ■

$$A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A^T\mathbf{v} \tag{12}$$

Theorem 3.5.6 *If A is an $m \times n$ matrix, then the solution space of the homogeneous linear system $A\mathbf{x} = \mathbf{0}$ consists of all vectors in R^n that are orthogonal to every row vector of A.*

Theorem 1.2.6 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in R^n, and if k is a scalar, then:*

(*a*) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ **[Symmetry property]**

(*b*) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ **[Distributive property]**

(*c*) $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$ **[Homogeneity property]**

(*d*) $\mathbf{v} \cdot \mathbf{v} \geq 0$ *and* $\mathbf{v} \cdot \mathbf{v} = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$ **[Positivity property]**

Theorem 7.5.8 *If A is an $m \times n$ matrix, then:*

(a) *A and A^TA have the same null space.*

(b) *A and A^TA have the same row space.*

(c) *A^T and A^TA have the same column space.*

(d) *A and A^TA have the same rank.*

For the remaining parts, see that:

- in part (d), recall that A and A^TA have the same number of columns, n, and then, from this and part (a), they should have the same rank (dimension theorem);
- parts (b) and (c) follow from part (d).

The following companion to Theorem 7.5.8 follows on replacing A by A^T in that theorem and using the fact that A and A^T have the same rank for part (d).

Theorem 7.5.9 *If A is an $m \times n$ matrix, then:*

(a) *A^T and AA^T have the same null space.*

(b) *A^T and AA^T have the same row space.*

(c) *A and AA^T have the same column space.*

(d) *A and AA^T have the same rank.*

Theorem 7.5.10 *If A is an $m \times n$ matrix, then the following statements are equivalent.*

(*a*) $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(*b*) $A\mathbf{x} = \mathbf{b}$ *has at most one solution for every* $\mathbf{b}$ *in* R^m.

(*c*) *A has full column rank.*

(*d*) *A^TA is invertible.*

It suffices to prove that statements (*c*) and (*d*) are equivalent, since the remaining equivalences follow immediately from Theorem 7.5.6.

***Proof* (c) ⇔ (d)** Since A^TA is an $n \times n$ matrix, it follows from statements (*c*) and (*q*) of Theorem 7.4.4 that A^TA is invertible if and only if A^TA has rank n. However, A^TA has the same rank as A by part (*d*) of Theorem 7.5.8, so A^TA is invertible if and only if $\text{rank}(A) = n$, that is, if and only if A has full column rank. ■

Theorem 7.5.11 *If A is an $m \times n$ matrix, then the following statements are equivalent.*

(a) $A^T\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

(b) $A^T\mathbf{x} = \mathbf{b}$ *has at most one solution for every vector* $\mathbf{b}$ *in* R^n.

(c) *A has full row rank.*

(d) AA^T *is invertible.*

Theorems 7.5.10 and 7.5.11 make it possible to use results about square matrices to deduce results about matrices that are not square. For example, we know that A^TA is invertible if and only if $\det(A^TA) \neq 0$, and AA^T is invertible if and only if $\det(AA^T) \neq 0$.

Thus, it follows from Theorems 7.5.10 and 7.5.11 that A has full column rank if and only if $\det(A^TA) \neq 0$, and A has full row rank if and only if $\det(AA^T) \neq 0$.

EXAMPLE 7 A Determinant Test for Full Column Rank and Full Row Rank

We showed in Example 4 that the matrix

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

has full column rank, but not full row rank. Confirm these results by evaluating appropriate determinants.

Solution To test for full column rank we consider the matrix

$$A^TA = \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 14 & -1 \\ -1 & 2 \end{bmatrix}$$

and to test for full row rank we consider the matrix

$$AA^T = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -3 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 5 & -5 \\ -3 & -5 & 10 \end{bmatrix}$$

Since $\det(A^TA) = 27 \neq 0$ (verify), the matrix A has full column rank, and since $\det(AA^T) = 0$ (verify), the matrix A does not have full row rank. ■

APPLICATIONS OF RANK

The advent of the Internet has stimulated research on finding efficient methods for transmitting large amounts of digital data over communications lines with limited bandwidth. Digital data are commonly stored in matrix form, and many techniques for improving transmission speed use the rank of a matrix in some way.

Rank plays a role because it measures the "redundancy" in a matrix in the sense that if A is an $m \times n$ matrix of rank k, then $n - k$ of the column vectors and $m - k$ of the row vectors can be expressed in terms of k linearly independent column or row vectors.

The essential idea in many data compression schemes is to approximate the original data set by a data set with smaller rank that conveys nearly the same information, then eliminate redundant vectors in the approximating set to speed up the transmission time.